Rituximab for the treatment of rheumatoid arthritis

ERG Report

Commercial-In-Confidence information **removed**

Report commissioned by: NHS R&D HTA Programme

On behalf of: The National Institute for Health and Clinical Excellence

Produced by: *Liverpool Reviews and Implementation Group*

[redacted]

Tel: [redacted]
Fax:
Email:

Members of the review team:

[redacted]
Health Economics
The University of Liverpool Management School

[redacted]
Research Fellow, *Health Economics*
Liverpool Reviews and Implementation Group

[redacted]
Research Fellow, *Clinical Effectiveness*
Liverpool Reviews and Implementation Group

[redacted]
Director
Liverpool Reviews and Implementation Group

[redacted]
Research Fellow, *Clinical Effectiveness*
Liverpool Reviews and Implementation Group

[redacted]
Research Fellow, *Clinical Effectiveness*
Liverpool Reviews and Implementation Group

[redacted]
Research Fellow, *Clinical Effectiveness*
Liverpool Reviews and Implementation Group

[redacted]
Consultant Rheumatologist
Royal Liverpool and Broadgreen University Hospital

[redacted]
Consultant Rheumatologist
Royal Liverpool and Broadgreen University Hospital

[redacted]
Director
North West Medicines Information Centre
Pharmacy Practice Unit, Liverpool

[redacted]
Medical Statistician
Centre for Medical Statistics and Health Evaluation (CMSHE)
University of Liverpool

Correspondence to:

[redacted]

Tel: [redacted]
Fax:
Email:

Date completed: February 12th 2006

Acknowledgements:

The authors are pleased to acknowledge [redacted](Robertson Centre for Biostatistics, University of Glasgow) who provided economic advice and commented on drafts of the report and [redacted](Staffordshire Rheumatology Centre) and [redacted](The Freeman Hospital, Newcastle Upon Tyne) who provided clinical advice and commented on drafts of the report.

Contents:

Tables:

Figures:

Abbreviations:

ACR	American College of Rheumatology
BSR	British Society for Rheumatology
CRP	C-reactive protein
DAS	Disease Activity Score
DMARDs	Disease modifying anti-rheumatic drugs
EMEA	European Medicines Evaluation Agency
ERG	Evidence review group
ESR	Erythrocyte sedimentation rate
EULAR	European League Against Rheumatism
FACIT-F	Functional Assessment of Chronic Illness Therapy–Fatigue
HAQ-DI	Health Assessment Questionnaire (HAQ) Disability Index (DI)
ICER	Incremental cost effectiveness ratio
ITT	Intention to treat
K-M	Kaplan-Meier
LYG	Life years gained
MeSH	Medical subject headings
MTX	Methotrexate
NICE	National Institute for Health and Clinical Excellence
NSAIDs	Non-steroidal anti-inflammatory drugs
PSA	Probabilistic sensitivity analysis
QALY	Quality adjusted life year
RA	Rheumatoid arthritis
REFLEX	Randomized Evaluation of Long-Term Efficacy of Rituximab in RA
RF	Rheumatoid factor
RCT	Randomised controlled trial
SA	Sensitivity analysis
SF-36	Short form 36
TNFi	Tumour necrosis factor alpha inhibitor(s)

Definition of terms:

ACR20/50/70	A specified percentage improvement (20, 50, 70%) is required in the swollen and tender joint count along with improvement in three of the following: i) global disease activity assessed by observer, ii) global disease activity assessed by patient, iii) patient assessment of pain, iv)physical disability score (e.g. HAQ-DI), v) acute phase response (ESR or CRP level)
C-reactive protein	A plasma protein produced by the liver in which plasma concentrations vary in response to inflammation
Disease Activity Score 28	A continuous measure based on 28 joint evaluations and calculated using an equation that includes the tender joint count, swollen joint count, ESR and patient global assessment of general health. The system converts scores into categorical outcomes of: remission (≤2.6), low disease activity (≤ 3.2), moderate disease activity (> 3.2 ≤ 5.1) or high disease activity (>5.1)
Erythrocyte sedimentation rate	A non-specific measure of inflammation used in the diagnosis of RA in which the distance (in millimetres) that red blood cells have fallen after one hour in a vertical column of anticoagulated blood under the influence of gravity is measured
European League Against Rheumatism	The organisation which represents the patient, health professional and scientific societies of rheumatology of all the European nations and which aims to stimulate, promote, and support the research, prevention, treatment and rehabilitation of rheumatic diseases
Functional assessment of chronic illness therapy–fatigue	A 13-item questionnaire that assesses self-reported fatigue and its impact upon daily activities and function using a scale from 0–4 (0 = not at all; 1 = a little bit; 2 = somewhat; 3 = quite a bit; 4 = very much) and so the range of possible scores is 0–52, with 0 being the worst possible score and 52 the best response
Health Assessment Questionnaire Disability Index	HAQ-DI scores a patient's ability to perform daily activities from 0 (least disability) to 3 (most severe disability). In day-to-day practice, the term HAQ is often used instead of HAQ-DI
Monoclonal antibodies	Identical antibodies that are produced by one type of immune cell and are all clones of a single parent cell
Rheumatoid factor	An antibody which can bind to other antibodies and which is not normally found in the general population but is present in around 80% of adults who have RA. High levels of RF are associated with more severe RA and RF is also associated with a higher tendency to develop non-joint manifestations of RA such as rheumatoid nodules and rheumatoid lung disease
Short form 36 survey	A commonly used generic multi-purpose, short-form health survey with 36 questions yielding an 8-scale profile of functional health and well-being scores as well as psychometrically-based physical and mental health summary measures and a preference-based health utility index
Tumour necrosis factor	A pro-inflammatory cytokine that plays a central and hierarchical part in the pathogenesis of rheumatoid arthritis
Tumour necrosis factor alpha inhibitor	A biological agent designed to interrupt the inflammatory pathway of tumour necrosis factor

1 SUMMARY

1.1 Scope of the submission

The remit of the Evidence Review Group is to comment on the clinical and cost-effectiveness evidence submitted to the National Institute for Health and Clinical Excellence (NICE) as part of the single technology appraisal process. Evidence has been submitted to NICE from Roche in support of the use of rituximab for the treatment of severe rheumatoid arthritis (RA) following failure of previous therapy, including one or more tumour necrosis factor alpha inhibitor (TNFi), when compared to current standards of care.

The company submission presents a case for the use of rituximab in adult patients with severe RA. In their analysis of the decision problem, the company describes two different rituximab management strategies for patients. The first scenario is described as a "NICE recommended" strategy as it allows patients to fail on a single TNFi, receive rituximab and then go on to receive a series of disease modifying anti-rheumatic drugs (DMARDs), excluding any subsequent treatment with a TNFi. Patients are not permitted to receive a second TNFi in this scenario. The second scenario is described as a "sequential TNFi" strategy as it allows patients to fail on one TNFi, receive rituximab, and subsequently receive a second and a third TNFi before going on to receive a series of DMARD therapies. In both cases, the comparator is the same scenario without rituximab. The company has presented both scenarios to reflect their belief that although NICE does not recommend the sequential use of TNFi, there is evidence from clinical practice suggesting that a proportion of patients in the NHS in England and Wales are nonetheless receiving sequential TNFi treatment.

1.2 Summary of submitted clinical effectiveness evidence

The objective of the systematic review conducted by the company should be to identify clinical evidence to answer the questions outlined in the statement of the decision problem, However, there are no randomised controlled trials (RCTs) of rituximab versus a relevant comparator (e.g. leflunomide or a second, third TNFi). It is therefore appropriate that the company submission to details the only RCT evidence available.

The company submission provides clinical evidence from one randomised, placebo-controlled, double blind trial (REFLEX) that compares the effects of rituximab plus methotrexate (MTX), with placebo plus MTX, in a study population of 517 patients with long-standing refractory RA. Data from other RCTs are pooled to demonstrate the re-treatment efficacy of rituximab and for the analysis of safety data. Evidence from a further five trials is presented as the basis for indirect comparisons with other DMARDs.

The results from the REFLEX trial at 24 and 48 weeks confirm that rituximab plus MTX is more effective than placebo plus MTX. These findings are consistent across a range of primary and secondary outcome measures including American College of Rheumatology responses (ACR20/50/70), disease activity score (DAS28), European League Against Rheumatism (EULAR) response, Health assessment questionnaire (HAQ) disability index (DI) and radiographic scores. Given that the patients in the trial are difficult to treat, have severe disabling disease with marked impairment of quality of life, the results of the REFLEX trial are convincing for this trial population. However, whether or not the patients in the REFLEX trial are similar enough to the patients described in the rituximab management strategies put forward in the company submission is debateable, as 40% of the REFLEX trial patients had received at least two prior TNFi before receiving rituximab.

Long-term efficacy data for re-treatment with rituximab from the REFLEX trial are favourable, but the results are limited by the small number of patients available for follow-up. The estimated mean time to re-treatment from the REFLEX trial is 307 days (N=164). The available safety data from the REFLEX trial show that rituximab patients had slightly higher rates of adverse reactions than the placebo patients. The European Medicines Evaluation Agency (EMEA) particularly stresses the risks of infusion reactions and infection associated with rituximab. This mirrors the belief that patients taking any of the newer biological drugs require close surveillance and monitoring.

The only RCT evidence available for rituximab is the comparison with placebo plus MTX. It is therefore appropriate for the company to conduct indirect comparisons to calculate absolute efficacy values for use in the economic model in order to answer the questions outlined in their statement of the decision problem. However, the ERG is not confident that the adjusted ACR scores described by the company are valid. In particular, it is not clear from the evidence presented by the company that all relevant clinical studies have been included in the indirect comparison exercise. The rationale for the choice of the indirect comparison method adopted is unclear and the indirect comparison method used to adjust the ACR responses only uses a single value for the reference placebo (RP).

In summary, the ERG agrees with the company that rituximab plus MTX is more effective than placebo plus MTX in the REFLEX trial where 40% of patients had received at least two prior TNFi before receiving rituximab.

1.3 Summary of submitted cost-effectiveness evidence

The economic model submitted in support of the company submission is a micro-simulation Markov model based upon the phase III randomised control trial of rituximab plus MTX versus placebo plus MTX (REFLEX trial). Patient disease progression is tracked within the

model according to health assessment questionnaire (HAQ) score. By using micro-simulation of 10,000 RA patients, patient history is kept in memory and cost utility values are assigned to each individual at each cycle. The company conclude that rituximab is considered to be a cost-effective treatment option in RA. For the "NICE recommended" scenario, the original company submission reports an incremental cost effectiveness ratio (ICER) of £14,690 per quality adjusted life year (QALY) gained. For the "sequential TNFi" scenario, the ICER is estimated at £11,601 per QALY gained.

Early examination of the submitted economic model by the ERG identified some aspects of its implementation, which caused concern as to its reliability for generating estimates of cost-effectiveness. Two particular issues were raised with both NICE and the company concerning the method of randomisation and the representation of parameter uncertainty in the probabilistic sensitivity analysis (PSA). The company then submitted a revised model to NICE that addressed some of the ERG's concerns. The cost-effectiveness results of the revised model were very similar to the results of the original model. However, the ERG felt obliged to carry out a simple validation exercise before they could confirm that the model logic had been consistently implemented.

The ERG then identified a number of clinical and economic issues that call into question the validity of key model assumptions, and the credibility of the ICERs generated. In particular, these relate to errors in mortality rates, the evidence base for progression rates for HAQ scores, the calculation of treatment costs and errors/omissions in the estimation of in-patient costs. Some other issues were identified as potentially influencing model results, and sensitivity analyses have been carried out to show their impact on model results:

- whether the size of benefit from each treatment is overstated, because loss of efficacy is assumed to be instantaneous rather than cumulative;

- whether the assumed mean time between doses of rituximab is too conservative;

- whether the treatment sequencing in the submitted scenarios is sub-optimal.

Using alternative ERG assumptions and parameters in the model has the effect of generating substantially worsened cost-effectiveness results for the two management scenarios described by the company in their submission. The ICER for the "NICE recommended" scenario ranges from £37,002 per QALY gained to £80,198 per QALY gained and the ICER for the "sequential TNFi" scenario ranges from £28,553 per QALY gained to £65,558 per QALY gained. No patient sub-groups could be identified which exhibit significantly better economic results than the whole cohort.

The consequences of these corrections and amendments is that economic results for the use of rituximab no longer appear as unequivocally advantageous as suggested in the company submission, and may more reasonably be termed 'borderline' at best. There remain important areas where there is substantial uncertainty, which could easily invalidate economic results generated by the company model, most especially in relation to the long-term progression of disease and its effect on HAQ scores, and the duration of effective treatment for each of the active agents considered.

1.4 Commentary on the robustness of submitted evidence

The main strength of the submitted evidence is that the company makes a convincing case for the use of rituximab plus MTX versus placebo plus MTX using clinical evidence from the REFLEX trial in a specific population who are difficult to treat, have severe disabling disease with marked impairment of quality of life. However, this evidence cannot be used directly to answer the questions raised in the company's analysis of the decision problem because, in the REFLEX trial, rituximab was not compared to a relevant comparator (e.g. leflunomide or second or third TNFi).

In order to compare the management strategies using rituximab described in their analysis of the decision problem, the company carried out an indirect comparison exercise. However, given the criticisms outlined in Section 3.4.1, the ERG are not confident that the adjusted ACR responses used in the economic evaluation are wholly valid.

Finally, the ERG identified problems with the company submitted model in two stages. Early examination of the submitted economic model by the ERG identified some aspects of its implementation, which caused concern as to its reliability for generating estimates of cost-effectiveness. The company then submitted a revised model and addressed some of the ERG's concerns. However, the ERG subsequently identified a number of additional clinical and economic issues that call into question the validity of key assumptions in the revised economic model, and the credibility of the ICERs generated. In particular, the ERG question whether the size of benefit from each RA treatment is overstated, because loss of efficacy is assumed to be instantaneous rather than cumulative. This assumption merits further justification from the company. The ERG concludes that the robustness of the evidence base used in the company economic model is uncertain.

2 BACKGROUND

2.1 Introduction

The remit of the ERG is to comment on the clinical and cost-effectiveness evidence submitted to NICE as part of the single technology appraisal process. Evidence has been submitted to NICE from Roche in support of the use of rituximab for the treatment of severe RA following failure of previous therapy, including one or more TNFi, when compared to current standards of care.

Rheumatoid arthritis is a chronic, inflammatory systemic autoimmune disorder that affects the synovial joints but also many other parts of the body. If left untreated, irreversible joint damage results. Rheumatoid arthritis is associated with increased mortality.

Rituximab (known as MabThera® in the UK and Rituxan® in the USA) is a monoclonal antibody that depletes the CD20+ B cells implicated in the immunopathogenesis of RA. In July 2006 rituximab plus MTX was licensed in Europe for the treatment of severe RA following the failure of conventional treatments, including at least one TNFi. The licensing submission was supported by a phase III study[1] comparing rituximab plus MTX with placebo plus MTX along with evidence from phase II trials.[2,3] It is restricted to use by specialist physicians experienced in the diagnosis and treatment of RA.

2.2 Epidemiology

Rheumatoid arthritis affects between 0.5 and 1% of the population worldwide.[4,5] In the UK few records are kept on the numbers of people diagnosed with the disease and there are no national databases to record this information. One oft-cited research article,[6] based on the Norfolk Arthritis Register,[7] estimates the prevalence of RA in the general UK population to be 0.81% (1.16% women and 0.44% men, yielding a ratio of female:male 2.7:1). Although RA can strike at any age (with prevalence increasing with age) peak onset is said to be between the ages of 40 and 70 years.[4]

According to the British Society for Rheumatology[8](BSR) there are approximately 100 new cases of inflammatory joint disease per hundred thousand adults (16+ years) per year in the UK, of whom 24 would be diagnosed with RA. Figures relating to the number of patients in England and Wales with RA range between 344,000[9] and 426,800.[4] Of those, approximately 15% have a severe form of the disease.[4]

Disease severity ranges from self-limiting illness to chronic, progressive disease causing joint destruction and deformity. Within five years of diagnosis, a third of people with RA are unable to work.[4] Additionally, mortality is increased for people with RA when compared to

the general population[8] and total lifespan can be reduced by three to eighteen years.[10,11] Moreover, RA is a significant independent risk factor for ischaemic heart disease, with the risk related to the severity and duration of inflammation.[9,10]

2.3 Aetiology

Rheumatoid arthritis is an autoimmune disease characterised by a profound inflammatory cell infiltrate in the joints. A variety of immunological mechanisms is thought to underlie this disease including T and B lymphocytes, phagocytes, inflammatory cytokines and antibodies. The cause(s) of RA is unknown, but there is a tendency for it to run in families. If one member of a pair of identical twins has RA, then the other has a 16% chance of also developing the disease. This is higher than the 0.8% risk in the general population.[12]

Certain infections or factors in the environment might cause the immune system to attack the body's own tissues, resulting in inflammation in various organs of the body, but there seems to be no single trigger that is sufficient by itself. Smoking and obesity are known risk factors.[9] Since women seem to be more at risk than men, there is the possibility that RA is linked to female hormone production.[12] In pregnancy, RA tends to go into remission and RA is unlikely to begin at this time; however, post-delivery poses a higher risk of relapse and development of RA.[12] It has also been suggested that use of the oral contraceptive pill may be largely responsible for reducing the occurrence of RA in younger women during the last 30 to 40 years.[12]

Rheumatoid factor status

Rheumatoid factor (RF) is an antibody not usually found in the blood of the general population and which may be tested for in patients with suspected RA. Approximately 80% of RA patients are RF positive. Results of an RF test cannot be used to diagnose the disease since: i) not all patients with RA are RF positive and ii) the presence of RF may be due to other factors. Results are interpreted within the context of other signs and symptoms. High levels of RF (generally above 20 IU/mL, 1:40 or over the 95th percentile) are indicative of RA. The higher the levels of RF, the greater the possibility of a more destructive articular disease. Since prediction of the persistent cases (those that will suffer joint damage) is the key to successful treatment of early RA,[8] RF may be one means of selecting patients for more aggressive therapy.

2.4 Clinical features of RA

Rheumatoid arthritis is a chronic, inflammatory systemic autoimmune disorder that affects the synovial joints. These joints become inflamed leading to progressive erosion of cartilage and

bone. Affected joints are painful, swollen and stiff, particularly first thing in the morning or after periods of inactivity. If left untreated, irreversible joint damage results.

Along with swollen and tender joints, patients with RA may experience other extra articular complaints, including fatigue, anaemia, weight loss, general malaise, lymph node enlargement, lung diseases, pericarditis, vasculitis, skin nodules and eye diseases.[4,10]

The course of the disease is highly unpredictable: disease activity may be continuous, or there may be periods of partial or complete remission.

2.5 Diagnosis and assessment

A diagnosis of RA can be difficult to establish as there are no definitive tests that differentiate between RA and other types of inflammatory polyarthritis. However, since early treatment is known to arrest disease progress (and thereby limit joint damage) it is crucial that diagnosis is made without delay. Current BSR guidelines[8] advocate that a diagnosis of RA should be made as early as possible on the basis of persistent joint inflammation affecting at least three joint areas, involvement of the metacarpophalangeal or metatarsophalangeal joints or early morning stiffness of at least 30 minutes duration.

Measures to assess disease activity and response to treatment include the set of ACR responses, DAS (DAS28) and EULAR response. However, these measures are not routinely utilised in clinical practice.

The ACR response criteria allow measurement of response to treatment according to three levels of improvement, 20%, 50% or 70%. To achieve an ACR20, a 20% improvement from baseline is required in the swollen and tender joint count along with improvement in three of the following:

- Global disease activity assessed by physician
- Global disease activity assessed by patient
- Patient assessment of pain
- Physical disability score (e.g. HAQ-DI)
- Acute phase response

Likewise, to achieve an ACR50 or ARC70, the swollen and tender joint count plus three of the above must have improved by 50% or 70%.

The BSR[8] advocates that clinicians aim for a patient DAS28 score of ≤ 2.6 or at least ≤ 3.2. and further advise[13] that an adequate response to treatment can be characterised by a decrease in DAS28 of 1.2 (or greater). DAS scores are utilised within the EULAR response criteria to classify patients as good, moderate or non-responders to treatment.

The following tools are often used alongside the ACR20/50/70 assessment: health assessment questionnaire (HAQ) disability index (DI)**,** short form 36 (SF-36), FACIT-F and radiographic scores.

2.6 Current treatment options

Rheumatoid arthritis is not curable and therefore the main aims of treatment are to control the symptoms and signs of disease, maintain function and promote self-efficacy.[8] Irreversible damage occurs within the first two years of the disease; consequently, early diagnosis and treatment along with 'tight control' of disease activity are stressed.[5,8] Disease progression and outcomes are variable within and between individuals hence long-term planning and regular clinical evaluation are essential to the management approach.

2.6.1 Clinical guidance

Two main types of drug therapy may be offered to patients with RA: those that provide relief from symptoms without modifying the disease process, and DMARDs. There are also surgical options available to patients with severe joint damage; these include joint replacement, tendon construction and synovectomy.

Drugs, which provide symptom relief, include non-steroidal anti-inflammatories (NSAIDs) and analgesics. Patients may be managed exclusively with varying combinations of these, but more often, symptom-relieving agents are used as an adjunct to DMARDs.

Drugs classified as DMARDs fall into two distinct categories and these are offered at different points in the treatment pathway. Conventional DMARDs include azathioprine, leflunomide, hydroxychloroquine, sulfasalazine, injectable gold and MTX, with the latter generally regarded as the treatment of choice.[14] Many rheumatologists opt for MTX as the initial DMARD. Methotrexate has a favourable efficacy and toxicity profile, low cost, and established track record in the treatment of RA and has become the standard by which new DMARDs are evaluated.[14] Conventional DMARDs may be used alone or in combination.

The second type of DMARD is the relatively new group of biologics (e.g. agents that target the action of TNF-alpha). These include adalimumab (Humira®; Abbott Laboratories), etanercept (Enbrel®; Wyeth Laboratories) and infliximab (Remicade®; Schering-Plough), all of which are TNFi.

In RA, current NICE guidance[15] states that one TNFi can be prescribed when disease is active and severe and has not responded to treatment with at least two conventional DMARDs including MTX (unless there is intolerance to MTX). A trial period for a DMARD is defined as six months, with at least two months at the standard dose.

The guidance[15] goes on to advise that treatment with TNFi should be continued only if there is an adequate response at six months following initiation of therapy. Different TNFi should not be used sequentially (unless treatment is withdrawn because of an adverse event); thus after failure on one TNFi, a patient should receive standard care. In practice, however, it may be the case that patients are offered an alternative TNFi after failing to respond to an initial course (Tom Kennedy and Robert Moots, personal communication, December 2006). In fact, the BSR Biologics register newsletter confirms that 10% of patients on the BSR register (approximately 1000) had switched to a second TNFi.[16]

In Europe, rituximab is the only licensed drug for RA patients who fail to respond to DMARDs and then fail on a TNFi.

Rituximab was reviewed by the Scottish Medicines Consortium and was accepted in November 2006 for restricted use within NHS Scotland in combination with MTX. Outside of Europe, rituximab has regulatory approval in 28 countries (company submission p.4).

2.6.2 Number of patients treated

NICE estimates that there are approximately 400,000 people with RA in England and Wales.[15] The company submission estimates that there are approximately 363,000 patients with RA in England and Wales. It is estimated that 10-15% of RA patients have a severe form of the disease.[17,15]

In the NICE appraisal of etanercept and infliximab[18] it was estimated that 15,000 people with RA are eligible for TNFi therapy with an additional 950 becoming eligible each year. If we assume that between 15%[10] and 50%[17] of patients fail on TNFi, then this means that between 2,250 and 7,500 RA patients would be potentially eligible to receive rituximab after having failed a TNFi.

2.7 Critique of company background

The company submission provides an accurate and thorough discussion of the background to the disease of RA and its treatments.

2.7.1 Choice of comparator

In the clinical section of the company submission, it is acknowledged that the literature search conducted did not identify any RCTs that compared rituximab directly with the appropriate comparators (second or third TNFi, or leflunomide). In Europe, there is no other drug licensed for use in patients who have severe active RA and who have failed previous treatments (including at least one TNFi) and so valid direct comparisons were not available to inform the clinical and cost-effectiveness questions raised in the company's decision statement. The

company's clinical evidence focuses on the comparison of rituximab plus MTX versus placebo plus MTX. Indirect comparisons of rituximab with other DMARDs and TNFi are presented.

The company submission offers two different RA management strategies. The first strategy adheres to current NICE guidance,[15] whereas the second falls outside current guidance. However, the latter scenario does reflect the final scope issued by NICE and was included in the company's decision statement. Both scenarios begin with the assumption that all patients have failed one TNFi (etanercept, currently the most widely-used TNFi in England and Wales[19]).

In scenario 1, after patients have failed on etanercept, they either receive rituximab plus MTX or leflunomide plus MTX[1]. With the exception of rituximab, treatment options are the same for all patients: leflunomide plus MTX, gold, ciclosporin and finally MTX. This comparator option is described as "NICE recommended" as patients only ever receive one TNFi and after they fail, they return to using conventional DMARDs alone.

In scenario 2, after patients have failed on etanercept, they receive either rituximab plus MTX or adalimumab plus MTX. With the exception of rituximab, treatment options are the same for all patients: adalimumab plus MTX, infliximab plus MTX, leflunomide plus MTX, gold, ciclosporin and finally MTX. This comparator option is described as "sequential use of TNFi" as more than one TNFi is tried by patients before they return to using conventional DMARDs alone.

Given the uncertainty surrounding the many treatment pathways for patients with RA, it is unlikely that a consensus of opinion from medical experts can be reached regarding (i) choice of drugs in the sequence or (ii) order of drugs in the sequence. The appropriateness of the treatment pathways considered by the company in both scenarios may therefore be subject to debate within the medical community.

---

[1] The ERG note that the current summary of product characteristics for leflunomide states that "…recent or concurrent treatment with hepatotoxic or haematotoxic DMARDs (e.g. methotrexate) may result in an increased risk of serious adverse reactions; therefore, the initiation of leflunomide treatment has to be carefully considered regarding these benefit/risk aspects".[20] electronic Medicines Consortium. Arava 10, 20 and 100mg Tablets. 21st March 2006 [cited 6th February 2006]; Available from: http://emc.medicines.org.uk/emc/assets/c/html/displaydoc.asp?documentid=7480#PRODUCTINFO

3 CLINICAL EFFECTIVENESS

3.1 Critique of systematic clinical review

The company submission includes a systematic review of rituximab for the treatment of RA; both direct and indirect comparisons of clinical trial evidence were carried out. Firstly, we provide a critique of the systematic review methods used and then go on to critique the clinical trial evidence and data analysis presented by the company.

Key aspects of the methodological quality of the company's review of the clinical literature were assessed based on an accepted quality assessment tool[21] and the results are summarised in Table 3-1.

Table 3-1: Quality assessment of the clinical effectiveness review

Quality assessment checklist item	Yes/No
Did the review address a clearly focused research question?	✓
Was the search strategy adequate? (i.e. did the reviewers identify all relevant studies?)	✓/✗
Are the inclusion/exclusion criteria specified?	✓/✗
Did the review include the right type of studies?	✓
Is there a statement of completeness from the company?	✗
Did the reviewers assess the quality of the included studies?	✓/✗
Was the method of data extraction reported?	✗
Were appropriate measures of outcomes used?	✓
If the results of the studies have been combined, was it reasonable to do so?	N/A
Are appropriate sub-group analyses presented?	✓
Are the main results of the review reported? (e.g. numerical results included with the CIs)	✓/✗
Are issues of generalisability addressed?	✓

✓=yes, ✓/✗=partially, ✗=no, N/A= not applicable

3.1.1 Search strategy

Direct comparison

Three electronic databases were searched (Medline, EMBASE and the Cochrane Library) covering the period 1993 to 17th October 2006. In addition, two sets of conference abstracts, EULAR (annual meetings 2002-2005) and ACR (annual meetings 2002-2006), were searched via the following websites: http://www.eular.org and http://www.rheumatology.org respectively.

Search terms for electronic databases appropriately included a combination of free-text and index terms (rheumatoid arthritis) combined with drug name (rituximab) used as free-text

terms. However, the search strategy details do not include any information on the subject index headings (for example, MeSH) and the relationship between the search terms (for example, Boolean), details of any additional searches (e.g. searches of company databases). We were therefore unable to reproduce these searches. However, the ERG is confident that all relevant published trials were identified by the company.

Indirect comparisons

No information about the search strategy used to identify papers for inclusion in the indirect comparison analysis was originally included in the company submission. However, on request the company did provide this information.

Searches were conducted focussing on (i) treatment failure and (ii) RCTs of anti-rheumatic drugs that satisfied the following criteria: ACR response criteria as an outcome measure; adult RA patients and trial duration of at least six months. The first search was carried out in Medline (1996 to March Week 4 2005) while the second was carried out in Medline (1966 to March Week 4 2005) and EMBASE (1988 to 2005 Week 21) with additional searches of the Cochrane Library and NICE HTA reports.

For both searches, search strategy details include information on the subject index headings (for example, MeSH) and the relationship between the search terms (for example, Boolean). However, the search was conducted in April 2005 and has not been updated to inform the current submission (November 2006).

Regarding the first search, the company recognise and discuss the difficulties in targeting a search to the concept of treatment failure due to limited MeSH terms and reliance on text string searching.

Regarding the second search which is limited to RCTs, recognised filters which have been tested, validated and proven to be effective in systematically retrieving RCTs[22,23] were not utilised.

3.1.2 Inclusion and exclusion criteria

Direct comparison

Details of inclusion and exclusion criteria are provided in Table 3-2 and are considered appropriate and complete.

Table 3-2: Scope of the literature review for direct comparisons

	Clinical effectiveness
Population	Adults with severe active rheumatoid arthritis who have had an inadequate response to, or are intolerant of, other DMARDs including one or more TNFi agent
Intervention	Rituximab plus MTX
Comparators	Not specified
Outcomes	RA efficacy outcomes
Study design	Clinical trial data
Inclusion criteria	• Clinical trial data from controlled studies of human populations in which rituximab, RA and RA efficacy endpoints were the main focus of the paper • Patient population should consist of those patients who had an inadequate response to, or are intolerant of, one or more TNFi
Exclusion criteria	• Papers covering the use of rituximab in other autoimmune diseases, non-Hodgkin's Lymphoma or other haematological malignancies • Papers providing a review, update or commentary on data published elsewhere • Animal studies or in vitro research

Indirect comparisons

Information about the inclusion and exclusion criteria used for the indirect comparisons were not included in the original company submission. However, on request, the company did provide broad reasons as to why papers were excluded from the treatment failure search. For the ACR response criteria search, the company provided the inclusion criteria and broad reasons why studies were included and excluded (Table 3-3). The criteria described appear to be appropriate.

However, from a methodological perspective, the exclusion of trials not reporting ACR responses merits discussion. Whilst this would seem reasonable when looking to extract ACR20 responses, there is a large body of evidence to suggest that outcome reporting bias can mean that non-significant outcomes are less likely to be reported in a paper even if they have been measured.[24] Therefore, to exclude studies (N=32) that do not report ACR responses might result in excluding trials where these were measured but not reported.

Table 3-3: Scope of the literature review for indirect comparisons

	Clinical effectiveness	
	Treatment failure	**ACR response criteria**
Population	Adults with severe active RA	Adults with severe active RA
Intervention	DMARDs following TNFi	DMARDs and/or TNFi
Comparators	Not specified	Not specified
Outcomes	RA efficacy outcomes	RA efficacy outcomes
Study design	Clinical trial data	Clinical trial data
Inclusion criteria	• Use of DMARDs following treatment failure with TNFi	• DMARD trials • Biological trials
Exclusion criteria	• Studies comparing biological agents to (an)other biological agent(s) • Wrong indication or focus (includes case studies) • Studies with no post-biological treatment discussed (includes post-marketing studies, reviews, HTA reports) but with a clinical focus • Trials with the wrong outcomes (include pharmacokinetic, safety studies etc)	• Study design (not an RCT, no comparator, open label, pharmacokinetic study etc) • Wrong indication (not adult RA) • Wrong intervention (no drug intervention or wrong drug class) • Secondary analysis or duplicate publication • Wrong outcomes (no ACR response criteria) • Disease or treatment duration too short

3.1.3 Application of inclusion criteria

Direct comparison

A flow diagram in the company submission indicates that for direct comparisons, 69 citations were identified by the electronic database search (of which 23 were excluded on the basis of title) and 56 conference abstracts were also identified. One citation[1] and six abstracts[25-30] describing the phase III, double-blind REFLEX trial (WA17042) and open-label extension to the REFLEX trial (WA17531) met the inclusion criteria set out in the company report.

For the long-term efficacy analyses and safety analyses, the company included data from two phase II RCTs (WA17043 and WA16291); no explanation was given for including additional data sources in their analyses. Both of these trials were appropriately excluded from the systematic review because they included patients who had no prior exposure to a TNFi and who had received unlicensed doses of rituximab.

Indirect comparisons

Information provided by the company, upon request, shows the number of results found by each search term and the number of citations excluded by the relevant exclusion criteria; 99/99 citations were excluded for treatment failure leaving no relevant articles; 264 citations were found for ACR response criteria of which 44 were included in the review. However, only six of these studies were used for calculating specific treatment adjusted ACR response rates. The criteria used to select these six studies were not stated. In addition, the company did not provide a flow diagram or present details of the excluded studies or the reasons for their exclusion.

3.1.4 Quality assessment

Direct comparison

The company submission did not include a formal quality assessment but did partly discuss the methodological limitations of the one included trial as specified by NICE.[31] The dates of recruitment and flow diagrams of participants through each stage were unclear. Upon request, the company provided further information on the flow of patients. However, this additional data did not fully explain the flow of patients analysed in the repeat treatment analyses.

The company submission states that this was a blinded study, in which the study sponsor, investigators, and patients were all unaware of the patient's trial arm. A dual assessor approach was employed in which an efficacy assessor only had access to efficacy data while a safety assessor "had access to all clinical and laboratory (safety) data and was able to make any necessary changes to the patient's medical therapy, thus minimizing the chance of unblinding of the efficacy assessor who only had access to efficacy data"(company submission p.34). Radiographic assessments were collected and scored by two independent readers blinded to treatment assignment and time point.

It is noted that some patients did become unblinded due to vial breakage and it is also stated in the peer reviewed journal article[1] that the blinding of the efficacy assessor was potentially compromised at one of the recruiting centres. While these patients were subsequently excluded from the 24-week intention to treat (ITT) analyses, and a sensitivity analysis demonstrated no change in the significance of the primary results, this issue should be considered when interpreting long-term results that did include these patients.

The quality of data reporting in the company submission was poor: no confidence intervals were presented for any of the results and presentation of p-values was inconsistent.

Indirect comparisons

The company submission did not provide any quality assessment of the studies included in the indirect comparison analysis. However, it did not acknowledge that the trials included in the indirect comparison analysis (excluding rituximab and abatacept trials) were made up of patients who were from less severe RA populations and not necessarily comparable to the patient population of interest.

3.1.5 Combination of studies

The company submission states (p.59) "due to the study selection a meta-analysis was not required". However, analyses of the efficacy of repeated treatments and adverse events do include pooled analyses.

In the analysis of long-term efficacy of repeated treatments, clinical data on 279 patients are analysed in the company submission. However, the actual number of patients from individual trials (WA17043 and WA16291) in the analysis is unclear.

In the analysis of adverse reactions (N=938), data from two Phase II studies are pooled and presented alongside data from the REFLEX trial. In all other safety analyses, an all exposure population (N=1039), from phase II and phase III trials, is described.

3.2 Direct comparison: rituximab plus MTX versus placebo plus MTX

Data presented in this report on the direct comparison have been extracted from both the company submission and the primary published, peer-reviewed clinical paper.[1] Additional information was provided by the company in clarification of questions raised by the ERG. P-values are stated where reported in the company submission and or clinical paper.[1]

One parallel group, multinational, multicentre, randomised, placebo-controlled, double-blind study (REFLEX) involving 517 patients is included in the direct comparison (rituximab 309: placebo 208). Patients were randomly assigned in a 3:2 ratio and given rituximab, at the licensed dose of 1000mg, or placebo, on day 1 and then again on day 15. The placebo concentrate was diluted and administered exactly as for the active rituximab concentrate. Patients in both arms also received MTX and glucocorticoids.

From week 16, patients were allowed to exit the trial or receive rescue treatment for reasons of treatment failure (Figure 3-1). Rescue treatment for those receiving rituximab constituted "standard care" (which was determined by individual investigators at their discretion and so could not be standardised but may have included a return to one or more DMARDs or a return

to TNFi therapy) whilst those in the placebo group were eligible to enter the open-label arm of the REFLEX trial where they received rituximab plus MTX. Only 54% of the patients randomized to placebo and 82% of the patients randomized to rituximab completed the full 24 weeks of the treatment period.

The REFLEX trial is a complex, non-standard randomised controlled trial; the design of the trial may therefore attract criticism. For example, patients were not allowed to exit the trial before week 16. This means that valuable information regarding the natural timing and reasons for withdrawal are unavailable for this period. In addition, it would have been more informative if the cross-over of patients had commenced at week 24 (time of the primary efficacy analysis) instead of week 16.

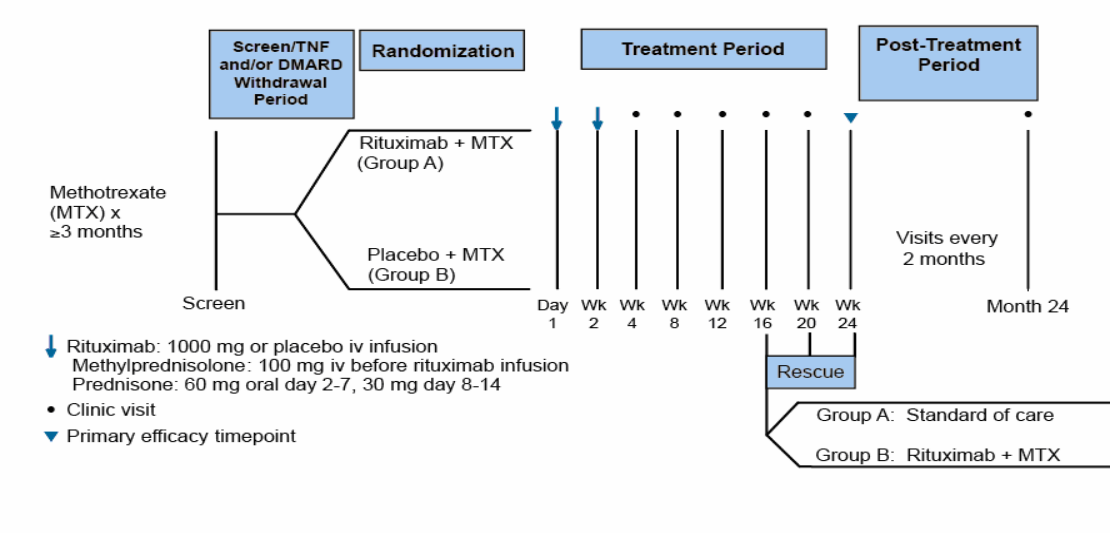


Figure 3-1: Trial flow diagram (company submission p.27)

Table 3-4: Number of patients in each of the analyses

	Placebo+MTX	Rituximab+MTX	Total
Patients enrolled and randomised	n=209	n=311	n=520
Patents treated	n=208	n=309	n=517
A. Primary (ITT) analysis	n=201	n=298	n=499
➢ Sub-group (ITT) analyses	n=201	n=298	n=499
B. Long-term efficacy after one course of rituximab	n=209	n=308	n=517
C. Radiographic endpoints after 56 weeks	n=186	n=277	n=463
D. Long-term efficacy following repeated courses	n=117*	n=164	n=281**
E. Analysis of adverse events ❖ Analysis of adverse reactions ❖ Analysis of acute infusion reactions ❖ Analysis of infections	n=398	n=540	n=938 n=1039*** n=1039***

* Includes 45 placebo patients who entered via "rescue therapy"

** Despite clarification from the company, the numbers included in the long-term efficacy following repeated measures are still unclear. The quoted figure of 281 does not appear to correspond with the 279 used in the analyses.

*** n=1039 for first course, n=570 for second course, n=191 for third course and n=40 for fourth course

The numbers of patients included in each of the analyses undertaken are shown in Table 3-4. The flow diagrams requested from the company are provided in Appendix 1.

3.2.1 Rescue therapy

Eighty patients in the placebo group were allowed to exit the trial and enter into rescue therapy (receive rituximab) between weeks 16 and 24. These patients were included in the ITT population as 'non-responders'. The company submission (p.74) stated, "due to the mechanism of action, non response to rituximab can only really be determined from 4 months". Time to respond to MTX is not explicitly stated in the company submission. In the published literature it is assumed to be approximately three[9] to six[32,33] months. As the patients in the REFLEX trial had been on MTX for at least 12 weeks prior to the start of the trial, it is likely that patients in the placebo arm would have had sufficient time to respond to MTX before being considered as non-responders.

3.2.2 Trial characteristics

Trial characteristics are summarised in Table 3-5.

The trial was conducted in 114 rheumatology centres in the US, Europe, Canada, and Israel. As the company submission acknowledges, there are differences in clinical practice between the US and other countries. For this reason, randomisation was stratified by region. However, even within regions it is recognised that differences in clinical practice will exist.[8] Randomisation was also stratified by RF status.

It was stated in the company submission that patients in the trial should have had no more than five DMARDs (and this is also reproduced in the EMEAreport[34]) but later in the company submission it is reported that some patients had received as many as eight or nine previous DMARDs. Later clarification from the company states that the reference to no more than five DMARDs was incorrectly inserted and the inclusion exclusion criteria did not state a maximum number of previous DMARDs. Certainly, in clinical practice there would appear to be no reason to exclude patients who had had more than five previous DMARDs.

Table 3-5: Study characteristics

Study Name	Interventions drug & dose, N	Study enrolment	Study Design	Outcomes	Location & centres	Inclusion criteria	Exclusion criteria	Follow-up
REFLEX (WA17042) REFLEX open label extension (WA17531)	Rituximab + MTX (N=309) Placebo + MTX (N=208)	Dates not given	RCT Phase III	Primary outcome: • At least a 20% improvement in the swollen and tender joint counts and at least three of five other disease activity measures Secondary outcomes: • At least a 50% and 70% improvement in the swollen and tender joint counts and at least three of five other disease activity measures • Changes in disease activity • Assessment of physical function • Quality of life measurements • Fatigue measurements • Radiographic assessment of joint damage	International (US, Europe, Canada, and Israel), multi-centre (114 rheumatology centres)	Patients must have had: • Rheumatoid arthritis for at least 6 months, according to the American College of Rheumatology 1987 revised criteria • An inadequate response to, or intolerance of, at least one previous or current anti-TNF therapy • MTX (10–25mg/week) for at least 12 weeks prior to screening, with the last 4 weeks at a stable dosage • Active disease, defined as ≥ 8 swollen joints (of 66 joints assessed) and ≥ 8 tender joints (of 68 joints assessed) • C-reactive protein (CRP) level ≥1.5 mg/dl or an erythrocyte sedimentation rate (ESR) ≥28 mm/hour • Radiographic evidence of ≥1 joint with a definite erosion attributable to RA, as determined by a central reading site (a centralised organisation independent of the sponsors that provided blinded radiographic assessments)	Patients must not have: • A history of a rheumatic autoimmune disease other than RA (except secondary Sjögren's syndrome • Significant systemic involvement secondary to RA (vasculitis, pulmonary fibrosis, or Felty's syndrome) • ACR functional class IV disease	Primary analysis was carried out after 24 weeks follow-up Secondary analysis was carried out after 48 weeks follow-up (56 weeks with regard to radiographic assessment)

3.2.3 Participant characteristics

Information relating to participant characteristics was reported in both the company submission and in the published paper. The participants in the two treatment groups appear to be comparable.

The patients enrolled would appear to match the population of interest as all patients had severe RA (as defined by a range of measures). The average duration of disease was 12 years meaning that most patients had experienced RA for a considerable length of time. Another baseline finding to note is that in the REFLEX trial there are around four times as many females as males. Published literature reports there are around three times as many females as males with RA.[6]

All of the patients in the trial had received at least one TNFi in the past. However, 40% of patients had received at least two TNFi; this means that 40% of REFLEX trial patients do not match the target patients described in the NICE "recommended" scenario which does not allow the use of sequential TNFi. Also, the most commonly used TNFi cited by the patients in the REFLEX trial is infliximab, whereas the most commonly used TNFi to treat RA in Europe, and also discussed in the company base case scenarios, is etanercept.[19]

3.2.4 Comparator

In the REFLEX trial rituximab plus MTX is compared to placebo plus MTX. This comparison does not reflect any of the treatment options described in the company's decision statement (scenario 1 or scenario 2) or the alternatives pursued in the economic evaluation. It is noted that a trial comparing rituximab against a combination of DMARDs would be a challenge in terms of identifying what combination is likely to be successful in patients who have failed a TNFi (Dr Andrew Hassell, personal communication, February 2006). Also, current published literature[15] suggests that increasing evidence is accumulating to support the trial of a second TNFi in patients who fail on one TNFi.

3.3 Clinical results

Table 3-6 shows the key results of the REFLEX trial at 24 weeks for the ITT population. It is noted that the ITT population described by the company excludes 21 randomised patients (13 from the rituximab arm, 8 from the placebo arm). The company submission focuses on appropriate efficacy outcomes.

Table 3-6: Key results from REFLEX trial

Outcome	Placebo (N=201)	Rituximab (N=298)
Primary		
ACR20(%)	18%	51%
Secondary		
ACR50(%)	5%	27%
ACR70(%)	1%	12%
Change in DAS [mean (SD)]	-0.4 (1.17)	-1.9 (1.6)
EULAR response (%)		
None	78%	35%
Moderate	20%	50%
Good	2%	15%
Change in ACR core set [mean (SD)]		
Swollen joint count	-2.6 (10.35)	-10.4 (12.95)
Tender joint count	-2.7 (15.48)	-14.4 (17.48)
Patient global assessment	-5.3 (22.88)	-26.0 (29.56)
Physician global assessment	-6.2 (27.70)	-29.5 (27.40)
Health assessment questionnaire[1]	-0.1 (0.45)	-0.4 (0.60)
Pain assessment	-2.5 (23.30)	-23.4 (29.35)
CRP (mg/dL)	0.0 (3.59)	-2.1 (3.48)
ESR (mm/h)	-4.1 (25.05)	-18.5 (22.56)
Change in SF36 domains) [mean (SD)]		
Mental health[2]	1.3 (9.43)	4.7 (11.75)
Physical health[3]	0.9 (5.65)	5.8 (8.47)
Changes in FACIT-F [4] [mean (SD)]	-0.5 (9.84)	-9.1 (11.3)

For SF-36 a positive change = improvement. For all other continuous variables a negative change = improvement.

1: clinically relevant improvement = decrease> 0.22

2: clinically relevant improvement = increase>6.33

3: clinically relevant improvement = increase >5.42

4: clinically relevant improvement = decrease >4

3.3.1 Primary efficacy outcome: ACR20

Table 3-6 shows that at week 24, 51% of patients receiving rituximab had achieved an ACR20 response compared to 18% of patients in the placebo group. Supporting ACR20 data presented in the published paper[1] showed a statistically significant separation between rituximab and placebo treatment at week eight. This outcome is clinically significant for this group of patients, who have severe, difficult to control disease, and have failed a TNFi.

3.3.2 Secondary efficacy outcomes

All secondary efficacy outcomes, including ACR50 and ACR70 responses, were significantly different between the two groups ($p<0.002$) in favour of rituximab. Two of the quality of life outcomes reported are of particular interest. Whilst 65% of patients treated with rituximab showed a moderate or good improvement in EULAR scores, 35% failed to show any improvement. Similarly, DAS28 scores fell by 1.9 in the rituximab group.

Both the erythrocyte sedimentation rates (ESRs) and the C-reactive protein (CRP) levels in the rituximab groups fell close to normal levels, whilst those in the placebo arm remained high. Baseline scores were not provided for the SF-36 or FACIT-F.

3.3.3 Radiographic endpoints

Radiographic endpoints are reported in the published paper[1] and in the submission (p.48). At 24 weeks, only the score for joint space narrowing was significantly different between the treatment and placebo arms. The short time frame was not sufficient for significant effects to be observed and so radiographic data in the submission are fully reported at 56 weeks (see Section 3.3.5).

3.3.4 Sub-group analyses

Three sub-group analyses were carried out to assess the impact of rituximab versus placebo on (i) patients who had failed only one prior TNFi compared to patients who had failed two or more (ii) US and non-US patients (as US patients tend to receive biologics earlier) (iii) RF positive and RF negative patients.

Table 3-7 shows that there was a greater response to both rituximab and placebo when patients had received only one TNFi prior to treatment. This finding could indicate that it is more effective to use rituximab after only one TNFi, or that patients who have received more than one TNFi have RA that is resistant to treatment. Analysis of variance on ACR20 responses by the ERG shows that both treatment and number of prior TNFi are independent predictors of response, but there is no significant interaction term (p=0.28). Nonetheless, the significant difference attributable to number of prior TNFi suggests that it is prudent to consider the sensitivity of economic results to this factor (see Section 4.3.6).

The ACR20 response rates for rituximab patients were different according to geographical region; the company submission reports the difference to be significant at the 10% level (p=0.08). Nonetheless, the effects of rituximab versus placebo were significantly different for both regions.

Reported ACR20 response rates were significantly different (rituximab versus placebo) for both RF positive and RF negative patients. Rituximab seems to be more effective in RF positive patients (Table 3-7).

Table 3-7: ACR20 responses: number prior TNFi, US versus non-US, RF status

		Placebo (N=201)	Rituximab (N=298)	p-value
1 prior TNFi*		(n=121)	(n=179)	
	ACR20	21%	58%	<0.0001
	ACR50	7%	30%	<0.0001
	ACR70	1%	14%	<0.0001
≥2 prior TNFi*		(n=80)	(n=119)	
	ACR20	14%	42%	<0.0001
	ACR50	3%	22%	<0.0001
	ACR70	3%	10%	<0.0403
US no.(%)		[redacted]	[redacted]	[redacted]
	ACR20	[redacted]	[redacted]	[redacted]
	ACR50	[redacted]	[redacted]	[redacted]
	ACR70	[redacted]	[redacted]	[redacted]
Non-US no.(%)		[redacted]	[redacted]	[redacted]
	ACR20	[redacted]	[redacted]	[redacted]
	ACR50	[redacted]	[redacted]	[redacted]
	ACR70	[redacted]	[redacted]	[redacted]
RF positive		n=160	n=234	
	ACR20	31 (19%)	127 (54%)	<0.0001
	ACR50	9 (6%)	69 (29%)	<0.0001
	ACR70	3 (2%)	31 (13%)	<0.0001
RF negative		n=41	n=64	
	ACR20	5 (12%)	26 (41%)	0.0009
	ACR50	2 (5%)	11 (17%)	0.0739
	ACR70	0 (0%)	6 (9%)	0.0450

* p-values calculated by the ERG

3.3.5 Outcomes at 24 and 48 weeks

This section characterises the long-term effects of a single course of rituximab compared with placebo and shows results at 24 and 48 weeks for patients. For all outcomes (excluding DAS28) there are two sets of analyses: one for patients who received only the treatment they were randomised to (observed) and all patients regardless of treatment received (non-responder imputation (NRI)). The NRI data also included the 21 patients excluded from the ITT population. As the number of patients in the observed category is small, especially for 48 week data, discussion is focussed on NRI values rather than observed values.

The company submission (p.44) reports that 37% (114/308) of patients in the rituximab group remained in follow-up at week 48 compared to 11% (24/209) in the placebo arm. This indicates that a substantial proportion of patients continued to receive clinical benefit almost a year after a single course of rituximab.

Primary efficacy outcome: ACR20 response

The NRI data in Table 3-8 verify that at 24 and 48 weeks there are a larger number of patients in the rituximab group compared with the placebo group showing an ACR20 response after

just one treatment. At week 48, there has been a reduction in the number of patients achieving ACR20 in both groups of patients.

Table 3-8: Proportion of patients with ACR responses at weeks 24 and 48

		Placebo (N=209)				**Rituximab (N= 308)**			
		Observed data (N=112)		NRI (N=209)		Observed data (N=254)		NRI (N=308)	
		No.	%	No.	%	No.	%	No.	%
Week 24	ACR20	36	32	36	17	159	63	159	52
	ACR50	11	10	11	5	82	32	82	27
	ACR70	3	30	3	1	37	15	37	12
		n=24		N=209		n=114		n=308	
Week 48	ACR20	8	33	8	4	58	51	58	19
	ACR50	2	8	2	1	39	34	39	13
	ACR70	1	4	1	<1	16	14	16	5

Secondary efficacy outcomes

There are reductions in the number of patients achieving ACR50 and ACR70 responses over time, but differences between groups remain (Table 3-8).

Data at 48 weeks show that both groups have clinically relevant reductions in DAS28 (greater than 1.2). NRI results are not reported.

At 48 weeks, the number of patients receiving clinically significant improvement as measured by EULAR and HAQ-DI responses remained higher in the rituximab group than in the placebo group. Remission data show that 5% of the rituximab group achieved remission compared with only 1% of the placebo group (Table 3-9).

Table 3-9: Mean change in DAS28, EULAR, HAQ-DI at weeks 24 and 48

	Time point		Placebo (N=209)				Rituximab (N= 308)			
			Observed		NRI		Observed		NRI	
DAS28			n=112				n=254			
	Week 24	Mean (SD)	5.8	(1.38)			4.6	(1.49)		
		Mean change (SD)	-0.9	(1.28)			-2.3	(1.47)		
			n=24				n=114			
	Week 48	Mean	5.1	1.39			4.5	1.47		
		Mean change	-1.41	1.38			-2.1	1.72		
			No	**%**	**No**	**%**	**No**	**%**	**No**	**%**
EULAR	Week 24		n=112		n=209		n=254		n=308	
		Moderate	41	37	41	20	154	61	154	50
		Good response	4	4	4	2	46	18	46	15
		Low disease activity	4	4	4	2	47	19	47	15
		Remission	2	2	2	1	27	11	27	9
			n=24		n=209		n=114		n=308	
	Week 48	Moderate	13	54	13	6	54	47	54	18
		Good response	2	8	2	1	26	23	26	8
		Low disease activity	2	8	2	1	27	24	27	9
		Remission	1	4	1	<1	14	12	14	5
Change in HAQ-DI			n=111		n=209		n=254		n=308	
	Week 24	>=0.22	48	43	48	23	179	71	179	58
		>=0.3	41	37	41	20	149	59	149	48
		>=0.5	29	26	29	14	123	48	123	40
		>=0.8	9	8	9	4	62	24	62	20
			n=23		n=209		n=113		n=308	
	Week 48	>=0.22	13	57	13	6	70	62	70	23
		>=0.3	11	48	11	5	62	55	62	20
		>=0.5	7	30	7	3	47	42	47	15
		>=0.8	3	13	3	1	23	20	23	7

Radiographic endpoints

Radiographic information was available for patients in the rituximab and placebo groups for whom baseline and 56 week radiographs were taken. Table 3-10 shows that the differences between the rituximab and placebo groups were significant for all radiographic endpoints.

Table 3-10: Radiographic endpoints at 56 weeks

	Placebo + MTX (N=186)	Rituximab + MTX (N=277)
Primary radiographic endpoint		
Change from screening / baseline in total Sharp-Genant score		
N	184	273
Mean (SD)	2.31 (5.283)	1.00 (2.751)
Median	0.5	0
Min, Max	-7.4, 38.1	-7.5, 17.4
p-value		0.0046

3.3.6 Repeat treatment (re-treatment)

As mentioned in Section 3.1.5, the company state that they use data from two phase II RCTs in addition to data from the REFLEX trial for the analysis of long-term efficacy. Neither of these phase II RCTs met the inclusion criteria for the review.

Efficacy analyses of re-treatments were conducted on 279 patients who had received a second treatment with rituximab. All of the patients who received a repeat treatment had responded to the first course of rituximab (response defined as >= 20% improvement in swollen joint count and tender joint count). Of these 279 patients, only 155 had completed their second course of treatment at least 24 weeks before the analysis of ACR responses.

Long-term efficacy data for re-treatment with rituximab are favourable, but the results are limited by the small number of patients available for follow-up.

Primary efficacy outcome

Twenty-four weeks after the second course of rituximab, patients in the re-treatment group (N=155) show better ACR responses from the original baseline following their second course of rituximab (Table 3-11). Thus, the results appear to demonstrate that there can be an additional response to second treatments. The company submission further notes (p.53) that there is no evidence of decreased response with a third treatment dose but provides no evidence to support this statement.

In addition, ACR20 responses collected before 24 weeks also appear to reflect the assumption that the clinical effects of subsequent doses of rituximab may be additive.

Table 3-11: ACR responses and repeat treatment

	N	First course		Second course	
At week 12	226	No.	%	No.	%
ACR20		130	58	159	70
ACR50		58	26	84	37
ACR70		15	7	32	14
At week 16	205	No.	%	No.	%
ACR20		131	64	140	68
ACR50		64	31	83	40
ACR70		25	12	42	20
At week 20	181	No.	%	No.	%
ACR20		121	67	130	72
ACR50		64	35	77	43
ACR70		25	14	38	21
At week 24	155	No.	%	No.	%
ACR20		101	65	111	72
ACR50		51	33	65	42
ACR70		19	12	33	21

Secondary efficacy outcomes

All secondary efficacy outcomes support the argument that there are additional clinical benefits to be gained with each subsequent course of rituximab. ACR50 and ACR70 scores show increased benefit for the second treatment of rituximab (Table 3-11). DAS28 scores decreased after the initial course and remained lower than the baseline score prior to second treatment course. EULAR responses reflect DAS28 scores. HAQ-DI scores show treatment gains were maintained across two courses of rituximab.

Time to re-treatment

Time to re-treatment is provided for both the pooled analysis as above (N=279) and for the 164 patients randomised to rituximab in the REFLEX trial. Of the 279 patients from the pooled analysis who received a second treatment of rituximab, the majority of patients required re-treatment between week 24 and week 48 after their first treatment (Table 3-12).

Table 3-12: Time to re-treatment

Weeks	(N=279)	%
<16	0	0
16-<24	6	2.2
24-<48	176	63.1
48-<72	84	30.1
72-<96	12	4.3
>=96	1	0.4

For the 164 patients from the REFLEX trial who were randomised to rituximab and received a second treatment there was a mean time of 307 days between treatments (Kaplan-Meier

analysis). However, the range of time to re-treatment is very broad with some patients requiring re-treatment as soon as 173 days after initial treatment whilst others did not need re-treatment until 742 days (Table 3-13). According to the REFLEX trial protocol, patients were not allowed re-treatment within 168 days, so in clinical practice this interval could be shorter.

Estimates of time to re-treatment vary in the company submission. Firstly, on p.5 it is stated that the mean first to second course treatment interval is 33.2 weeks (232.4 days). On p.12 it is stated that "the mean time to repeat treatment observed in the phase III randomised control trial was 301 days". On page 58/59, the mean time to treatment is estimated as 307 days. The mean time to repeat treatment used in the economic model is 293 days.

A sub-group analysis was presented (Table 3-13) that assessed whether there were differences in time to re-treatment depending on the number of prior TNFi treatments. This showed that the re-treatment time was greater for patients who had received one TNFi prior to rituximab than for those who had received two or more TNFi.

Table 3-13: Time to re-treatment overall and by number of prior TNFi

	N	Mean (days)	Median (days)	Range (days)
Failed 1 TNF	95	320.2	301	176-585
Failed >=2 TNF	69	289.5	271	173-742
Overall	164	307	289	173-742

3.3.7 Safety

In the company submission safety data are discussed under five headings: all adverse reactions, acute infusion reactions, infections, malignancy and management of rituximab non-responders (see Table 3-14 to Table 3-17 for details). For the analysis of all adverse reactions, data are pooled from two phase II trials (WA17043 and WA16291) excluding patients with lower doses of rituximab; these studies did not meet the inclusion criteria for inclusion in the review. Phase II data are presented alongside adverse reactions data from the phase III REFLEX trial (WA17042). All other safety analyses are conducted on the "all exposure population" (N=1039) i.e. all individuals who have received rituximab for the treatment of RA in phase II and phase III trials regardless of randomisation and dose.

All adverse reactions

As shown in Table 3-14 the most commonly reported adverse reactions are acute infusion reactions or infections and infestations (mainly upper respiratory tract infections). In the REFLEX trial all adverse reactions rates, excluding chills and urinary tract infections, are higher in the rituximab arm. None of the differences between groups are statistically significant.

Table 3-14: Adverse reactions reported in phase II and phase III studies

	Pooled Phase II Study Population		Phase III Study Population	
	Placebo+ MTX (N=189) No.(%)	Rituximab+ MTX (N=232) No.(%)	Placebo+ MTX (N=209) No.(%)	Rituximab+ MTX (N=308) No.(%)
Acute infusion reactions*				
Hypertension	10(5%)	22(9%)	11(5%)	21(7%)
Nausea	14(7%)	19(8%)	5(2%)	22(7%)
Rash	6 (3%)	18 (8%)	9 (4%)	17 (6%)
Pyrexia	1(<1%)	12 (5%)	7 (3%)	15 (5%)
Pruritis	1 (<1%)	14 (6%)	4 (2%)	12 (4%)
Urticaria	0	2 (<1%)	3 (1%)	10 (3%)
Rhinitis	2 (1%)	6 (3%)	4 (2%)	8 (3%)
Throat irritation	0	5 (2%)	0	6 (2%)
Hot Flush	4 (2%)	2 (<1%)	0	6 (2%)
Hypotension	11 (6%)	10 (4%)	1 (<1%)	5 (2%)
Chills	3 (2%)	13 (6%)	6 (3%)	3 (<1%)
Infections and infestations				
Any infection	56 (30%)	85 (37%)	78 (37%)	127 (41%)
Urinary tract infections	8 (4%)	14 (6%)	17 (8%)	15 (5%)
Upper respiratory tract	28 (15%)	31 (13%)	26 (12%)	48 (16%)
Lower respiratory tract infection/pneumonia	10 (5%)	9 (4%)	5 (2%)	8 (3%)
General disorders				
Asthenia	0	3 (1%)	1 (<1%)	6 (2%)
Gastrointestinal disorders				
Dyspepsia	3 (2%)	9 (4%)	0	7 (2%)
Abdominal pain upper	3 (2%)	7 (3%)	1 (<1%)	4 (1%)
Metabolism and nutritional disorders				
Hypercholesterolemia	1 (<1%)	3 (1%)	0	6 (2%)
Musculoskeletal disorders				
Arthralgia/musculoskeletal pain	8 (4%)	18 (7%)	6 (3%)	17 (7%)
Muscle Spasms	0	1 (<1%)	2 (1%)	7 (2%)
Osteoarthritis	1 (<1%)	4 (2%)	0	6 (2%)
Nervous system				
Paraesthesia	2 (1%)	4 (2%)	1 (<1%)	8 (3%)
Migraine	0	4 (2%)	2 (1%)	5 (2%)

*Reactions occurring within 24 hours of infusion

In addition to the adverse reactions highlighted in Table 3-14, the company submission (p.70) identifies "medically significant events reported uncommonly in the rituximab treated population and considered potential reactions to treatment". These include generalised oedema, respiratory disorders (bronchospasm, wheezing and laryngeal oedema), skin and subcutaneous disorders (angioneurotic oedema and generalised pruritis) and immune system disorders (anaphylaxis, anaphylactoid reaction).

Acute infusion reactions

In the trials programme (unspecified) acute infusion reactions were observed in 79/540 (15%) patients receiving rituximab. The numbers of patients receiving placebo who experienced acute infusion reactions are not reported.

Reported infusion reactions for re-treatments are shown in Table 3-15. Data presented show that the first infusion of the first, second and third courses led to 269 (26%), 81(14%) and 20(10%) acute infusion reactions. However, the number of acute infusion reactions rose to 15% with the fourth course but numbers of patients followed up to the fourth course were small (N=40).

No more than 1% of infusion reactions were classified as either serious or severe and only one case was classified as life-threatening. No deaths from acute infusion reactions were reported.

Table 3-15: Acute infusion reactions reported in the "all exposure" safety population

	1st course (N=1039) No. (%)	2nd course (N=570) No. (%)	3rd course (N=191) No. (%)	4th course (N=40) No. (%)
First infusion				
Any adverse event (AE)	269 (26)	81 (14)	20 (10)	6 (15)
Total no. AEs	446	104	26	7
Serious AEs	5 (<1)	1 (<1)	0	0
Severe (CTC grade 3)[a]	11 (1)	2 (<1)	0	0
Life-threatening (CTC grade 4) [a]	1 [b]	0	0	0
Required dose modification	100 (10)	32 (6)	6 (3)	4 (10)
Led to discontinuation	12 (1)	0	0	0
Second infusion				
Any adverse event (AE)	95 (9)	30 (5)	4 (2)	1 (3)
Total no. AEs	124	39	5 (3)	1
Serious AEs	2 (<1)	1 (<1)	0	0
Severe (CTC grade 3)[a]	21(<1)	1 (<1)	0	0
Life-threatening (CTC grade 4) [a]	0	0	0	0
Required dose modification	13 (1)	6 (<1)	0	0
Led to discontinuation	0	1 (<1)	0	0

a Intensity values originally captured as "severe" or "life threatening" have been converted to CTC grade 3 and CTC grade 4 respectively. Only the most severe intensity is counted for multiple occurrences of the same adverse event in the one individual.

b patient 4584: serious anaphylactic reaction (originally reported as CTC grade 3 but later upgraded to CTC grade 4).

Infections

After the first course of treatment a total of 954 infections and infestations were reported by 587 (56%) patients and rates of infections were stable around 81-83 per 100 patient years over repeated courses.

Further clarification from the company identified 11 deaths due to infection but only one of these was considered to be related to rituximab treatment (Table 3-16). The company did not describe the method used to assess whether or not death was related to rituximab in the trial.

Table 3-16 : Infections resulting in death (all exposure population)

Cause of death	Last treatment day	Day of death	Relation to trial treatment
Bronchopneumonia	15	157	No
Unknown cause	14	353	No
Intestinal adenocarcinoma	15	271	No
Myelodysplastic syndrome	28	245	No
Neutropenic sepsis	15	424	No
Cerebrovascular accident	15	167	No
Haemorrhagic stroke	1	176	No
Coronary artery disease	15	234	No
Sepsis	137	366	Yes
Pancreatic neoplasm	18	291	No
Unknown cause	568	623	No

Malignancy

The reported incidence of malignancy was 1.5 per 100 patient years and lies within the expected range of an age and gender matched population. However, this incidence rate is based on the analysis of short-term data.

Management of rituximab non-responders

It is unclear from the company submission how many rituximab non-responders were included in the all exposure group. However the submission does state that 110 patients treated with rituximab received a subsequent DMARD and 78/110 (71%) patients went on to receive one or more TNFi.

Of the 78 patients receiving one or more TNFi after failure of rituximab, four patients (5%) experienced a serious infection.

Table 3-17 shows that the number of serious infections and the number of serious infections per 100 patient years are similar both before and after TNFi. Furthermore, the rates of 5.23 and 7.62 serious infections per 100 patient years are consistent with published data on patients starting their first TNFi (6.4 per 100 patient years).[35] Given the relatively low patient numbers and relatively short-term follow up (i.e.1-2 years in the main), data regarding long-term side effects of rituximab are limited, particularly with respect to patients subsequently treated with a TNFi or further DMARDs. In particular, it is difficult to know how long the immunological effect of rituximab lasts in contrast to the clinical effect. Adding a DMARD or TNFi to a patient who is clinically ill but still B cell depleted, may be a cause for concern and additional data would be helpful to best inform how this should be best managed (Dr R. Moots, personal communication, February 2006).

Table 3-17: Serious* infection rates in patients receiving subsequent treatment with a TNFi

	"all exposure" (N=78)
Before	
Total patient years	57.38
Number of serious* infections	3
Serious* infection events per 100 patient years (and 95% confidence intervals)	5.23 (1.69, 16.21)
After TNFi	
Total patient years	52.50
Number of serious* infections	4
Serious* infection events per 100 patient years (and 95% confidence intervals)	7.62 (2.86, 20.30)

* Reported as serious and/or treated with intravenous antibiotics. Multiple occurrence of the same event in one individual are counted multiple times. Events with a start date prior to the last valid visit date are included.

Comment

The EMEA have reviewed the data on AEs and conclude that treatment with rituximab in RA imposes a significant risk of serious AEs. They recommend that, through the risk management plan (RMP) submitted with the licence application, the company must continue to identify, characterise and prevent or minimise risks relating to rituximab. The RMP identifies one indicated risk, i.e. acute infusion reactions and six potential risks: infection, immunogenicity, neoplasm, immunization response, pregnancy/lactation and drug interactions. The EMEA particularly stress the risks of infusion reactions and infection. The risks associated with rituximab should be considered alongside the fact that rituximab may offer benefits to patients with severe RA who are resistant to all other medications.

The EMEA's concern regarding the use of rituximab mirrors the general belief that the close surveillance of patients treated with any of the new biologics, both during treatment and post-treatment, is important. In the UK, a national biologics register has been established to record data on patients receiving newer biologic therapies.[36]

3.4 Indirect comparisons

Six of the 44 studies identified in the literature review (see Section 3.1.3) are included in the indirect comparison analysis; study data are summarised in Table 3-18. The Genovese trial is included, despite abatacept not being licensed at the time of submission, in order to increase the power and sample of the ACR indirect comparison.

Table 3-18: Published ACR responses

Treatment	Placebo			Treatment			Reference:
	>ACR20	>ACR50	>ACR70	>ACR20	>ACR50	>ACR70	
Rituximab + MTX	18%	5%	1%	51%	27%	12%	Cohen et al 2006[13]
Leflunomide	46%	23%	9%	52%	34%	20%	Strand et al 1999[33]
Etanercept + MTX	27%	3%	0%	71%	39%	15%	Weinblatt et al 1999[34]
Infliximab + MTX	20%	5%	0%	50%	27%	8%	Maini et al 1999[35]
Adalimumab + MTX	30%	10%	3%	63%	39%	21%	Keystone et al 2004[36]
Abatacept + MTX	20%	4%	2%	50%	20%	10%	Genovese 2005[37]
Gold	NA	NA	NA	18%	5%	1%	Assumed equal to MTX
Ciclosporin	NA	NA	NA	18%	5%	1%	Assumed equal to MTX
Palliative care (MTX)	NA	NA	NA	18%	5%	1%	Cohen et al 2006[13]

All trials were considered to have a common comparator (MTX) and therefore be suitable for "adjusted indirect comparison analyses." The company employed an adaptation of the method used by Choi et al[37] and Bansback[38] where a weighted average of placebo ACR rates from all the included trials was calculated i.e. a 'reference placebo' (RP). This figure was then used to adjust the specific treatment ACR rates for each of the trials. The company submission included both the formulae and the calculations of adjusted ACR rates. The adjusted ACR rates are shown in Table 3-19 and are used as absolute values in the company submitted economic model.

Table 3-19: Adjusted ACR responses

Treatment	>ACR20	>ACR50	>ACR70	Reference:
Rituximab + MTX	63%	33%	15%	Cohen et al 2006[1]
Leflunomide	31%	21%	12%	Strand et al 1999[39]
Etanercept + MTX	70%	39%	15%	Weinblatt et al 1999[40]
Infliximab + MTX	59%	32%	9%	Maini et al 1999[41]
Adalimumab + MTX	60%	37%	20%	Keystone et al 2004[42]
Abatacept + MTX	60%	24%	12%	Genovese 2005 [43]
Gold	26%	9%	3%	Assumed equal to MTX
Ciclosporin	26%	9%	3%	Assumed equal to MTX
Palliative care (MTX)	26%	9%	3%	Cohen et al 2006[1]

3.4.1 Critique of indirect comparison

Identification of papers

It is unclear from the company submission how the six trials in Table 3-19 were selected for inclusion in the main indirect comparison analysis. Further clarification from the company did include details of the literature searches but did not explain the reasons for excluding 38 of the 44 trials identified. For the analysis to be conducted correctly all relevant trials should be included and used to inform the range of estimates taken for the RP rate. Furthermore, the search was conducted in April 2005 and has not been updated for the submission in November 2006.

On inspection it would appear that the Strand paper[39] has been wrongly identified for use in the indirect comparison. This paper compares ACR responses at 52 weeks for leflunomide versus MTX in patients who have had RA for a mean of seven years and have not previously received MTX.

In the adjusted indirect comparison exercise, there is an error in the calculation of the RP. The sample size of the placebo group in the Strand[39] paper is stated as 133, it should be 182. This error does not significantly affect the ICERs calculated by the company.

Appropriateness of methodology

Computationally the methodology is sound. However, there are several issues that should be noted. Firstly, whilst it is assumed that the comparator arms are the same for all trials i.e. MTX, this is not wholly accurate. Four of the trials included a comparator arm of placebo plus MTX,[1,40-42] one MTX only[39] and one placebo only (although 76% and 82% of rituximab and placebo respectively were receiving concomitant MTX).[43]

Secondly, as no trials reporting ACR responses for treatment with either gold or ciclosporin were found, the adjusted ACR analyses were conducted assuming equivalence to MTX. There is no evidence for this assumption.

For the analyses to produce meaningful results the trials need to be exchangeable e.g. similar patients from similar disease populations. As the company comments, the severity and duration of disease differs between trials with the disease being slightly more severe in both the REFLEX trial and the abatacept trial[43] (See Table 3-20). Only the trials assessing abatacept and rituximab were conducted in patient populations where patients had failed on a previous TNFi, an important feature of the REFLEX population and licensed indication for rituximab.

Table 3-20: Severity of disease in trial populations

Treatment	Source	Swollen joints (mean no.)	Tender joints (mean no.)	DAS28 (mean score)	HAQ(mean score)	Duration of disease (mean no. of years)
Infliximab	Maini[41]	19	24,32*	NS	1.8	7,9*
Etanercept	Weinblatt[40]	17,20*	28	NS	1.5**	13
Adalimumab	Keystone[42]	19	27,28*	NS	1.5	11
Abatacept	Genovese[43]	22	31,33*	6.5	1.8	12
Leflunomide	Strand[39]	13,15*	16	NS	0.9	7
Rituximab	Cohen[1]	23	33,34*	6.9	1.9	12

*Values for placebo, treatment, **HAQ-DI, Ns=not stated

Finally, the RP rate is a weighted average of the ACR responses reported in the six included trials. It is unclear of what population the RP rate is indicative and how sensitive the results are to plausible values for particular relevant patient groups. A more appropriate method would be to calculate a range of RP rates that include/exclude trials based on population characteristics thereby allowing for heterogeneity between the trials. For example, one RP rate could be calculated for the combination of the abatacept trial[43] and the REFLEX trial. These trial populations resemble most closely the patients identified by the company in the statement of their decision problem.

Having calculated a RP rate the company use an adapted version of Choi's[37] adjustment method (published method), to adjust ACR rates for all six trials. A rationale for the use of a non-standard method is not included so it is unclear which is preferable in particular circumstances.

These issues around the use of a single RP rate and the lack of justification for the methodology used highlight concerns about the validity of the adjusted ACR rates and their use in the economic model.

3.5 Summary of clinical evidence

3.5.1 Clinical results

Direct comparison: rituximab plus MTX versus placebo plus MTX

- In the REFLEX trial, at 24 weeks, 51% of patients in the rituximab group reached an ACR20 response compared to 18% of patients in the placebo group. At 48 weeks, 19% of patients in the rituximab group reached an ACR20 response compared to 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
- Estimated mean time to re-treatment in the REFLEX study was 307 days for those patients randomised to rituximab. Pooled analyses reveal that rituximab patients show better ACR responses from their original baseline response following their second course of rituximab
- Safety analyses show that in the REFLEX study all adverse reactions rates, excluding chills and urinary tract infections, are higher in the rituximab group compared to the placebo group. Of the 110 patients in the all exposure population who received a subsequent DMARD, 78 patients received one or more TNF inhibitor therapy, four patients (5%) experienced a serious infection

Indirect comparison: DMARDs (including TNFi) efficacy values adjusted by reference placebo

- Reference placebo value: MTX (26%)
- Absolute adjusted ACR20 efficacy values: rituximab (63%); leflunomide (51%); etanercept plus MTX (70%); infliximab plus MTX (59%); adalimumab (60%); abatacept (60%); gold (26%), ciclosporine (26%)

3.5.2 Clinical issues

Direct comparison: rituximab plus MTX versus placebo plus MTX

- The REFLEX trial does not compare rituximab with a relevant comparator (leflunomide or second or third TNFi) to directly answer the key clinical and cost-effectiveness questions explored by the company
- Long-term efficacy data (re-treatment) and safety data (including use of DMARDs post rituximab) are limited by short duration of follow up, small numbers of patients and unclear description of patient flows. Long-term efficacy analysis includes patients from trials excluded from the systematic review
- Interpretation of clinical evidence is hampered by inconsistent presentation of p-values in the company submission

Indirect comparison: DMARDs (including TNFi) efficacy values adjusted by reference placebo

- Inclusion and exclusion criteria used in adjusted indirect comparison exercise is unclear; it is unknown whether or not all relevant trials have been included
- Comparator arms of the trials which are used in the calculation of the RP are not the same
- No clinical evidence to support the equivalence of MTX, gold and ciclosporin
- No rationale for method of indirect comparison used

4 COST EFFECTIVENESS

4.1 Critique of cost-effectiveness review

4.1.1 Health economics literature search for rituximab related articles

The submission identifies two abstracts[44,45] describing the cost effectiveness of rituximab in the treatment of RA. No details of the search strategy used are provided. The abstracts are neither summarised nor discussed in the submission; the company state that they reflect the economic evaluation in the company submission. Other conference abstracts presented by the company discussing the cost effectiveness of rituximab in RA have also been published[44,46-50] but are not identified in the submission.

4.1.2 Health economics literature search for TNFi related articles

The company conducted a review which was intended to update and supplement the health economics review that was published in the recent Health Technology Assessment report entitled "A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of RA in adults and an economic evaluation of their cost-effectiveness".[4] The company did not provide a summary of the methods or the results of this previously published review of the health economics literature.

As part of their review, the company developed a search strategy to "identify economic models, information on costs and cost effectiveness of TNFi for the treatment of RA" (company submission p.79).

Identification and description of studies

The submission included full details of the electronic search strategy used in the review update. The ERG was therefore able to replicate the electronic searches undertaken by the company. The databases searched were described with dates. The total number of papers initially found and the number of papers excluded from the review were reported. Reasons for excluding papers were also provided.

Stated inclusion criteria were:

- **Study design**

  Cost-consequence analysis, cost-benefit analysis, cost-effectiveness analysis, cost-utility analysis, cost studies (UK only), quality of life studies

- **Population**

  People with RA; other forms of arthritis are excluded

- **Intervention**

  Etanercept, infliximab or adalimumab

- **Comparator**

  Disease modifying anti-rheumatic drugs

- **Outcome**

  Quality of life estimates, cost estimates, cost-effectiveness

- **Time horizon**

  February 2005 to October 2006

Using these inclusion criteria, the company identified three studies[51-53] for inclusion in the review; none of the studies included rituximab as a comparator to TNFi.

Data extraction

The company extracted data from the three studies identified for inclusion in the review. The key features of the studies are presented and discussed in the main body of the submission with detailed descriptions of the studies provided in an appendix. In the appendix, details of the three studies are summarised in a format based on a simplified version of the original Drummond and Jefferson checklist[54] for the critical appraisal of published economic evaluations.

Quality assessment

The submission states that the Consensus on Health Economic Criteria[55] was used to assess the quality of the included studies; it is reported that each study was considered to be of adequate quality as at least 15 checklist points were met by each study. However, the results of the quality assessment conducted by the company are not fully described in the text.

4.1.3 Summary and conclusions

The company's review of the health economics literature is brief. The submission does not include a search strategy for the identification of their two rituximab abstracts. It is therefore not possible to determine whether other relevant rituximab papers/abstracts exist without

conducting a new literature search. It would have been useful if the company had provided a summary of the results of the previously published Health Technology Assessment report[4] referred to in the company submission. It would have been appropriate to discuss the results of the update review in light of the results of the previous review. Finally, reading of the literature review was hampered by the inconsistent and often inaccurate use of references and erroneous labelling of tables.

4.2 Overview of company economic evaluation

4.2.1 Description of company model

A micro simulation Markov model was constructed in Microsoft Excel based upon the phase III randomised control trial of rituximab (REFLEX trial). Patient disease progression is tracked within the model according to their HAQ score. By using micro-simulation of 10,000 RA patients, patient history is kept in memory and cost utility values are assigned to each individual at each cycle.

All patients enter the model after failing an initial TNFi and at the start of their next treatment option. Patients may then respond within one of the three ACR response categories. Next, patients are allocated a pre-defined drop in HAQ according to which ACR category (ACR20, ACR50 and ACR70) was achieved. Patients who respond are assumed to receive their respective treatment for a pre-specified length of time before stopping/failing therapy. A fixed rate of HAQ progression/deterioration will occur during a patient's time on treatment. At the point of treatment failure the patient will experience a further increase in HAQ (rebound effect), before commencing the next pre-defined treatment within the sequence, where the above process starts again. At every six monthly cycle in the model, patients are subject to an age, sex and RA adjusted risk of death.

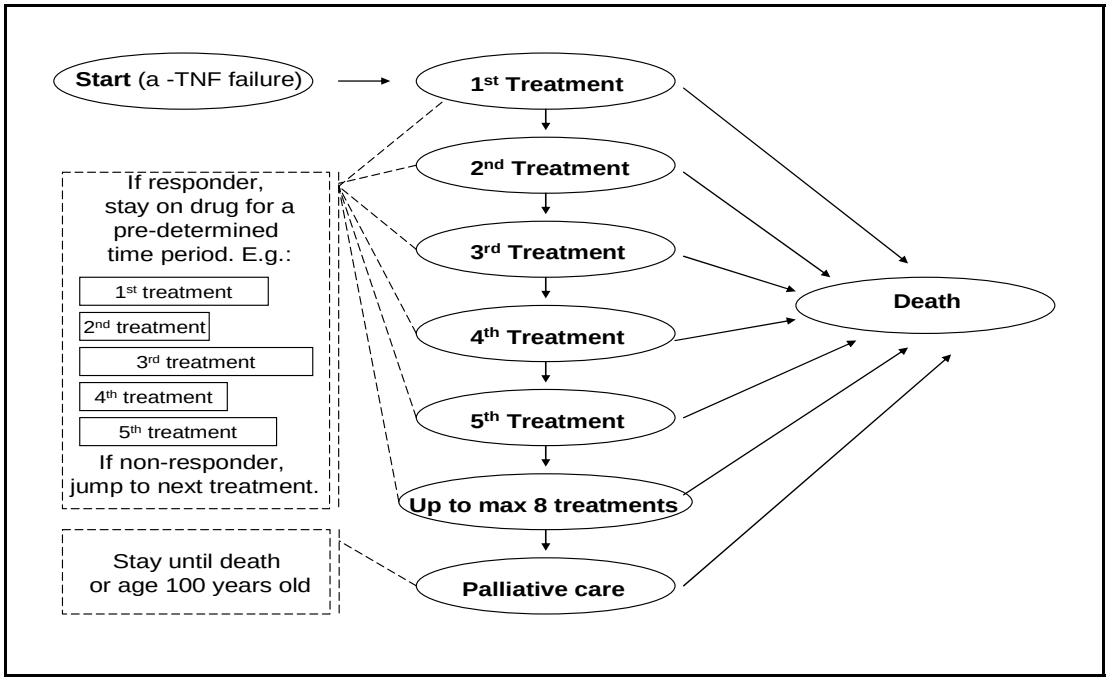


Figure 4-1: Structure of the company model

Table 4-1: Model variables

Parameter	Parameter value			Source
Treatment	>ACR20 <ACR50	>ACR50 <ACR70	>ACR70	
Rituximab + MTX	30%	19%	15%	REFLEX[1]
Leflunomide	11%	8%	12%	Strand[39]
Adalimumab + MTX	23%	17%	20%	Keystone[42]
Gold/ciclosporin/MTX	18%	6%	3%	REFLEX [1]
Infliximab + MTX	27%	22%	9%	Maini[41]
ACR response	HAQ score update			
Non-responder	-0.1			REFLEX[1]
ACR20	-0.45			REFLEX[1]
ACR50	-0.85			REFLEX[1]
ACR70	-1.11			REFLEX[1]
Time on treatment for responding patients	Years			
Rituximab + MTX	4.25			Estimate
Leflunomide	4.10			Adapted from Barton[56]
Adalimumab + MTX	4.25			Adapted from Barton[56]
Gold	3.85			Adapted from Barton[56]
Ciclosporine	1.7			Adapted from Barton[56]
Infliximab + MTX	2.46			Adapted from Crnkic[57]
Rituximab mean time until repeat treatment	9 months			REFLEX[1]
HAQ score progression per model cycle – all therapies	0.017			Scott[58]
HAQ score progression per model cycle – palliative care (MTX)	0.065			Bansback[38]
Mortality risk adjustment	1.33^{HAQ}			Barton[56]

4.2.2 Population

The patient population in the model is assumed to have equivalent characteristics to the patients in the REFLEX trial (WA17042), which compares rituximab plus MTX with placebo plus MTX. Sub-group analysis is carried out for patients according to their number of prior TNFi.

The characteristics of the patients in the model appear to be consistent with the licensed indication of rituximab – adult patients, severe RA, inadequate response or intolerance to other DMARDs including one or more TNF inhibitor therapies. However, as noted in Section 3.2.3, 40% of patients in the REFLEX trial have received more than one TNFi before rituximab.

4.2.3 Perspective and time horizon

An NHS perspective is adopted, in line with current NICE guidance. The economic evaluation purports to capture direct costs and benefits only. However, there is an option in the economic model which allows the user to include the indirect costs of unemployment. The lifetime costs and benefits of rituximab are included in the model in order to capture the long-term chronic nature of RA. Patients are followed from entry into the model until they either die or reach 100 years of age.

4.2.4 Comparator

In their statement of the decision problem, the company compares the intervention (rituximab plus MTX) with clinical management strategies without rituximab (see Section 2.7.1).

Given the uncertainty surrounding the many treatment pathways for patients with RA, it is unlikely that a consensus of opinion from medical experts can be reached regarding (i) choice of drugs in the sequence or (ii) order of drugs in the sequence. The appropriateness of the treatment pathways considered by the company in both scenarios may therefore be subject to debate within the medical community.

4.2.5 Efficacy

The primary measures of efficacy used in the REFLEX trial and other relevant RCTs are ACR response rates (<ACR20, ACR20, ACR50 and ACR70).

The economic evaluation uses adjusted ACR response rates from the indirect comparison described in Section 5.6 of the company submission. An appropriate refinement is made to the adjusted ACR response rates in order to calculate the actual proportion of patients falling within each ACR category.

ACR responses are used in the derivation of cost per QALY values. Given the criticisms of the company's indirect comparison exercise outlined in Section 3.4.1, the ERG is not confident that the adjusted ACR responses employed in the economic model are valid.

The ERG notes that there is an option to use unadjusted response rates in the economic model.

Time on treatment for responding patients

Once a patient responds to treatment, there is no other stopping rule (except for death and age 100) until the pre-determined average time on treatment. As the company state (p.101) "robust estimates of this parameter do not exist". In the base case evaluation, rituximab is assumed to have an equivalent average time on treatment as etanercept and adalimumab and an additional 1.79 years on treatment compared to infliximab. Estimates of average times on treatment are used to calculate the final patient monitoring costs described in the company model.

4.2.6 Health benefits and utilities

In the economic model, patients' disease severity is measured by the HAQ. Baseline HAQ scores and change in HAQ scores relative to ACR response are taken from the REFLEX trial; this relationship is assumed to be equivalent across treatments. The rate of HAQ progression is also assumed to be equivalent across all therapies except for palliative care.

In the economic model, there is a mechanism which permits HAQ scores (intermediate outcome) to be mapped to QALYs (final outcome). The company note that recently NICE has endorsed the link between HAQ scores and QALYs.[15]

The company uses three different methods to transpose HAQ scores into QALYs. The HUI-3 transformation is used in the base case as it is based on the largest sample (N=2000) of patients (treated with adalimumab) and the data were collected in a clinical trial setting. The company note that this approach is inconsistent with the NICE reference case, as the utility scores were not derived from the EQ-5 D instrument using standard gamble or time trade off methods. The use of the two alternative transformations using EQ-5D scores is explored in a one-way sensitivity analysis (SA).

The costs and benefits of adverse events of drug treatments are excluded from the economic model.

4.2.7 Resources and costs

Three different categories of resource use are identified in the model: drug and drug administration resource use, patient monitoring resource use and inpatient visit resource use (see Table 42 – Table 46 in the company submission for more detail).

Drug administration costs are presented only for rituximab, infliximab, etanercept and adalimumab. The company assumes that the average time taken to administer rituximab is five hours compared to three hours for infliximab and this difference is reflected in the calculation of required health care personnel attendance times.

Appendix 8 of the company submission presents detailed information on typical resource use and associated frequencies of patient monitoring visits and examinations; this information was derived from an interview with a clinician. The resource patterns described in Appendix 8 are not wholly reflected in the calculation of patient monitoring costs in the economic model.

In the company submission, estimates of inpatient resource use are derived from NOAR[59] data and grouped into six HAQ strata. Using the national average rheumatology inpatient cost per day, inpatient resource use costs are calculated. Whether or not the rheumatology inpatient cost includes the cost of surgery is unclear. Upon request, the company clarified this matter by generating an expected cost of surgery by HAQ category and recalculating total inpatient resource use costs.

4.2.8 Discounting

Health benefits and costs were discounted at 3.5% in line with current NICE guidance.[60]

4.2.9 Results

The results of the original company economic evaluations are shown in Table 4-2 and Table 4-3. In terms of cost per QALY, the company conclude that rituximab can be considered a cost-effective treatment option in RA.

Table 4-2: Scenario one: no sequential use of TNFi

	Total LYG	Total QALYs	Difference	Total Drug Costs	Total Costs	Difference	ICER
Rituximab	17.99	3.051	0.727	£36,003	£41,229	£10,675	**£14,690**
No rituximab	17.71	2.324		£24,254	£30,554		

Table 4-3: Scenario two: sequential use of TNFi

	Total LYG	Total QALYs	Difference	Total Drug Costs	Total Costs	Difference	ICER
Rituximab	18.50	3.933	0.526	£62,608	£66,583	£6,103	**£11,601**
No rituximab	18.18	3.407		£55,744	£60,480		

4.2.10 Sensitivity analysis

Univariate sensitivity analysis and probabilistic sensitivity analysis (PSA) were conducted by the company. The results of the SA are presented in Table 4-4. As can be seen from the results of the univariate SA, the model is most sensitive to variations in patient age (Scenario 1) and the assumed interval between those patients who respond to treatment (Scenario 2). The one-way sensitivity analysis demonstrates that the cost per QALY of rituximab varies from £5000 to £31,500 per QALY.

In terms of the PSA, the ERG noted that in the original company submission (based on the original version of the model) the parameter sets are subjected to variations that are governed by the estimated standard deviation of each variable, rather than the standard error of each estimated statistic. The ERG also noted the use of an irregular sampling method from the primary distribution. The ERG concluded that the PSA results (scatterplots and cost-effectiveness analysis curves) should be disregarded as the PSA methodology has not been applied correctly. This is more fully discussed in Section 4.3.1).

Table 4-4: Sensitivity analysis results

Variables	Assumptions	Scenario 1: Result (cost per QALY)	Scenario 2: Result (cost per QALY)
Rituximab treatment frequency	Every 6 months Every 12 months	£23,774 £9,759	£24,151 £4,789
Discount rate (QALYs)	1.5% 6%	£12,528 £17,686	£9,198 £15,274
Discount rate (costs)	1.5% 6%	£15,337 £13,967	£15,067 £8,127
Baseline HAQ score	1 2.5	£12,166 £20,302	£9,643 £15,643
Negative QALYs	Allowed Dis-allowed	£14,690 £15,400	£11,601 £12,175
Time on treatment for responders (all drugs)	2 years 6 years	£13,228 £17,267	£10,360 £10,148
Time on treatment for responders (rituximab)	2 years 6 years	£13,929 £15,218	£9,023 £12,307
HAQ to QALY equation	Hurst (1997) Hawthorne (2000)	£12,756 £18,872	£10,113 £14,415
Rebound effect	100% 50%	£14,690 £9,190	£11,601 £7,549
Response rates	Adjusted Unadjusted	£14,690 £15,790	£11,601 £12,456
RA mortality risk multiplier	1 2	£13,266 £17,675	£10,376 £13,461
HAQ progression rate	0.0085 0.0325 (palliative) 0.034 0.13 (palliative)	£17,521 £14,610	£13,730 £12,327
Baseline age	35 75	£12,729 £31,518	£10,482 Rituximab dominant
Rituximab drug administration	8 hours	£15,008	£12,040
HAQ drop for ACR20, 50, 70	-50% +50%	£19,751 £11,780	£15,663 £9,217
No of prior TNFi	2 or more		£14,766
First TNFi in treatment	Adalimumab Infliximab		£11,260 £9,217

4.2.11 Model validation reported within the submission

To determine structural validity, the results of the model were calculated without using model formulae and the expected outputs were compared with the true outputs. The company conclude that all of the cases passed the test with acceptable minimum differences between expected and true outputs. This provides reassurance that no serious formula errors have gone undetected. To determine scenario validity, 16 parameters were changed in the model. The tests revealed two issues; one (negative QALY scores) has been fixed by a change in the programming code and the other (a costing error) has been left unchanged as it has no impact on results.

4.2.12 Budget impact analysis

The company submission estimates the five-year budget impact of introducing rituximab for use in RA following the failure of one TNFi. In year one cost savings are estimated to be in the region of £5 million rising to £11 million in year five. In the budget impact analysis, it is assumed that patients sequence through TNFi.

4.3 Critique of company model

4.3.1 Model implementation and validation

Early examination of the submitted economic model identified some aspects of its implementation, which caused concern as to its reliability for generating estimates of cost-effectiveness. Two particular issues were raised with both NICE and the company concerning the method of randomisation, and also the representation of parameter uncertainty in PSA.

Randomisation

The model requires sets of random numbers to be generated which are used to determine the occurrence of two key events experienced by patients:

- response to therapy;
- death.

In the model originally submitted by the company, these random numbers were linked to specific model periods (1-100) and were used for both an intervention patient and the corresponding comparator patient. However, since the intervention arm involves an additional treatment phase (with rituximab) the treatment received for each period is not the same - being offset by 1 or 9 periods (depending on response to rituximab). The consequence of this is that corresponding patients could have widely different experiences, including quite different patterns of response to the same treatments. The model was structured to make direct pairwise comparisons of outcomes and costs for the two treatment sequences as applied to individual patients, resulting in many simulated patients appearing to experience extreme differences in relative survival times, treatment outcomes and costs.

It is possible to carry out such micro-simulations in two ways:

- case-controlled, in which the essential characteristics of each simulated patient are preserved, except for the effects of the intervention (equivalent to a simulated case-controlled clinical trial);

- a cohort comparison where the characteristics of simulated patients are randomly drawn from a pre-specified distribution for each study arm separately; this should result in patient sets with approximately equal average characteristics, but without any specific linking of patients.

Since the submitted model was presented as though it were a case-controlled design, yet yielded unrealistic pairwise differences for a number of pairs, the ERG expressed concern that they could not be confident that the economic results of the model were not subject to inbuilt error or bias, particularly since the bespoke programming of the simulation was not easy to follow without proper user documentation. In discussion with the model authors (NICE, Roche and ERG teleconference on 14th December 2006), it transpired that they had not intended for the model to be interpreted as case-controlled. However, they acknowledged that the mismatch in patient experience between the two arms could potentially persist as an aggregate bias, and undertook to consider whether the model could be modified to overcome some or all of these problems.

Probabilistic sensitivity analysis

Examination of the PSA carried out in the model indicated that for some model parameters the authors had employed standard deviations (measures of sample dispersion) rather than standard errors (measures of uncertainty in parameters estimates) - in particular for parameters of normal and Weibull distributions. This led to strange scatterplots of incremental costs and outcomes quite untypical of expected PSA results. The PSA was therefore of no value in resolving questions of confidence in cost-effectiveness results. The model authors undertook to rework the PSA to correct these mistakes.

Revised company model

On January 3rd 2007, the ERG received a modified version of the company model with some additional documentation to assist interpretation of the program code. The changes made to the model were:

- random numbers governing response to treatment were preserved for patient pairs in the two arms of the model;

- for PSA variables subject to normal distributions, standard deviations were replaced by standard errors;

- for resource use distributions, a truncated normal distribution was substituted for the previous gamma distribution in PSA;

- the authors state that they were unable to obtain the information necessary to correct the acknowledged problems with representing uncertainty in the Weibull parameters for time on treatment. Within the model a note is included suggesting that in the new PSA this aspect of uncertainty was disabled, though the ERG was unable to verify this within the program code.

Model validation

In view of the complexity of the bespoke program code, and lack of annotation or documentation associated with the original model, the ERG was uncertain whether the model structure and assumptions described in the submission had been correctly translated in the model programming. To address this concern the ERG constructed a simple cohort validation spreadsheet to represent the base case scenario using the same assumptions and parameter values as in the company model. This exercise yielded estimates of costs and benefits (Table 4-5) which were generally somewhat higher, but led to incremental values and a cost-effectiveness ratio sufficiently close to the submitted values as to constitute a reasonable validation of the model logic. Indeed the observed discrepancies are likely to be in part due to the lack of half-cycle correction in the submitted model, but which was applied in the ERG validation spreadsheet.

Table 4-5: Validation of model logic-base case results from original model and ERG simple cohort spreadsheet

	Direct medical costs	QALYs	Incremental cost per QALY
Original submitted model			
Rituximab strategy	£41,229	3.051	
Comparator	£30,554	2.324	
Incremental	**£10,675**	**0.727**	**£14,690**
Validation spreadsheet			
Rituximab strategy	£44,192	3.161	
Comparator	£33,044	2.430	
Incremental	**£11,148**	**0.732**	**£15,235**

4.3.2 Mortality calculations

The submitted model uses mortality probabilities taken from the Government Actuary's Department published life tables.[61] Unfortunately, two errors have been made in the use of these estimates:

- the *annual* probabilities of death have been applied to each 6-month period in the model so that approximately double the number of deaths occur in each period than would be expected;

- a simple averaging of probabilities is carried out weighted according to the gender balance of patients at the baseline age (52 years). However, since female mortality rates are generally lower than those for males, the relative weighting of survivors at risk changes over time.

Applying corrected mortality probabilities to the revised submitted model substantially alters the life expectancy of patients in both arms of the comparison (from about 18 years to about 23 years). However, its impact on the cost effectiveness of rituximab is small - incremental cost per patient is reduced by less than £100, and incremental QALYs gained per patient increase by 0.05-0.06, resulting in a modest reduction in the ICER (from £14,700 to £13,600/QALY).

4.3.3 Progression of functional disability

Importance of HAQ to model logic

The logical structure of the submitted model mediates the impact of different treatments through the estimation of changes in mean HAQ scores. This is composed of two parts:

- direct modification of the HAQ during response to that treatment, and

- reduction in the rate of long-term deterioration in HAQ scores during response to treatment.

Regardless of which effect is considered these changes in HAQ impact on mortality/survival, patient utility (quality of life) and direct medical costs (both treatment-related and disease-related), as illustrated in Figure 4-2. Thus, we can expect that assumptions and parameter values governing mean HAQ scores will be highly influential on model results.

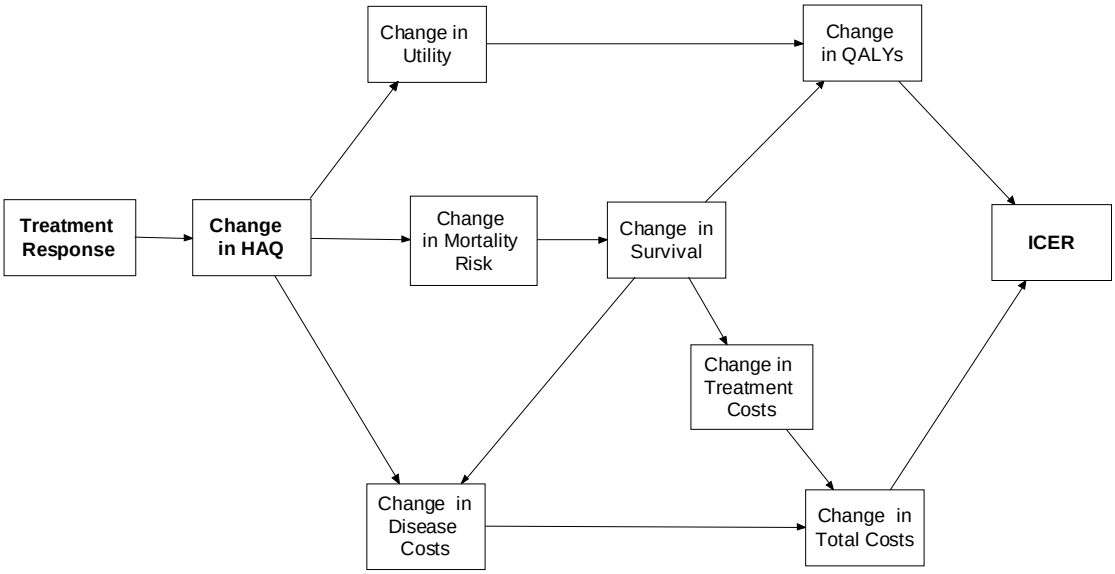


Figure 4-2: Effects of changes in HAQ in company model

HAQ progression and non-response to treatment

The direct modification of HAQ attributable to response to treatment is applied for each period in which the treatment continues to be effective. The proportion of patients considered to have achieved a response (at one of three levels) is estimated from trial data. However, even those patients not responding to treatment receive some benefit in the form of a reduction in HAQ for the 6-month period in which the treatment was trialled, corresponding to measured reductions in HAQ for non-responding patients observed in trials. During the first treatment period (regardless of response) no long-term progression of HAQ scores is applied - presumably as the observed treatment effects in trials are considered to have already accounted for any underlying deterioration.

However, this algorithm leads to anomalies in the model. Consider a patient who undergoes a sequence of seven different treatments but fails to respond to any of them. According to the model, at the end of this process, the patient's HAQ will return to exactly the same value as at the start of the process 3.5 years before without any progression in functional disability. By contrast, the patient who receives an efficacious treatment over the same period will return to the initial HAQ score worsened by 3 years of steady functional deterioration. In other words the best option appears to be to give patients an endless succession of placebo treatments, at virtually no cost, which appear somehow to stave off the development of disability indefinitely - the ultimate cost-effective solution. Clearly this is implausible, and indicates that periods of ineffective treatment should be subject to progressive deterioration in HAQ within the model.

The impact of this modification has been tested in the validation spreadsheet and produces very minor changes in incremental costs and QALYs so that the ICER increases by a trivial amount (from £15,235 to £15,277/QALY gained).

Impact of different values for HAQ progression rates

Two annual progression rates are used in the model to represent worsening HAQ scores in the long-term: a rate of an additional 0.034 points per year whilst undergoing any active treatment, and a greater rate of 0.13 points per year when all active treatment options have been exhausted and the patient is deemed to receive only palliative therapies. As change in HAQ scores is the prime driver of both benefits and costs in the model it is not surprising that these two parameters are influential in the estimation of the cost effectiveness of rituximab. Figure 4-3 presents a 2-way sensitivity analysis from the revised model illustrating the impact of various values of the progression parameters on the base case scenario. This illustrates how the ICER varies with the assumed rate of increase in HAQ per model period whilst on active treatment, and with various possible ratios between the progression rate in palliative care and that on active treatment. For convenience, the analysis was carried out using the validation spreadsheet.

The company's base case scenario assumes that progression on palliative care is nearly four times the rate on active treatment (bottom line on the chart), and suggests that beneficial ICER estimates are obtained over a wide range of progression rate values. By contrast if it were demonstrated that in fact long-term HAQ progression after the failure of all active treatment options is little different from that experienced previously, then it is unlikely that rituximab could be considered cost effective under any assumptions. The threshold ratio for cost-effectiveness (where the ICER = £30,000 per QALY gained) using the company's assumption of 0.017 increase in HAQ per period whilst on active treatment is 1.257, corresponding to a long-term progression rate of at least 0.021 per period. It is important, therefore, to examine the evidence supporting the progression rates employed in the submitted model

The dominant source of the variations shown in Figure 4-3 is the impact of HAQ progression on incremental utility, since incremental costs are very insensitive to changes in HAQ. This indicates that the key pathways for ICER changes in the submitted model are those which involve converting changes in HAQ into utility differences, and into survival differences (see Figure 4-2).

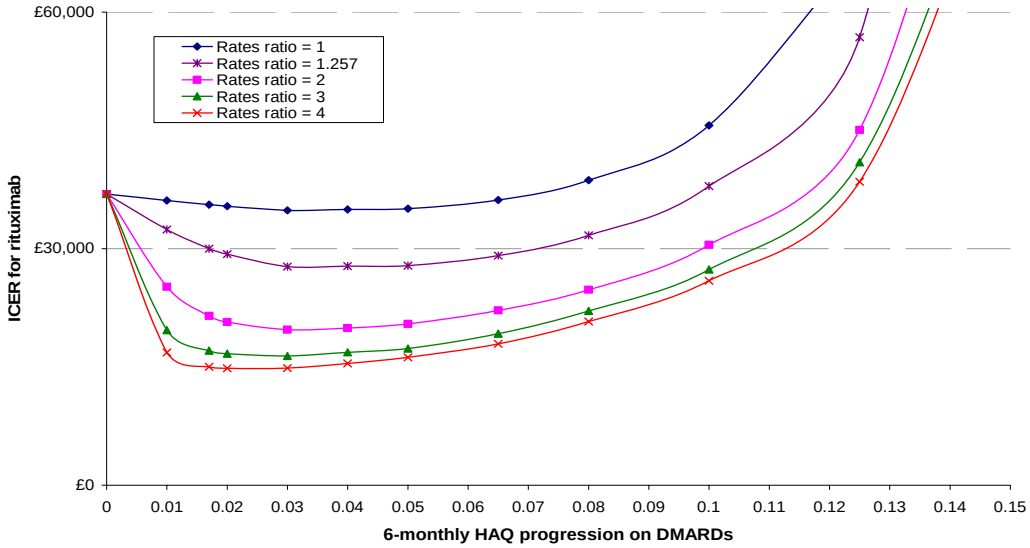


Figure 4-3: Illustration of the sensitivity of the estimated ICER for rituximab to different values of the progression rates of HAQ scores

Theoretical considerations concerning HAQ response and progression

The HAQ is a self-reported tool designed to capture important aspects of functional disability and to allow assessment of impairment to be given a quantitative value. The scoring procedure is based on eight separate items (constructed from answers to 20 questions) each of which may take integer values from 0 to 3. By simple averaging, these yield a single HAQ score that ranges from 0.0 to 3.0 in steps of 0.125 - i.e. 25 distinct possible values.

There is considerable published literature discussing the relative merits of HAQ and similar indices, each of which has particular strengths and weaknesses. We concentrate here on the properties of the HAQ as they affect the measurement of functional impairment over extended periods of time, with special interest in the implications of these properties for the way HAQ changes in cohorts of patients are represented in the submitted model.

Closed scale. The restriction of HAQ to values falling between two boundaries (0 and 3) causes anomalous affects for the representation of treatment effects. The submitted model associates a fixed decrement in HAQ score with each of the three levels of response to treatment, independent of the prevailing HAQ score. This means that patients with relatively low initial HAQ and a good response to treatment could be assigned an off-scale negative HAQ while on treatment. Conversely, patients with high initial HAQ score who undergo HAQ progression during a prolonged period of treatment may then 'rebound' on treatment

failure to a HAQ score exceeding the maximum allowed value (3). The model copes with these problems by truncating the calculated scores to the relevant minimum or maximum boundary value. However, this implies that in fact it is not appropriate to use a single fixed effect parameter to model the effect of treatments on HAQ, since there is clearly a diminishing effect as the underlying score approaches either boundary.

Similar logic applies to the model assumption that HAQ progression over time can be represented by a simple linear function of disease duration. If patients live long enough it is inevitable that at some point the HAQ score will exceed 3 and must thereafter be truncated. Clearly the current model does not adequately represent the characteristics of the HAQ scale, potentially leading to distortion and bias.

Score dynamics. The submitted model is very basic in its representation of the HAQ score. There is no recognition that there is variation in scores within a cohort, so that each modelled patient is afforded the same HAQ score. Also, there is no attempt to consider the effects of inherent uncertainty/variability in the scores of patients over time. This is an important aspect of all self-reported instruments and involves alterations in patient perceptions of their condition (responder variability) as well as the inherently variable nature of the entities being measured (disease variability). The extent of such changes are clearly seen in Figure 4-4 (reproduced from Scott [58]). Although the mean score appears to increase slowly and steadily over time, the extent of individual fluctuation from year-to-year is considerable. Of particular note is the experience of patients close to the top of the scale; the notion that any patient arriving at the maximum scale point is thereafter doomed to remain there indefinitely is clearly refuted.

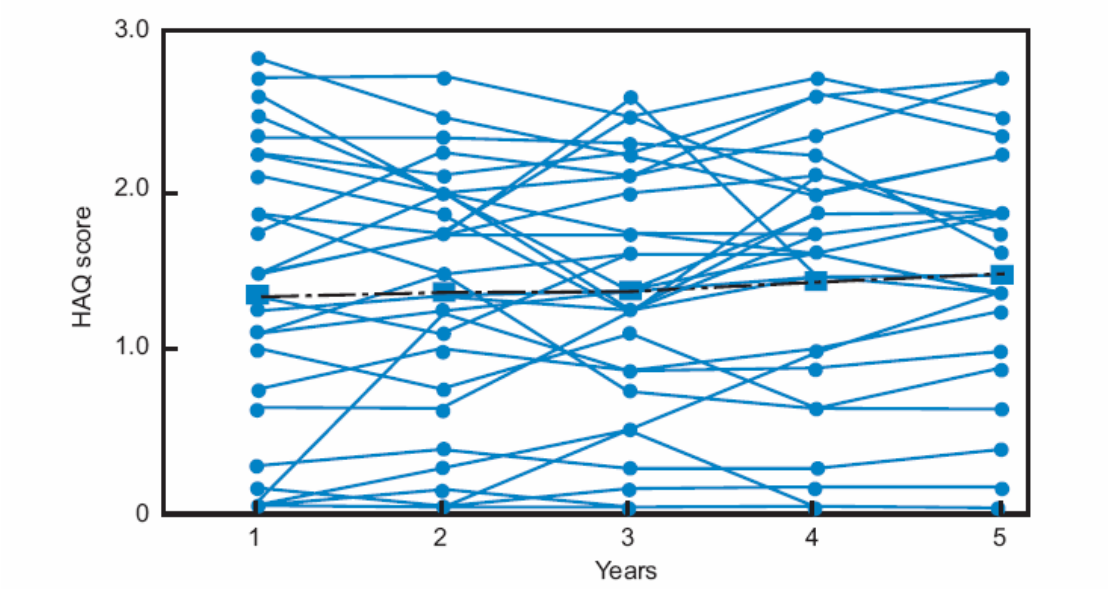


Figure 2. Individual (solid lines) and mean (broken lines) changes in Health Assessment Questionnaire (HAQ) score in 30 rheumatoid arthritis patients followed for 4 years in a single unit.

Figure 4-4: Variations in HAQ scores for individual patients reproduced from Scott[58]

The NOAR report[62] provided by the company is valuable in offering greater insight into how HAQ scores change over time. Table 4-6 reproduces a summary of annual movements of patients between six HAQ score bands over a 5-year period. Of particular note is that the proportion of patients remaining within the same band from one year to the next is remarkably low in the four intermediate bands (35-44%), and also that 39% of patients in the highest band show improvement within 12 months.

Table 4-6: Frequency distribution of annual changes in HAQ scores for 1246 early-stage patients followed for 5 years (from Table 4 of NOAR report[62]

Frequency of transitions in each cell

		HAQ band at time (x + 1) (12 months later)					
		0.00 - 0.375	0.50 - 0.875	1.00 - 1.375	1.50 - 1.875	2.00 - 2.375	2.50 - 3.00
HAQ band at time (x)	0.00 - 0.375	75.8	17.3	4.6	1.8	0.3	0.1
	0.50 - 0.875	31.8	39.1	20.6	6.8	1.6	0.2
	1.00 - 1.375	14.0	24.5	35.2	18.0	7.2	1.1
	1.50 - 1.875	8.6	7.3	22.3	36.7	20.7	4.5
	2.00 - 2.375	3.4	3.6	6.7	24.4	43.5	18.4
	2.50 - 3.00	0.6	0.6	1.2	9.2	27.2	61.3

The impact of this degree of variability can be gauged by repeatedly applying the transition rates shown in the NOAR report[62] to a specified cohort of patients to simulate trends over several years; Figure 4-5 shows the effects for the NOAR early-stage RA cohort, and also for a more severe illustrative cohort chosen with a starting mean HAQ of 1.88. The estimated score for the NOAR[62] cohort increases steadily but non-linearly with a decreasing rate each year until converging at a 'steady-state' level after 15-20 years. Clearly it is unlikely that this will be an accurate estimate of the long-term prognosis, since the transition rates were only measured over a 5 year period and are likely to change in later years. Nonetheless this shows that we should expect to see large changes in the early years, reducing in size over time. The second line (with initial mean HAQ of 1.88) shows a downward non-linear convergence to the same steady-state value (which is wholly determined by the NOAR[62] probabilities). Clearly this is not realistic, and demonstrates that it is not appropriate to use evidence of progression rates in early-stage RA patients as the basis for estimating long-term changes in the later stages of disease, since both the transition probabilities and the initial case-mix will be quite different.

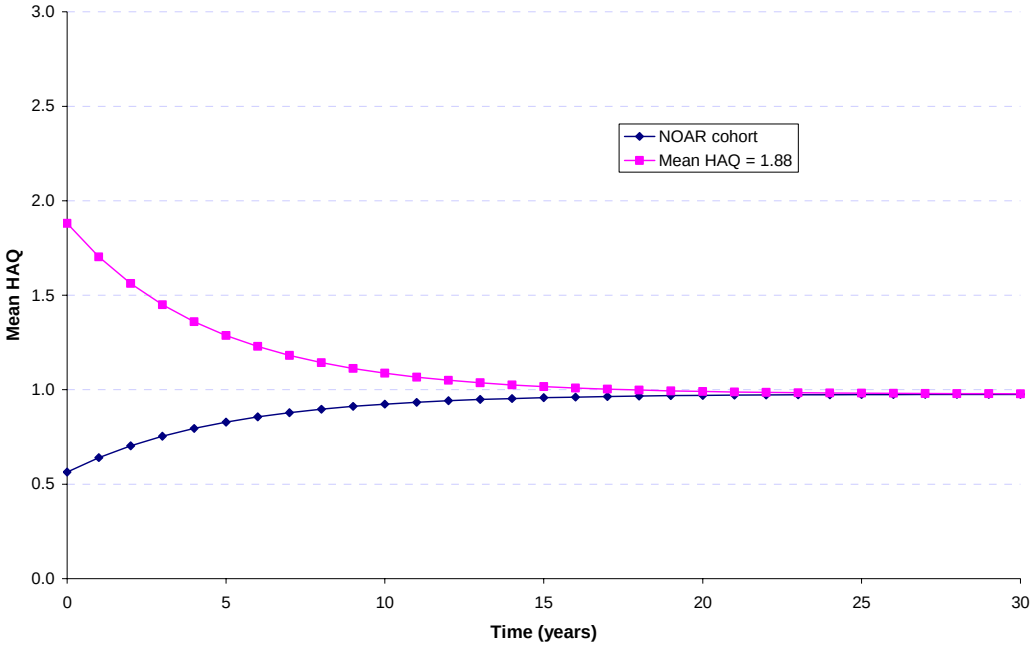


Figure 4-5: Illustration of non-linear trends in mean HAQ scores using NOAR transition rates

In Figure 4-6 we show a scenario much closer to that implied by the submitted model - with an initial mean HAQ of 1.88, and transition probabilities strongly weighted towards steady deterioration in function year by year. Even here it is apparent that a linear trend would not be considered a realistic basis for representing the long-term progression of loss of functional capacity as measured by HAQ. We would expect progression rates to be diminishing steadily over time, and stabilising at a mean value rather less than the maximum of the scale. This contrasts sharply with the model assumptions:

- that all patients progress to the maximum score (3.0);

- that the same numerical increase in HAQ will occur annually during treatment; and

- that in the long-term, progression rates on palliative treatments will be up to four times the earlier rate.

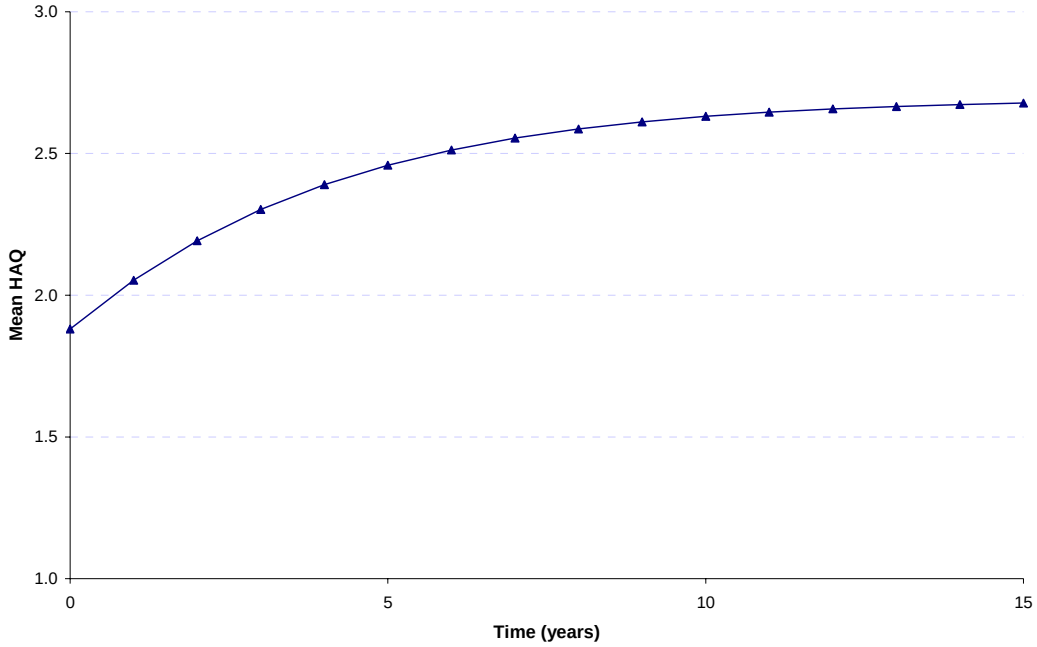


Figure 4-6: Illustration of projected trend in mean HAQ using transition probabilities weighted strongly toward progression

Evidence for HAQ progression rate on DMARDs

The submitted model features a progression rate for HAQ scores of 0.034 per year (or 0.017 per six-months), obtained from Table 6 in Scott and Garrood's review paper[58] published in 2000. This mentions results from nine observational studies of different types and durations. Scott and Garrood[58] combined trend rates they obtained from each study to obtain an 'average' rate, though without a description of how the calculation was carried out. The importance of this parameter to the model results warranted the ERG revisiting the cited studies. Table 4-7 summarises our findings which differ in important respects from those of Scott and Garrood.[58] Various factual and interpretive corrections were identified, and we chose to prefer long-term rates over early-stage disease rates (the latter being unrepresentative of the patient cohort being modelled). It also seemed important to separate cross-sectional studies, from those in which patients were followed up over extended time periods, since cross-sectional studies are more susceptible to case-mix bias. Weighted mean rates were then re-estimated separately from cross-sectional and longitudinal studies, providing revised values both of which were considerably smaller than the 'average' of Scott and Garrood.[58] Moreover, the cross-sectional studies yielded an estimate twice the size of the longitudinal studies.

Table 4-7: Evidence of long-term HAQ progression in RA - studies used by Scott & Garrood[58] in estimating average annual progression rate

Study	Data period	Cases	Study Type	Rates quoted	Rates calculated	Comments
Wolfe[63]	1976-90	561 in total: 264 *0-2* years disease 143 *2-7* years disease 67 *7-12* years disease 57 *12-17* years disease 30 *17-22* years disease	Cross-sectional study of new cases followed-up for 2 (early disease) or 5 years on treatment	0.020 pa	0.0159 pa	Weighted average linear trend of 0.0167 per year, from unadjusted data (Table 2). Weighted average linear trend of 0.0159 per year, from adjusted data (Fig 1). Authors recognise that this is a non-representative sample of patients presenting with serious needs, so long-term differences are not representative of true natural history of disease. Also confounded by treatments given during observation period.
Lassere[64]	1992	358 seen in last 2 years (excluding those who had died, poor English, cognitively impaired, and non-respondents)	Cross-sectional study	0.045 pa	0.0369 pa	No information on duration sample sizes - equal sizes assumed. Linear trend in means gives 0.0397 pa for all points, and 0.0369 pa excluding early stage group (off linear). Linear trend in medians gives 0.0544 pa for all points, and 0.0449 pa excluding early stage group (off linear). Several likely sources of bias present.
Sherrer[43]	1966-82	681 new cases followed for average of 11.9 years (excluding 281 deaths and 81 lost to follow-up). Mean duration of illness at start 10 years	Cross-sectional	0.072 pa	0.0217 pa	Longitudinal regression analysis did not identify duration of disease as a significant indicator for HAQ - no HAQ comparison possible (no HAQ at baseline). Cross-sectional (unadjusted) trends in HAQ by duration of disease: - weighted linear trend gives 0.0367 pa - wtd linear trend for duration >15 years gives 0.0217 pa
Ward 1[65]	1979-91	282 volunteers with >=2 years RA, followed up for 10 years	Prospective longitudinal	0.012 pa	0.019 pa without specialist care, 0.007 pa with specialist care.	Authors report separate linear trends in adjusted HAQ for 3 sub-groups: - no specialist care 0.019 pa - intermittent specialist care 0.019 pa - continued specialist care 0.007 pa Overall wtd average is 0.0161 pa

Gardiner[66]	1984-1989	175 IP and OP patients seen in 1 month in 1984	Prospective longitudinal	0.030 pa	0.036 pa	Mean increase in HAQ of 0.18 (S.D. 0.66) over 5 years.
Callahan[67]	1984-91	100 US OP patients	Prospective longitudinal	-0.006 pa	-0.006 pa	Mean increase in MHAQ of -0.06 over 5 years.
Leymarie[68]	1991-?	370 French & Dutch patients with duration <5yrs	Prospective longitudinal	0.000 pa	0.000 pa	Annual assessment over 2 years. For 34% HAQ was worse, 39% stable, 27% improved at 2 years. Mean HAQ 1.06. Mean duration of disease 2.1 years.
Ward 2[69]	1981-94	182 volunteers adults with minimum 1.5 years follow-up	Prospective longitudinal	0.017 pa	0.0163 pa	Baseline duration 13.7 years, 10.4 years follow-up, mean HAQ 1.02
Munro[70]	1986-95	160 completing patients of 440 original started on gold therapy	Prospective longitudinal	0.119 pa	- means not estimable	5 year follow-up. Only median HAQ values given. 160 cases (not 440 as stated by Scott). Non-homogeneous sub-groups. Need to discount trends for treatment effect in first year.
Overall weighted average		**1603 patients**	**Cross-sectional studies**		**0.023 pa**	
Overall weighted average		**1109 patients**	**Longitudinal studies**		**0.012 pa**	

pa: per annum:

Evidence for HAQ progression rate on palliative care

Tracing back the reference given in support of the submitted model's assumed long-term progression rate on palliative care (0.13 per year), we find that the source is a paper by Young et al. published in 2000 reporting results of a 5 year observational study of early RA patients presenting at nine UK rheumatology hospitals (the ERAS study[59]). It is notable that Young[59] does not anywhere give results for mean HAQ for any time point nor for any sub-group of the cohort. The estimate for the change in HAQ appears to have been derived by Bansback et al 2005[38] by manipulation of information extracted from figure 2 of the ERAS[59] paper, which shows only medians, and for functional groups at the *end* of the 5 year period. This is fundamentally flawed since it selects out those patients known to end the study with the worst scores, and therefore with the greatest scope for deterioration. The use of medians on a seriously biased end-point subset, as the basis for inferring a predictive temporal trend in mean HAQ scores, is completely inappropriate and without merit.

Even if the derivation of the long-term progression rate used in the model could be justified its application to patients receiving only palliative treatments is questionable. The ERAS paper[59] gives no indication of how many (if any) of the 84% who had started treatment with one or more DMARDs had exhausted all treatment options within 5 years, nor how many (if any) of those were included in the end-point sub-group used as the basis for the palliative care progression rate. Thus it is difficult to see how these patients (all with duration of disease less than 6 years) could be considered a suitable source for projecting the experience of patients with RA of duration 10-20 years or more.

Summary concerning the use of HAQ scores

Assumptions about the nature and extent of progressive functional disability, as measured by mean HAQ scores, are highly influential in the submitted model, especially in determining the size of health utility gains from use of rituximab.

The nature of the closed HAQ scale and the natural variability of HAQ scores (both patient and disease related) suggest that the model assumption of a fixed increment in HAQ score per time period, irrespective of the current HAQ score, is simplistic and misleading especially over extended projection periods.

The analysis of observational studies cited to support a progression rate of 0.034 per annum whilst on active treatment fails either to give an accurate representation of the quoted sources, or to recognise the incompatibility of data derived from cross-sectional and longitudinal studies. The best estimate derived from these studies by the ERG is an average progression rate of 0.012 per annum.

The ERAS paper[59] which was the original source for the HAQ progression rate on palliative care is not an appropriate basis for estimating the experience of such patients. In addition the derivation of a mean rate from end-result median values is erroneous and untrustworthy. No evidence has been provided to support the idea that a substantially different long-term progression rate should apply when the DMARD options are exhausted.

The ERG view is that HAQ progression would be best estimated by use of a simple non-linear trend line, consistent with a long-term stable maximum mean value a little below the scale maximum (say about 2.75) to reflect the inherent variability in HAQ measurement.

4.3.4 Resource use and costs

Therapy costs

In the submitted model the costs of treatment (drugs, administration and monitoring) are averaged over the estimated mean duration of treatment, and then the average applied to patients on treatment in any given ½ year period. Where there are significant initial additional costs - loading doses for some drugs, and generally for administration costs - this process depends on estimated mean duration of treatment. Unfortunately, the model rounds each estimated mean to the nearest six-month time point, which can lead to unjustified cost differences between treatments. In particular, rituximab, etanercept and adalimumab all have a duration of 4.254 years increased to 4.5 years, whereas leflunomide is reduced from 4.10 to 4.0 years and ciclosporin from 1.70 to 1.50 years.

All costs have been amended by the ERG to ensure that the derivation of annual costs is consistent with the rounded durations of treatment used in the model to generate outcome effects. In addition, the Weibull distributions underlying the mean durations have been employed to reflect fully the expected timing of treatment withdrawal. The dose levels and adjuvant treatments have been checked and where necessary amended in accord with British National Formula [71] entries. Also, the range and frequency of monitoring activities for each drug have been checked against new BSR guidelines, leading to amendments to the frequency of out-patient visits in the first year of treatment. For infliximab, the drug cost has been based on the distribution of body weight from the clinical trial, allowing for 50% of wastage from part-used vials to be saved by vial sharing. For the cost of administration of rituximab and infliximab we have used the relevant NHS tariff category.

For most treatments the re-estimated treatment costs are little different from those used in the submitted model. However, the estimated costs of both rituximab and infliximab are increased substantially by these changes: by £637 per treatment-year in the case of rituximab, and by £1,850 per treatment-year for infliximab (Table 4-8).

Table 4-8: ERG amendments to model treatment costs

Therapy option	Annual cost in model	Re-estimated annual cost	Difference	Comments
Rituximab + MTX	£6,211	£6,848	£637 (+10.3%)	Adjusted to Weibull distribution - mean 4.5 years Administration costed as regular attender visit for "Chemotherapy with musculoskeletal primary diagnosis" (RDH98) - £267 [2005/6 NHS Reference Costs for rheumatology][72]
Etanercept + MTX	£10,656	£10,612	-£44 (-0.4%)	Adjusted to Weibull distribution - mean 4.5 years
Infliximab + MTX	£8,882	£10,732	£1,850 (+20.8%)	Adjusted to Weibull distribution - mean 2.5 years Estimated drug use based on weight distribution Assume 50% of wastage avoided by vial sharing Administration costed as for rituximab
Adalimumab + MTX	£10,514	£10,470	-£44 (-0.4%)	Adjusted to Weibull distribution - mean 4.5 years
Leflunomide	£1,733	£1,784	£51 (+2.9%)	Adjusted to Weibull distribution - mean 4.0 years Include folate supplementation (BNF)[71] 6% of patients on reduced treatment frequency[29] Fewer OP visits in year 1
Intra-muscular gold	£2,694	£2,547	-£147 (-5.4%)	Adjusted to Weibull distribution - mean 4.0 years All doses (after first) 50mg, reducing frequency progressively from weekly to 4-weekly (BNF)[71]
Ciclosporin	£3,954	£4,194	£240 (+6.1%)	Adjusted to Weibull distribution - mean 1.5 years Fewer OP visits in year 1
Palliative care (MTX)	£1,865	£1,755	-£110 (-5.9%)	Addition of folic acid Reduction in OP visits in year 1, averaged over 15 years survival

Disease costs

The submitted model only includes direct medical costs relating to in-patient admissions. No additional out-patient visits, GP consultations or prescribed medications are included in the cost estimates. The NOAR report[62] is the basis used for costing in-patient episodes, and provides the annual rate of days in hospital stratified by HAQ band. Clearly the use of NOAR[62] results involves the assumption that hospital admissions are determined solely by the HAQ score of patients, and does not change over time. If this assumption is not accepted then these data cannot be used for model costing since the NOAR patients are quite unlike those simulated in the model.

It was noted by the ERG at an early stage that there was no explicit mention of joint replacement surgery in the model, and that the costing of in-patient days with a single bed-day cost would miss the substantial additional costs of any major procedures carried out. In response to the ERG enquiry, the model authors provided an amended set of cost parameters for the model. However, we are of the view that these amended calculations involve some double-counting of hospital 'hotel' costs. We have therefore revised the calculations, and incorporated 2005/6 NHS reference costs.[72]

Examination of the program code also seems to suggest that there is an error in that costs calculated from annual event rates in the NOAR report[62] are applied per 6-monthly cycle thereby doubling the true in-patient costs incurred. The two sets of Roche parameters and the ERG amended calculations are shown in Table 4-9.

Table 4-9: Amendments to model disease-related costs: in-patient cost per patient-cycle

HAQ scores	0 <0.5	0.6<1	1.1<1.5	1.6<2.0	2.1<2.6	2.6<3.0
Roche original estimate	£63	£31	£123	£174	£448	£1,003
Roche revised estimate	£80	£77	£226	£340	£654	£1266
ERG revised estimate	£49	£25	£96	£136	£351	£786

4.3.5 Treatment effect prior to treatment failure

In the submitted model the assumption is made that any patient who responds to a therapy benefits from an immediate reduction in HAQ score, which is sustained throughout the period on treatment. When treatment ceases for any reason the benefit is withdrawn and the HAQ returns to the previous level, altered only by the underlying rate of deterioration (A in Figure

4-7). This may appear reasonable in situations where the cause of discontinuation is the sudden appearance of a significant adverse event or drug reaction in a patient previously in a stable condition and good response. However, not all patients terminate a treatment for this reason. In many the stated reason is "loss of efficacy", and it is reasonable to expect this to be a more gradual process becoming increasingly apparent to patient and physician or time until most or all previous gains are lost (closer to B in Figure 4-7).

To resolve uncertainty on this issue, the ERG requested that Roche provide

> "..a detailed table of withdrawals by week and by reason (serious AE, reaction, clinical advice, patient request, lack of efficacy, etc)."

Unfortunately, the response received was as follows:

> "Data relating to withdrawals by week of study is not available from the clinical study report, only data by the time of last dose. These details are provided in a separate PDF attachment, labelled appendix 5."

Appendix 5 included 53 patients in the rituximab arm who withdrew prior to week 24, in only eight of whom was an adverse event or illness cited as the reason - presumably lack of response being the main cause of early withdrawals. Given the design of the trial it is unlikely that it could have furnished useful evidence to help determine how initial efficacy is lost. Information was also provided by the company on the time to second treatment with rituximab, but these data also do not offer any help in tracking the extent of response over time.

A brief literature search on this question proved largely unproductive. However, one small registry study of patients switching between infliximab and etanercept in Sweden[73] gave rise to some suggestive observations. For patients treated with etanercept, it was noted that most switched treatment due to loss of efficacy, and that the DAS28 score at that time was close to the original baseline level (prior to starting etanercept). By contrast, for most infliximab patients the reason for switching was adverse events but at the time of switching "the response had become somewhat less".

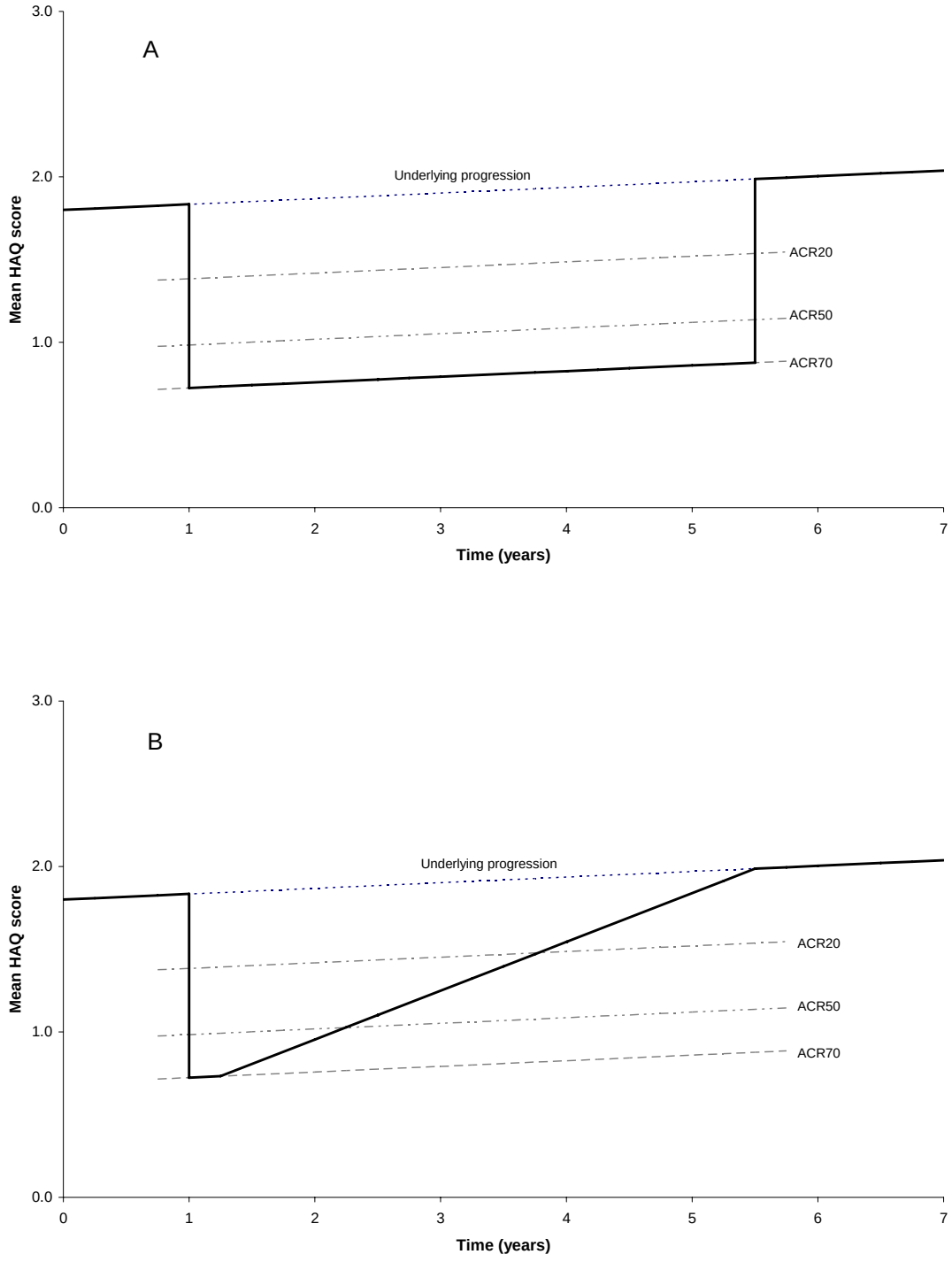


Figure 4-7: HAQ profiles under treatment response - A in submitted model, B with steady loss of effect

Thus it appears that there may be grounds to consider that the model assumption of sudden deterioration in response (case A in Figure 4-7) is optimistic, and that some measurable loss of efficacy should be anticipated where 'loss of efficacy' is the cited reason for ending treating, and may also be relevant in cases of adverse events. We have therefore attempted to

replicate a 'worst case' scenario (similar to case B in Figure 4-7) by reducing by half the HAQ gains attributable to each of the three degrees of response. This is a rather crude approximation, but is necessary as the structure of the company model does not easily lend itself to a more sophisticated adjustment.

4.3.6 Other model issues

Mortality risk and HAQ

The submitted model employs the relationship between HAQ score and mortality risk derived by Wolfe in 1994[74] and used by Barton et al[56]:

$$\text{Mortality relative risk} = 1.33^{\text{HAQ}}$$

Although there are wide confidence intervals on this parameter estimate (1.099 - 1.61), it appears that this uncertainty has only a limited effect on economic results. Substituting the upper and lower confidence limits into the validation spreadsheet had the effect of altering the ICER from £15,235/QALY to either £14,151 or £16,647. It therefore appears that this model assumption is unlikely to be important in determining cost effectiveness.

Utility and HAQ

In the company submission, three linear models are described relating utility estimates to HAQ scores. The authors explain their preference for the Bansback[38] equation on the grounds of its much larger sample size and being collected in a trial setting. In practice, model estimates obtained with the Bansback[38] model fall midway between those obtained with the other two equations. It therefore appears to be a reasonable assumption, and not likely to lead to any pronounced bias in results.

Cycle length

The submitted model is structured on the assumption that regular patient reviews take place at approximately six-monthly intervals. This simplification of real-life variability is probably acceptable for patients in a stable condition. However, this may not be so appropriate when a patient is receiving a new treatment. Clinical advice suggests that clinicians would normally expect to see a response within three months, and would consider switching to an alternative agent well before six months. It is not clear how a reduction in trial time to three months would affect economic results, since the option to use two separate time units is not available in the company model.

Duration of effective treatment

The mean time that each compound is assumed to be effective may be an important element in determining cost-effectiveness. However, the evidence base for these parameters is poor and relies on multiple sources that may not be comparable. A recent analysis of records from the UK General Practice Research Database[75] appears promising in helping to improve some of these estimates, using a single source, but unfortunately the GPRD data should not be considered reliable for this purpose. The patient records cover a long period (1987-2002) during which there were considerable changes in the way care was provided and in the drugs available for prescription. In particular it is clear that use of MTX increased 17-fold in that time, while use of gold fell by more than 75%. Since there is no information available to indicate the reasons for treatment changes, it is very likely that durations of treatment obtained from GPRD data cannot be considered unbiased measures of relative durability of effect.

Sensitivity analyses of variations in duration of treatment parameters provided by the company seem to indicate that increased values can give rise to moderate increases in ICER results. It is likely that this is closely related to delays in the inception of the palliative treatment phase, when the much higher rate of HAQ progression is assumed. If the higher progression rate is excluded it is probable that economic results will be less sensitive to the duration of treatment.

Interval between rituximab doses

At first sight the company submission is confusing in relation to the duration of time between doses of rituximab in patients who respond to treatment. In Section 1.8 the mean time between first and second treatments is given as 33.2 weeks (232 days), and between second and third treatments as 32.2 weeks (225 days). However, in the executive summary it is shown as 301 days. In Table 16 it is shown as 307 days. For the purposes of economic modelling a figure of 9 months (293 days) is adopted as representing the midpoint of the summary of product characteristics quoted range (6-12 months). Disregarding the figures given in Section 1.8 as uncorrected observations, it appears that the estimate of 307 days obtained by Kaplan-Meier (K-M) analysis (provided in response to the ERG's request) is probably more reliable. This might be construed as indicating that the model value of 9 months is unduly conservative (being close to the lower confidence limit of the K-M estimated mean). Therefore, the ERG has included a sensitivity analysis in which the costs of rituximab are re-estimated assuming 307 days between doses.

Treatment sequencing

The two scenarios used to generate the results shown in the company submission feature 3 or 5 active treatment steps other than rituximab and palliative care, in a fixed sequence. No justification is given for the ordering of these agents, which may well not represent the most cost-effective strategy, and may have implications for the cost effectiveness of rituximab. The ERG has therefore used the model to gain insight into the relative merits of alternative sequences.

PSA

In response to ERG criticism of the original PSA results and methods, the model authors provided an amended analysis using different measures of uncertainty. The original model included uncertainty on five different sets of parameters (assumed to be mutually independent):

1) probabilities of patients falling into one of four response categories, for each DMARD;

2) size of reduction in HAQ scores applicable to the four response levels;

3) progression rates for HAQ scores on DMARDs and on palliative care;

4) in-patient resource use rates; and

5) mean effective time of treatment for each DMARD.

In the amended model the PSA logic and/or parameters have been correctly amended for items 2, 3 and 4. The logic in respect of item 1 remains erroneous, since it does not accurately reflect that a choice between four categories only involves three degrees of freedom. The authors have used a 'rounding down' adjustment to correct for anomalous totals but this introduces substantial bias for some values. In the case of item 5, the authors were unable to obtain the information necessary to allow the correct calculations to be carried out, and have therefore excluded this factor from the PSA presented.

In should also be noted that no attempt was made to include the effect of either age or baseline HAQ scores in the PSA, arguing that these were better handled through one-way sensitivity analyses.

In the light of the limited scope of the revised PSA and the continued presence of logic errors, the ERG do not believe that the PSA results presented in January 2007 could be considered reliable. Unfortunately, it is not possible for the ERG to modify the company model to correct these problems in the time available.

4.3.7 ERG amended economic results

In order to assess the impact of the most important logic errors and alternative interpretation of evidence, several amendments have been made to the revised submitted model as follows:

- corrected mortality probabilities have been introduced;

- the two HAQ progression rates have been replaced by a single linear progression rate, from our re-analysis of the papers cited by Scott and Garrood[58];

- the costs of drugs, administration and monitoring have been amended; and

- revised disease cost parameter values have been used, correcting the detected error and incorporating the cost of joint replacement surgery.

With these changes in place, results are shown in Table 4-10 for the two scenarios presented in the company submission, set alongside the results obtained with the unamended revised model. In addition, these calculations are repeated using only half the HAQ treatment gains in the model to illustrate the differences between a 'worse case' assumption to be set against the 'best case' assumption of no attenuation of effect incorporated in the company submitted model, and also using reduced rituximab costs due to a longer interval between doses.

The ERG changes to the company model result in substantial increases in the cost-effectiveness ratio for both scenarios, to values somewhat beyond those normally considered cost effective (£40,900 and £32,900 per QALY gained). The 'worst case' amendment for graduated loss of efficacy doubles these ratios, as outcome gains are halved. A longer interval between doses yields modest improvements in the ICERs (to £37,000 and £28,600 respectively).

Finally, we present a 2-way sensitivity analysis for different values of the mean age and mean baseline HAQ score (Table 4-11). Results are not very sensitive to assumptions about either the initial age of patients or the baseline HAQ score, except that cost effectiveness is worsened for the very elderly.

Table 4-10: Cost-effectiveness results incorporating ERG corrections/amendments

	Rituximab simulation			Comparator simulation			Incremental			ICER
Scenario	**Life-years**	**QALYs**	**Costs**	**Life-years**	**QALYs**	**Costs**	**Life-years**	**QALYs**	**Costs**	**Cost/ QALY**
Base case (no TNFi) - revised model	12.747	3.045	£41,279	12.568	2.318	£30,588	0.179	0.728	£10,691	£14,694
Base case - revised model + ERG changes	**15.940**	**5.489**	**£44,636**	**15.890**	**5.157**	**£31,069**	**0.050**	**0.332**	**£13,567**	**£40,873**
Base case - ERG changes - 50% HAQ gains	15.792	4.626	£44,793	15.767	4.456	£31,212	0.025	0.169	£13,581	£80,198
Base case - ERG changes + longer interval	15.940	5.489	£43,351	15.890	5.157	£31,069	0.050	0.332	£12,282	£37,002
Alternate (TNFi) - revised model	13.028	3.963	£69,901	12.866	3.457	£63,996	0.162	0.506	£5,905	£11,666
Alternate - revised model + ERG changes	**15.999**	**5.954**	**£77,701**	**15.947**	**5.684**	**£68,853**	**0.053**	**0.269**	**£8,847**	**£32,855**
Alternate – ERG changes - 50% HAQ gains	15.843	4.870	£77,800	15.823	4.737	£69,070	0.021	0.133	£8,730	£65,558
Alternate – ERG changes + longer interval	15.999	5.948	£73,173	15.948	5.678	£65,456	0.051	0.270	£7,717	£28,553

N.B. All results are discounted at 3.5% per annum (pa)

Table 4-11: Two-way sensitivity analysis of cost-effectiveness by mean age and baseline HAQ score

Mean HAQ	**1.40**			**1.88**			**2.20**		
Mean Age	**IC**	**IQ**	**ICER**	**IC**	**IQ**	**ICER**	**IC**	**IQ**	**ICER**
30	£13,662	0.338	£40,401	£13,620	0.337	£40,382	£13,339	0.336	£39,649
52.2	£13,563	0.329	£41,161	**£13,567**	**0.332**	**£40,873**	£13,300	0.327	£40,716
65	£13,203	0.318	£41,470	£13,126	0.308	£42,575	£12,816	0.297	£43,159
80	£11,387	0.262	£43,523	£11,021	0.225	£48,961	£10,672	0.208	£51,204

IC = Incremental cost IQ = Incremental QALYs ICER = Incremental cost per QALY gained

Prior TNFi

The significant difference attributable to the number of prior TNFi (Section 3.3.4) suggests that it is prudent to consider the sensitivity of economic results to this factor i.e. are ICERs much different if sub-groups are considered separately, rather than combined as in the base case? Substituting response values for the single prior TNFi sub-group into the ACR20 indirect comparisons parts of the submitted model leads to a small difference in the results: survival is increased slightly in both arms, leading to small increases in QALYs and costs. However, as rituximab costs increase faster than those in the comparator, the ICER worsens slightly from £40,873 to £41,088 per QALY gained. Clearly, this is insufficient to warrant further consideration of the number of prior TNFi used.

Treatment sequencing

The initial stage in considering optimum treatment sequencing was to compare each drug as a sole intervention prior to palliative care, with palliative care alone. The results are summarised in Table 4-12, in the order of relative cost effectiveness. This suggests that the sequence adopted for the company submission may not be optimal.

Table 4-12: Model comparison of active treatments before palliative care to palliative care alone, ordered by reducing cost effectiveness

Treatment	Incremental cost per patient	Incremental QALYs per patient	ICER
Leflunomide	-£304	0.204	-£1,491
Gold	£711	0.118	£6,016
Ciclosporin	£1,513	0.054	£27,896
Rituximab	£13,677	0.375	£36,476
Adalimumab	£22,667	0.399	£56,825
Etanercept	£26,398	0.418	£63,098
Infliximab	£13,967	0.205	£68,093

A further set of tests was carried out by successively comparing the performance of pairs of drugs when their order was reversed. The outcome of this investigation identified three groups of broadly comparable economic performance:

- best performers: leflunomide and rituximab
- middle ranking: ciclosporin and gold
- worst performers: TNFi

The final stage of testing involved using the full treatment sequences specified in the model, and examining whether there were grounds to use any drug within a group before the other(s) in the group. This led to the conclusion that ciclosporin and gold could not be distinguished on either cost or outcome differences, nor could either of the TNFi 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), and therefore should normally be given prior to rituximab.

Using this resequencing of treatments results in slightly improved ICERs for rituximab: for the base case (no TNFi used) yields £37,028/QALY reduced from £40,873, and the alternate scenario with TNFi becomes £32,259 instead of £32,855.

4.4 Summary of cost-effectiveness evidence

4.4.1 Economic evaluation results

Base case: company

- The company report a revised ICER of £14,694 per QALY gained for the NICE recommended scenario of rituximab versus no rituximab
- The company report a revised ICER of £11,666 per QALY gained for the sequential use of TNFi scenario of rituximab versus no rituximab
- Limited PSA results (scatterplots and cost-effectiveness analysis curves) are presented by the company

Base case: ERG

- A number of key issues and parameters in the model do not seem to be clinically and/or economically justified, particularly in relation to long-term progression and its effect on HAQ scores
- After model assumptions are adjusted to more realistic estimates, the ICER for the NICE recommended scenario ranges from £37,002 per QALY gained to £80,198 per QALY gained and the ICER for the sequential use of TNFi ranges from £28,553 per QALYgained to £65,558 per QALY gained
- Varying mean age and baseline HAQ score had very little effect on costs and cost effectiveness

4.4.2 Economic issues

- Roche submitted a revised economic model after discussion with the ERG and NICE. The revised model included significant errors (mortality rates, estimation of in-patient costs) and issues (use of evidence for progression rates for HAQ scores, calculation of treatment costs, duration of effective treatment for each of the active agents considered)
- The ERG identified other influential issues: whether the size of the benefit from each treatment is overstated, because loss of efficacy is assumed to be instantaneous rather than cumulative; whether the assumed mean time between doses of rituximab is too conservative; whether the treatment sequencing in the submitted scenarios is sub-optimal
- The company PSAs (original and revised), due to limitations described by the ERG, are considered to be unreliable aids to decision-making

5 Discussion

The company submission presents a case for the use of rituximab in adult patients with severe RA. In their analysis of the decision problem, the company describes two different rituximab management strategies for patients. The first scenario is described as a "NICE recommended" strategy as it allows patients to fail on one TNFi, receive rituximab and then go on to receive a series of DMARD regimens (excluding any subsequent treatment with a TNFi). Patients are not permitted to receive a second TNFi in this scenario. The second scenario is described as a "sequential TNFi" strategy as it allows patients to fail on one TNFi, receive rituximab, receive a second and third TNFi before going on to receive a series of DMARD therapies. In both cases, the comparator is the same scenario without rituximab. The company have presented both scenarios to reflect their belief that although NICE does not recommend the sequential use of TNFi, evidence from clinical practice suggests that a proportion of patients in the NHS in England and Wales are nonetheless receiving sequential TNFi treatment.

The systematic literature review performed by the company did not yield any clinical studies which compare rituximab with an appropriate comparator (i.e. leflunomide or a second or third TNFi) to inform either of the rituximab scenarios described by the company. The literature search identified a single RCT (REFLEX trial) conducted by the company comparing rituximab plus MTX versus placebo plus MTX. All of the patients in the REFLEX trial had failed at least one prior TNFi. However, whether or not the patients in the REFLEX trial match the patients in either of the scenarios set out by the company is debateable. Forty percent of the patients in the REFLEX study had received two or more TNFi and the most popular TNFi received was infliximab. In the "NICE recommended" scenario patients are not allowed to receive more than one TNFi. Also, in the two scenarios proposed by the company etanercept is assumed to be the first TNFi of choice for all patients.

Results from the REFLEX trial furnish the principal clinical evidence presented in the company submission. The REFLEX trial appears to have been a well-conducted RCT, the results of which seem to demonstrate that rituximab plus MTX is more clinically effective than placebo plus MTX. At 24 and 48 weeks, ACR20/50/70 responses are greater in the active arm (rituximab) compared with placebo. As the patients who would be eligible to receive rituximab are difficult to treat, having severe disabling disease with marked impairment of quality of life, the results of the REFLEX trial are convincing. As with all biologics, strict surveillance and monitoring of the use of rituximab, during treatment and post-treatment, is merited.

Unfortunately, the clinical evidence from the REFLEX trial does not allow the company to answer the questions raised in their statement of the decision problem. The REFLEX trial provides evidence on the comparison of rituximab plus MTX versus placebo plus MTX. However, it does not provide any answers to the question of whether or not rituximab is more clinically effective and/or cost effective when compared to leflunomide or a second or third TNFi.

Although it was not possible to replicate the clinical searches conducted by the company, the ERG is confident that the company submission identified all relevant clinical studies. Consequently, it was appropriate for the company to undertake an indirect comparison exercise to identify absolute efficacy values for use in the economic evaluation. However, the ERG is not confident that the results of the indirect comparison exercise are valid. Firstly, it is not clear from the evidence presented by the company that all relevant clinical studies have been included in the exercise. Secondly, the company does not present a clear rationale for their choice of indirect comparison method. Thirdly, the indirect comparison method used to adjust the ACR responses uses a single value for the reference placebo. A more appropriate method would have been to calculate a range of reference placebo (RP) rates based on trial population characteristics, as this would have allowed for any heterogeneity between the trials to be explored. The methods and results of the indirect comparison are presented in the clinical effectiveness section of the company submission but are not discussed until the cost-effectiveness section.

In contrast to the clinical section of the company submission, the cost-effectiveness section concentrates wholly on the comparison of the two proposed management strategies with and without rituximab. In their economic analysis, the company conclude that rituximab should be considered to be a cost-effective treatment option in RA. For the "NICE recommended" scenario, the ICER is £14,694 per QALY gained, and for the "sequential TNFi" scenario, the ICER is £11,666 per QALY gained. However, there are a number of clinical and economic issues that call into question the validity of these claims, and the credibility of the ICERs generated.

The model submitted in support of the application was a bespoke 'micro-simulation' written in Visual Basic code within a Microsoft Excel workbook. Certain features of the model construction and coding gave rise to misgivings for the ERG relating to its reliability, and these were compounded by the complexity of the code used, and the lack of explanatory documentation. The company submitted a revised version of their model incorporating amendments that addressed some of the ERG's concerns. However, the ERG felt obliged to carry out a simple validation exercise before they could confirm that the model logic had been consistently implemented.

However, on detailed examination of the revised economic model submitted by the company, several significant additional errors and issues were identified by the ERG. In particular, these relate to errors in mortality rates, the evidence base for progression rates for HAQ scores, the calculation of treatment costs and errors/omissions in the estimation of in-patient costs. Although the revised model did address some of the failings in its generation of PSA results, the algorithm remains very limited in its coverage of parameter uncertainty, and therefore is not considered to be a reliable aid to decision-making in this instance.

The ERG also identified other key issues as potentially influencing model results, and carried out sensitivity analyses to show their impact on model results. For example, if it is assumed that the size of benefit from each treatment is overstated (because loss of efficacy is assumed to be instantaneous rather than progressive), the estimated cost-utility ratio may increase substantially - and may even double.

The ERG also included amendments that favour the company's case for the use of rituximab: assuming a longer mean time between doses and re-sequencing the treatment options in the scenarios described both lead to slightly improved outcomes for rituximab.

Using alternative ERG assumptions and parameters in the model has the effect of generating substantially worsened cost-effectiveness results for the two management scenarios described by the company in their submission. The ICER for the "NICE recommended" scenario ranges from £37,002 per QALY gained to £80,198 per QALY gained and the ICER for the "sequential TNFi" scenario ranges from £28,553 per QALY gained to £65,558 per QALY gained. No patient sub-groups could be identified which exhibit significantly better economic results than the whole cohort.

The consequences of these corrections and amendments is that economic results for the use of rituximab no longer appear as unequivocally advantageous as suggested in the company submission, and may more reasonably be termed 'borderline' at best.

The ERG also used the revised and ERG-amended model to investigate the influence of treatment sequencing on the cost-effectiveness of rituximab. This led to two potentially important conclusions:

- that if more than one TNFi is contemplated then they should be reserved for 'last resort' treatment when all other options have been exhausted, and

- that it appears to be preferable to use leflunomide before rituximab in a series of treatments, since it is unlikely to lead to meaningful life-time outcome differences, but does generate lower average costs.

The company presented estimates of the likely budget impact of using rituximab which suggest that rituximab will be cost-saving to the NHS. These estimates contradict their own model results, which show net cost increases per patient under all assumptions. Since in the long-run the budget impact must converge to the undiscounted incremental cost generated in the model, it is clear that the NHS must expect to budget for additional costs of £10,000 - £15,000 per patient with severe RA over their remaining lifetime, equivalent to additional annual costs of £12-18 million from use of rituximab in the proposed manner.

5.1 Implications for future research

There are no published RCTs of rituximab versus a relevant comparator (e.g. leflunomide or a second, third TNFi) to inform the management strategies with rituximab that are described in the company's analysis of the decision problem. Future trials are therefore necessary in order to undertake comprehensive comparisons of rituximab with all relevant treatment strategies for patients with severe RA who have failed therapy including a prior TNFi.

There is substantial uncertainty around important clinical issues, most especially in relation to the long-term progression of disease and its effect on HAQ scores, and the duration of effective treatment for each of the active agents considered. Further research in these areas is warranted.

Finally, due to the only recent use of rituximab in this patient population, there is a paucity of long-term evidence for both the continued clinical benefit of rituximab and its long-term comparative safety, including the safety of subsequently treating patients with other DMARDs including TNFi. Close monitoring and surveillance of these patients are therefore necessary.

6 REFERENCES

[1] Cohen SB, Emery P, Greenwald MW, Dougados M, Furie RA, Genovese MC, et al. Rituximab for rheumatoid arthritis refractory to anti-tumor necrosis factor therapy: results of a multicenter, randomized, double-blind, placebo-controlled, phase III trial evaluating primary efficacy and safety at twenty-four weeks. Arthritis & Rheumatism. 2006;54(9):2793-806.

[2] Emery P, Fleischmann R, Filipowicz-Sosnowska A, Schechtman J, Szczepanski L, Kavanaugh A, et al. The efficacy and safety of rituximab in patients with active rheumatoid arthritis despite methotrexate treatment: results of a phase IIB randomized, double-blind, placebo-controlled, dose-ranging trial. Arthritis & Rheumatism. 2006;54(5):1390-400.

[3] Edwards J, Szczepanski L, Szechinski J, Filipowicz-Sosnowska A, Emery P, Close D, et al. Efficacy of B-cell-targeted therapy with rituximab in patients with rheumatoid arthritis. New England Journal of Medicine. 2004;350(25):2572-81.

[4] Chen Y-F, Jobanputra P, Barton P, Jowett S, Bryan S, Clark W, et al. A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness. Health Technology Assessment. 2006;10(42).

[5] Emery P. Treatment of rheumatoid arthritis. British Medical Journal. 2006;332(7534):152-5.

[6] Symmons D, Turner G, Webb R, Asten P, Barrett E, Lunt M, et al. The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. Rheumatology. 2002;41(7):793-800.

[7] Symmons DPM, Silman AJ. The Norfolk Arthritis Register (NOAR). Clinical and Experimental Rheumatology. 2003;21 (Suppl 30):S94-S9.

[8] Luqmani R, Hennell S, Estrach C, Birrell F, Bosworth A, Davenport G, et al. British Society for Rheumatology and British Health Professionals in Rheumatology guideline for the management of rheumatoid arthritis (the first two years). Rheumatology. 2006;45(9):1167-9.

[9] Arthritis Research Campaign. Arthritis - the big picture. Chesterfield: Arthritis Research Campaign; 2002.

[10] Doan T, Massaroti, E. Rheumatoid arthritis: an overview of new and emerging therapies. Journal of Clinical Pharmacology. 2005;45:751-62.

[11] Moreland L. Unmet needs in rheumatoid arthritis. Arthritis Research & Therapy. 2005;7 (Supp 3):S2-S8.

[12] Symmons D. What is the cause of rheumatoid arthritis? Non-genetic factors. National Rheumatoid Arthritis Society 2002 [cited 31 January 2007]; Available from: http://www.rheumatoid.org.uk/article.php?article_id=69

[13] Ledingham J, Deighton C, on behalf of the British Society for Rheumatology Standards Guidelines and Audit Working Group. Update on the British Society for Rheumatology guidelines for prescribing TNF{alpha} blockers in adults with rheumatoid arthritis (update of previous guidelines of April 2001). Rheumatology. 2005;44(2):157-63.

[14] American College of Rheumatology Subcommittee on Rheumatoid Arthritis Guidelines. Guidelines for the management of rheumatoid arthritis: 2002 Update. Arthritis & Rheumatism. 2002;46(2):328-46.

[15] NICE. Final appraisal determination: adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis. 2006 [cited 31 January 2007]; Available from: http://www.nice.org.uk/page.aspx?o=388554

[16] The British Society for Rheumatology. The BSR Biologics Register and what it does. [BSR Biologics Register Newsletter, July 2006] [cited 6 February 2007]; Available from: www.boneandjointdecade.org/ViewDocument.aspx?ContId=1722

[17] National Horizon Scanning Centre. New and Emerging Technology Briefing: Rituximab (MabThera) for rheumatoid arthritis. Birmingham: Department of Public Health and Epidemiology, University of Birmingham; 2003.

[18] Jobanputra P, Barton P, Bryan S, Burls A. The effectiveness of infliximab and etanercept for the treatment of rheumatoid arthritis: a systematic review and economic evaluation. Health Technology Assessment. 2002;6(21).

[19] Penn H. Biologic therapies in autoimmune diseases. Clinical Medicine. 2006;6(105-8).
[20] electronic Medicines Consortium. Arava 10, 20 and 100mg Tablets. 21st March 2006 [cited 6th February 2006]; Available from: http://emc.medicines.org.uk/emc/assets/c/html/displaydoc.asp?documentid=7480#PRODUCTINFO
[21] The University of Sheffield School of Health and Related Research. MSc Health Informatics, Unit 5 [3. Appraising Systematic Reviews]. Critical Appraisal of Secondary Research [cited 31 January 2007]; Available from: http://www.shef.ac.uk/scharr/ir/mschi/unit5/3appraising.htm#casr
[22] Dickersin K, Scherer R, Lefebvre C. Systematic Reviews: Identifying relevant studies for systematic reviews. British Medical Journal. 1994;309(6964):1286-91.
[23] Glanville JM, Lefebvre C, Miles JN, Camosso-Stefinovic J. How to identify randomized controlled trials in MEDLINE: ten years on. Journal of the Medical Library Association. 2006;94(2):130-6.
[24] Chan A-W, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. British Medical Journal. 2005;330(7494):753-6.
[25] Cohen S, Emery P, Greenwald M, Dougados M, Williams S, Cravets M, et al. Prolonged efficacy of rituximab in rheumatoid arthritis patients with inadequate response to one or more TNF inhibitors: 1-year follow-up of a subset of patients receiving a single course in a controlled trial (REFLEX study). Annals of the Rheumatic Diseases. 2006;65(Suppl II):183.
[26] Keystone E, Emery P, Peterfly CG, Tak PP, Cohen S, Genovese MC, et al. Prevention of joint structural damage at 1 year with rituximab in rheumatoid arthritis patients with an inadequate response to one or more TNF inhibitors (REFLEX study). Annals of the Rheumatic Diseases. 2006;65(Suppl II):58.
[27] Keystone E, Fleischmann R, Emery P, Chubick A, Dougados M, Baldassare AR, et al. Long-term efficacy and safety of a repeat treatment course of rituximab in rheumatoid arthritis patients with an inadequate response to one or more TNF inhibitors. Annals of the Rheumatic Diseases. 2006;65(Suppl II):323.
[28] Kielhorn A, Bombardieri S, Aultman R, Jost F, Magrini F. Rituximab improves health-related quality of life as measured by the SF-36: domain score results from the reflex study. Annals of the Rheumatic Diseases. 2006;65(Suppl II):324.
[29] Kremer J, Tony H, Tak PP, Luggen M, Mariette X, Hessey E, et al. Efficacy of rituximab in active RA patients with an inadequate response to one or more TNF inhibitors. Annals of the Rheumatic Diseases. 2006;65(Suppl II):326.
[30] Tak PP, Mease P, Bombardieri S, Lue C, Isaacs JD, Schechtman J, et al. Repeat treatment with rituximab improves physical function and quality of life in patients with rheumatoid arthritis and an inadequate response to TNF inhibitors. Annals of the Rheumatic Diseases. 2006;65(Suppl II):503.
[31] NICE. Single Technology Appraisal (STA) Specification For Manufacturer/Sponsor Submission Of Evidence 2006 [cited 31 January 2007]; Available from: http://www.nice.org.uk/page.aspx?o=356870
[32] Savage I. Rheumatoid arthritis. Chemist & Druggist. 2002;259(6380):19-22.
[33] Parkinson S, Alldred A. Drug regimens for rheumatoid arthritis. Hospital Pharmacist. 2002;9(1):11-5.
[34] EMEA. MabThera-®. European Public Assessment Report (EPAR). 13 October 2006 [cited 31 January 2007]; Available from: http://www.emea.eu.int/humandocs/PDFs/EPAR/Mabthera/Mabthera-H-259-II-39.pdf
[35] Dixon WG, Hyrich KL, Watson K, Symmons DPM, Silman AJ. Serious infection rates in patients receiving biologic therapy in the United Kingdom: Results from the BSR biologics register (BSRBR). Annals of the Rheumatic Diseases. 2005;64(Suppl III):86.
[36] Griffiths I, Silman A, Symmons D, Scott D. BSR Biologics Registry. Rheumatology. 2004;43(1463-64).
[37] Choi HK, Seeger JD, Kuntz KM. A cost effectiveness analysis of treatment options for methotrexate-naive rheumatoid arthritis. Journal of Rheumatology. 2002;29(6):1156-65.

[38] Bansback NJ, Brennan A, Ghatnekar O. Cost effectiveness of adalimumab in the treatment of patients with moderate to severe rheumatoid arthritis in Sweden. Annals of the Rheumatic Diseases. 2005;64(7):995-1002.
[39] Strand V, Cohen S, Schiff M, Weaver A, Fleischmann R, Cannon G, et al. Treatment of active rheumatoid arthritis with leflunomide compared with placebo and methotrexate. Leflunomide Rheumatoid Arthritis Investigators Group. Archives of Internal Medicine. 1999;159(21):2542-50.
[40] Weinblatt ME, Kremer JM, Bankhurst AD, Bulpitt KJ, Fleischmann RM, Fox RI, et al. A trial of etanercept, a recombinant tumor necrosis factor receptor: Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate. New England Journal of Medicine. 1999;340(4):253-9.
[41] Maini R, St Clair EW, Breedveld F, Furst D, Kalden J, Weisman M, et al. Infliximab (chimeric anti-tumour necrosis factor alpha monoclonal antibody) versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: a randomised phase III trial. ATTRACT Study Group. Lancet. 1999;354(9194):1932-9.
[42] Keystone EC, Kavanaugh AF, Sharp JT, Tannenbaum H, Hua Y, Teoh LS, et al. Radiographic, clinical, and functional outcomes of treatment with adalimumab (a human anti-tumor necrosis factor monoclonal antibody) in patients with active rheumatoid arthritis receiving concomitant methotrexate therapy: a randomized, placebo-controlled, 52-week trial. Arthritis & Rheumatism. 2004;50(5):1400-11.
[43] Genovese MC, Becker JC, Schiff M, Luggen M, Sherrer Y, Kremer J, et al. Abatacept for rheumatoid arthritis refractory to tumor necrosis factor alpha inhibition. New England Journal of Medicine. 2005;353(11):1114-23.
[44] Lewis G, Porter D, Brown B, Diamantopoulos A, Kielhorn A. Economic consequences of providing rituximab as a new therapeutic option for rheumatoid arthritis in the UK. Annals of the Rheumatic Diseases. 2006;65(Suppl II):283.
[45] Porter D, Lewis G, Diamantopoulos A, Kielhorn A. Cost-effectiveness analysis of rituximab as a new therapeutic option for rheumatoid arthritis in the UK. Value in Health. 2006;9(3):A25.
[46] Schach S, Rubbert A, Diamantopoulos A, Aristides M, Kielhorn A. Annals of the Rheumatic Diseases. 2006;65(Supplement II):600.
[47] Rubbert A, Kielhorn A, Schach S, Brown B, Aristides M. Cost-effectiveness of rituximab as a new treatment modality for rheumatoid arthritis in Germany. International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Eleventh Annual International Meeting. Philadelphia, PA, USA 2006.
[48] De Vita S, Giuliani G, Diamantopoulos A, Brown B, Kielhorn A. Modelling of the cost-effectiveness of rituximab for treatment of rheumatoid arthritis in Italy. International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Eleventh Annual International Meeting. International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Eleventh Annual International Meeting 2006.
[49] Kielhorn A, Alvarez C, Aristides M, Diamantopoulos A. Using MabThera in patients with rheumatic arthritis in Spain: results of cost-effectiveness data based on micro-simulation. International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Eleventh Annual International Meeting. Philadelphia, PA, USA 2006.
[50] Kielhorn A, Rubbert A, Porter D, De Vita S, Brown B, Aristides M, et al. Cost-effectiveness of rituximab therapy for rheumatoid arthritis: a pan-European analysis. International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Eleventh Annual International Meeting. Philadelphia, PA, USA 2006.
[51] Tanno M, Nakamura I, Ito K, Tanaka H, Ohta H, Kobayashi M, et al. Modeling and cost-effectiveness analysis of etanercept in adults with rheumatoid arthritis in Japan: a preliminary analysis. Modern Rheumatology. 2006;16:77-84.
[52] Barbieri M, Wong J, Drummond M. The cost effectiveness of infliximab for severe treatment-resistant rheumatoid arthritis in the UK. Pharmacoeconomics. 2005;23(6):607-18.

[53] Coyle D, Judd M, Blumenauer B, Cranney A, Maetzel A, Tugwell P, et al. Infliximab and etanercept in patients with rheumatoid arthritis: a systematic review and economic evaluation. Canadian Coordinating Office for Health Technology Assessment 2006;64.
[54] Drummond M, Jefferson T. Guidelines for authors and peer reviewers of economic submissions to the BMJ. . British Medical Journal. 1996;313:275-83.
[55] Evans S, Goossens M, de Vet H, van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. International Journal of Technology Assessment in Health Care. 2005;21:240-5.
[56] Barton P, Jobanputra P, Wilson J, Bryan S, Burls A. The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis. Health Technology Assessment. 2004;8(11).
[57] Crnkic M, Petersson I, Saxne T, P G. Infiximab, etanercept and leflunomide in rheumatoid arthritis. Clinical experience in southern Sweden. BSR Conference Proceedings. 2001:231a.
[58] Scott DL, Garrood T. Quality of life measures: use and abuse. Baillieres Best Practice and Research in Clinical Rheumatology. 2000;14(4):663-87.
[59] Young A, Dixey J, Cox N, Davies P, Devlin J, Emery P, et al. How does functional disability in early rheumatoid arthritis (RA) affect patients and their lives? Results of 5 years of follow-up in 732 patients from the Early RA Study (ERAS). Rheumatology. 2000;39(6):603-11.
[60] NICE. Guide to the methods of technology appraisal. London: National Institute for Health and Clinical Excellence; 2004.
[61] Government Actuary's Department. Interim life tables. [cited 31 January 2007]; Available from: www.gad.gov.uk/Life_Tables/Interim_life_tables.htm
[62] Wiles N, Symmons DP, Harrison B, Barrett E, Barrett JH, Scott DG, et al. Estimating the incidence of rheumatoid arthritis: trying to hit a moving target? Arthritis & Rheumatism. 1999;42(7):1339-46.
[63] Wolfe F, Hawley DJ, Cathey MA. Clinical and health status measures over time: prognosis and outcome assessment in rheumatoid arthritis. Journal of Rheumatology. 1991;18(9):1290-7.
[64] Lassere M, Wells G, Tugwell P, Edmonds J. Percentile curve reference charts of physical function: rheumatoid arthritis population. Journal of Rheumatology. 1995;22(7):1241-6.
[65] Ward MM, Leigh JP, Fries JF. Progression of functional disability in patients with rheumatoid arthritis. Associations with rheumatology subspecialty care. Archives of Internal Medicine. 1993;153(19):2229-37.
[66] Gardiner PV, Sykes HR, Hassey GA, Walker DJ. An evaluation of the Health Assessment Questionnaire in long-term longitudinal follow-up of disability in rheumatoid arthritis. British Journal of Rheumatology. 1993;32(8):724-8.
[67] Callahan LF, Pincus T, Huston JW, 3rd, Brooks RH, Nance EP, Jr., Kaye JJ. Measures of activity and damage in rheumatoid arthritis: depiction of changes and prediction of mortality over five years. Arthritis Care & Research. 1997;10(6):381-94.
[68] Leymarie F, Jolly D, Sanderman R, Briancon S, Marchand AC, Guillemin F, et al. Life events and disability in rheumatoid arthritis: a European cohort. British Journal of Rheumatology. 1997;36(10):1106-12.
[69] Ward MM, Lubeck D, Leigh JP. Longterm health outcomes of patients with rheumatoid arthritis treated in managed care and fee-for-service practice settings. Journal of Rheumatology. 1998;25(4):641-9.
[70] Munro R, Hampson R, McEntegart A, Thomson EA, Madhok R, Capell H. Improved functional outcome in patients with early rheumatoid arthritis treated with intramuscular gold: results of a five year prospective study. Annals of the Rheumatic Diseases. 1998;57(2):88-93.
[71] British National Formulary. BNF 51. 2006 March [cited 2006; Available from: http://www.bnf.org/bnf/
[72] Department of Health. NHS Reference Costs 2005-06. 2006 [cited 2006 December 2006]; Available from: http://www.dh.gov.uk/PublicationsAndStatistics/Publications/PublicationsPolicyAndGuidance/PublicationsPolicyAndGuidanceArticle/fs/en?CONTENT_ID=4141135&chk=7wrsFg

[73] van Vollenhoven R, Harju A, Brannemark S, Klareskog L. Treatment with infliximab (Remicade) when etanercept (Enbrel) has failed or vice versa: data from the STURE registry showing that switching tumour necrosis factor{alpha} blockers can make sense. Annals of the Rheumatic Diseases. 2003;62(1195).
[74] Wolfe F, Mitchell D, Sibley J, Fries J, Bloch D, Williams C, et al. The mortality of rheumatoid arthritis. Arthritis & Rheumatism. 1994;37:481-94.
[75] Edwards CJ, Arden NK, Fisher D, Saperia JC, Reading I, Van Staa TP, et al. The changing use of disease-modifying anti-rheumatic drugs in individuals with rheumatoid arthritis from the United Kingdom General Practice Research Database. Rheumatology. 2005;44(11):1394-8.

7 APPENDICES

Appendix 1: Patient flow diagrams

The flow diagrams provided by the company are shown below.

Group A) Primary analysis

This is the ITT population from study WA17042, used for all the primary and sub-group analyses.

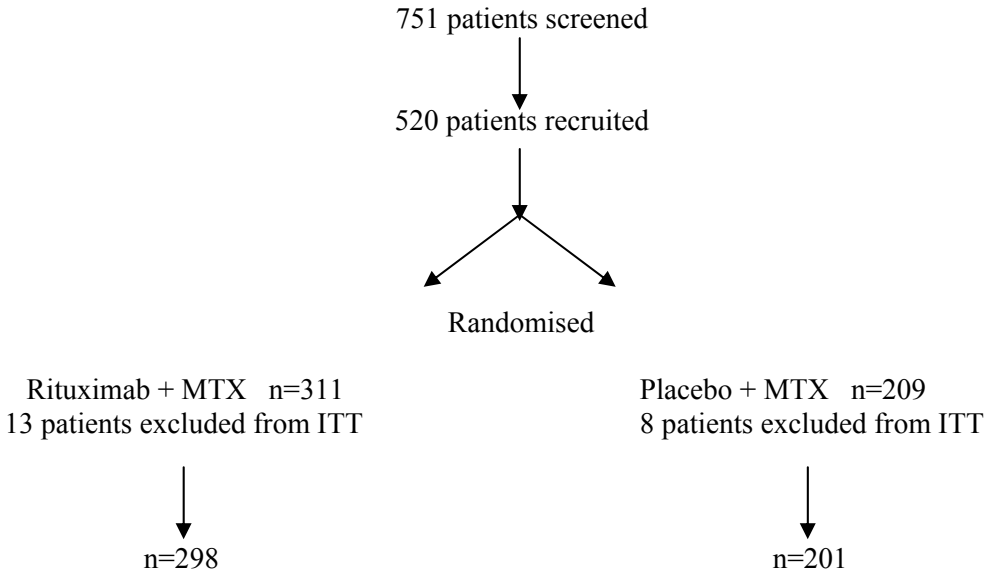
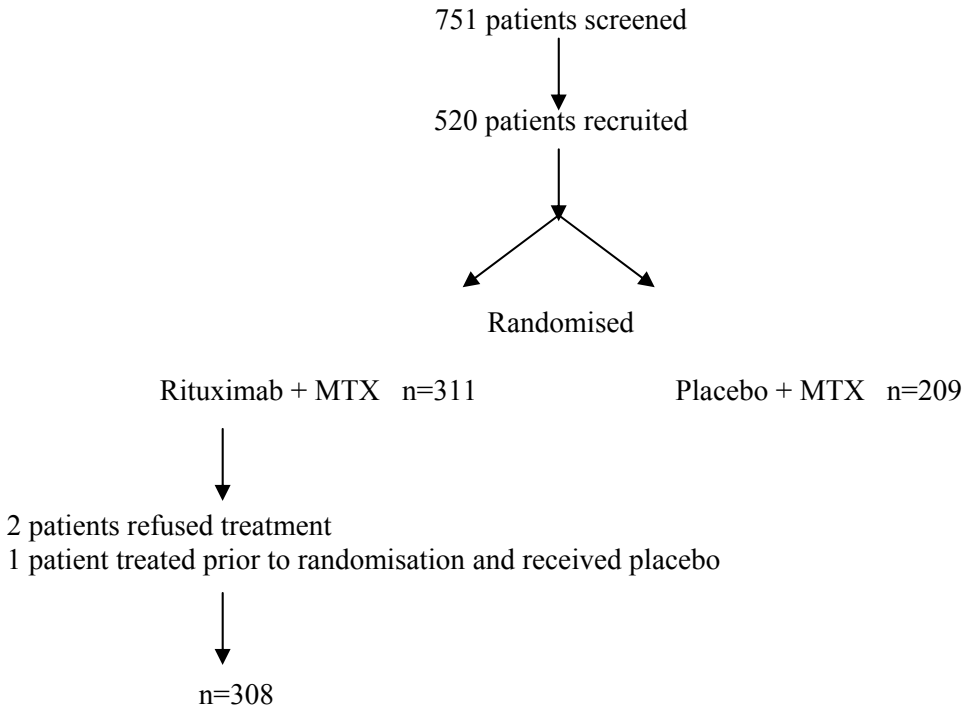


Table 7-1: Reasons patients were excluded from ITT

	Placebo (n)	Rituximab (n)
Blinding compromised	4	3
Received treatment but not randomised	2	4
Patient from site audited for compliance issues	1	4
Received no dose of study treatment	1	2
Total	**8**	**13**

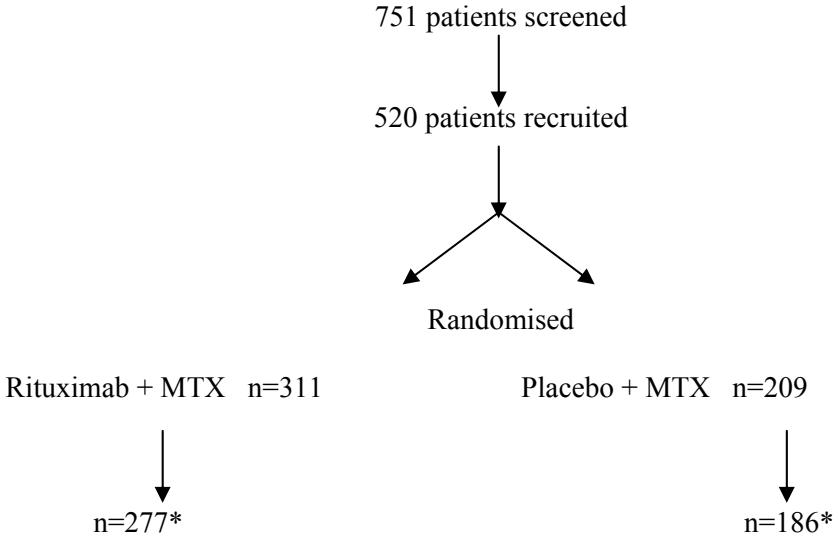
Group B) Long-term efficacy after one course of rituximab

This is based on all treated patients according to the treatment to which they were randomised.



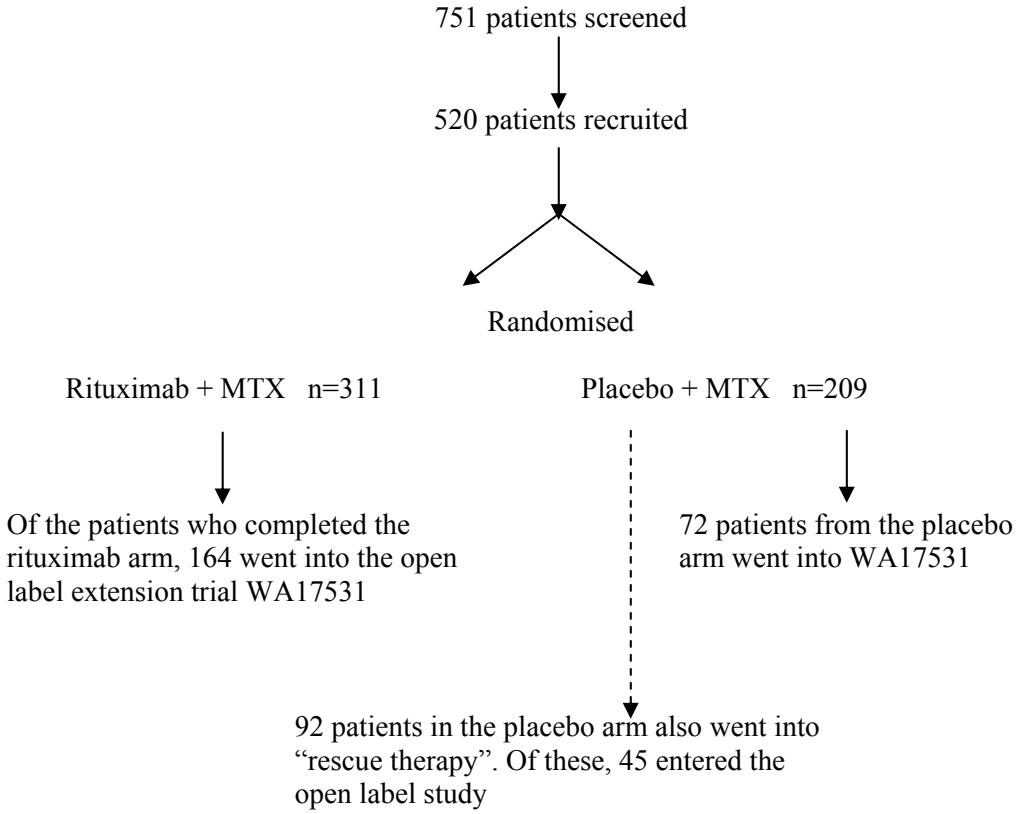
Group C) Radiographic endpoints after 56 weeks

This is based on a "completer population", i.e. all patients in the ITT for whom baseline and week 56 radiographs were taken.



* Numbers of patients in each group who received treatment and screening/post-screening radiographic assessment

Group D) Long-term efficacy following repeated courses



Total number of patients in WA17531 = 281

Group E) Analysis of adverse events

938 patients detailed in the tables of adverse events in the Summary of Product Characteristics arise from 540 rituximab-treated patients and 398 placebo-treated patients in randomised, double-blind Phase II and Phase III studies after one course of rituximab or placebo. This is different from the "All-Exposure" population as patients in open-label extension studies are not included.

Rituximab-treated patients:

308 patients from Phase III study WA17042
192 patients from Phase IIb study WA17043
40 patients from Phase IIa study WA16291

Placebo-treated patients:

209 patients from Phase III study WA17042
149 patients from Phase IIb study WA17043
40 patients from Phase IIa study WA16291

The numbers for the analysis of infusion reactions and infections (N=1039) are from the "all exposure" safety population, as shown in the table below. This also lists the total numbers of patients with 2, 3 and 4 courses of rituximab at the time of the data cut off (14th Oct 2005):

Table 7-2: Patients included in summaries of long-term data following first and repeated course of rituximab*

Source study		First course of rituximab	Second course of rituximab	Third course of rituximab	Fourth course of rituximab
Prior expose to TNFi therapy					
WA17042	Randomised treatment	308	164	50	10
	Open label	164	68	13	3
WA17043	Randomised treatment	53	27	13	4
	Open label	31	17	4	1
WA16291[b]	Randomised treatment	2	1	1	1
	Open label	3	2	1	0
Total patients		**561**	**279**	**82**	**19**
No prior expose to TNFi therapy					
WA17043	Randomised treatment	139	74	23	2
	Open label	76	35	8	0
WA16291[b]	Randomised treatment	36	23	11	1
	Open label	18	13	8	2
Total patients		**269**	**145**	**50**	**5**
Lower dose of rituximab+MTX					
WA17043	Rituximab 2x500mg[a]	124	90	32	5
Other rituximab regimens					
WA16291[b]	Rituximab 2x1000mg only[a]	38	23	12	5
	Rituximab 2x1000mg + ciclosporin[a]	37	23	10	3
All studies	**All regimens, randomised and open label**	**1039**	**570**	**191**	**40**

* All treatments were with rituximab 2x100mg + MTX unless it is indicated otherwise

a First course only then 2x1000mg+MTX for additional treatment courses

b Excluding ten patients who received additional blinded treatment (according to original randomisation) within WA16291

c Includes the above ten patients

Appendix 4

List of Modules and Their Descriptions

Calculation Entry Points (Front End)

- *Microsimulation* - main entry point of the Microsimulation calculation
- *PSA* - main entry point of the PSA calculation
- *One_way_SA* - main entry point of the One Way SA calculation
- *CEAC* - main entry point of the CEAC calculation
- *GoNext* - entry point for trial-by-trial Microsimulation calculation, one step forward

Optimized Calculations and General Utilities

- *mdlOptimization* - optimized implementations of all calculations
- *mdlRandom* - optimized implementation of random generator
- *mdlCopy* - utilities for handling worksheets
- *CEAC_Input* (form) - spreadsheet dialog form requesting parameters for the CEAC calculation

Connection with Interactive Tool (Director)

- *mdlInteractive* - basic communication (messages and flags) with the interactive tool (Director)
- *mdlFuncs* - utilities for efficient data transfer between the spreadsheet and the interactive tool (Director)
- *mdlControl* - process control functions used for the implementation of CANCEL button (see also CControl class module)
- *CControl* (class module) - class object used for the implementation of CANCEL button
- *ThisWorkbook* (workbook module) - workbook event handlers, used for prevention of back-end process termination, and for the implementation of CANCEL button

Unused and Obsolete

- *mdlTimer* - support for time profiling (not used in the production version)
- *GoPrevious* - obsolete implementation for trial-by-trial Microsimulation calculation, one step back (not supported in the optimized version)
- *Module1* - empty module

List of Variables of *mdlOptimization* Module

Global Variables

Name	***Description***	***Worksheet Reference***
Cycle	Cycle (increments of time)	Variables!C6
Base_Age	Population data: Base_Age	Variables!Z28
LinePal	Palliative care is line: 1st Sequence	Variables!C21
LinePal2	Palliative care is line: 2nd Sequence	Variables!D21
dACR20	HAQ score change wrt response: ACR20-49	Variables!Z74
dACR50	HAQ score change wrt response: ACR50-69	Variables!Z75
dACR70	HAQ score change wrt response: ACR70+	Variables!Z76
dACRNR	HAQ score change wrt response: Non responder	Variables!Z73
Rebound_effect	Rebound effect of new treatment	Variables!C35
Base_HAQ	Population data: Base_HAQ	Variables!Z29
RiskM	HAQ risk multiplier	Variables!C116
DisUtil	Discounting: Effects	Variables!C119
DisCos	Discounting: Costs	Variables!C120
retirement_age	Retirement age	Variables!C42
ciHAQ1-ciHAQ6	Resource use: Total indirect cost	Variables!Z97:AE97
cHAQ1-cHAQ6	Resource use: Total resource use costs	Variables!Z98:AE98
Negative_QALY_status	QALYs worse than death	Variables!C37
Allowed	"Allowed"	Variables!D37
female	Population weighting: Female	Variables!C33
male	Population weighting: Male	Variables!C32
nr_new_treatment_disc	Drug cost for non-responders: reduces anti-TNF costs by	Variables!C40
QoL_Equation	User selection	Qol_Equations!D2
ChangeRTX_ChangeHTX	Time on treatment: Years	Variables!AA58:AA66
RTXpACR20_HTXpACR20	Response rates: ACR20-49	Variables!Z46:Z54
RTXpACR50_HTXpACR50	Response rates: ACR50-69	Variables!AA46:AA54
RTXpACR70_HTXpACR70	Response rates: ACR70+	Variables!AB46:AB54
dHAQRTX_dHAQHTX	HAQ score long term deterioration on treatment	Variables!Z79:Z87
cRTX_cHTX	Drug costs (including administration and monitoring)	Variables!Z102:Z110
lifeTable	Life table (cached)	Life tables!A4:D104
treatmentIndices	Order of treatments in the 1st and 2nd Sequence, used to retain the treatment response probabilities between the sequences	Built from Variables!B9:D17

Local Variables of RunModel *Subroutine*

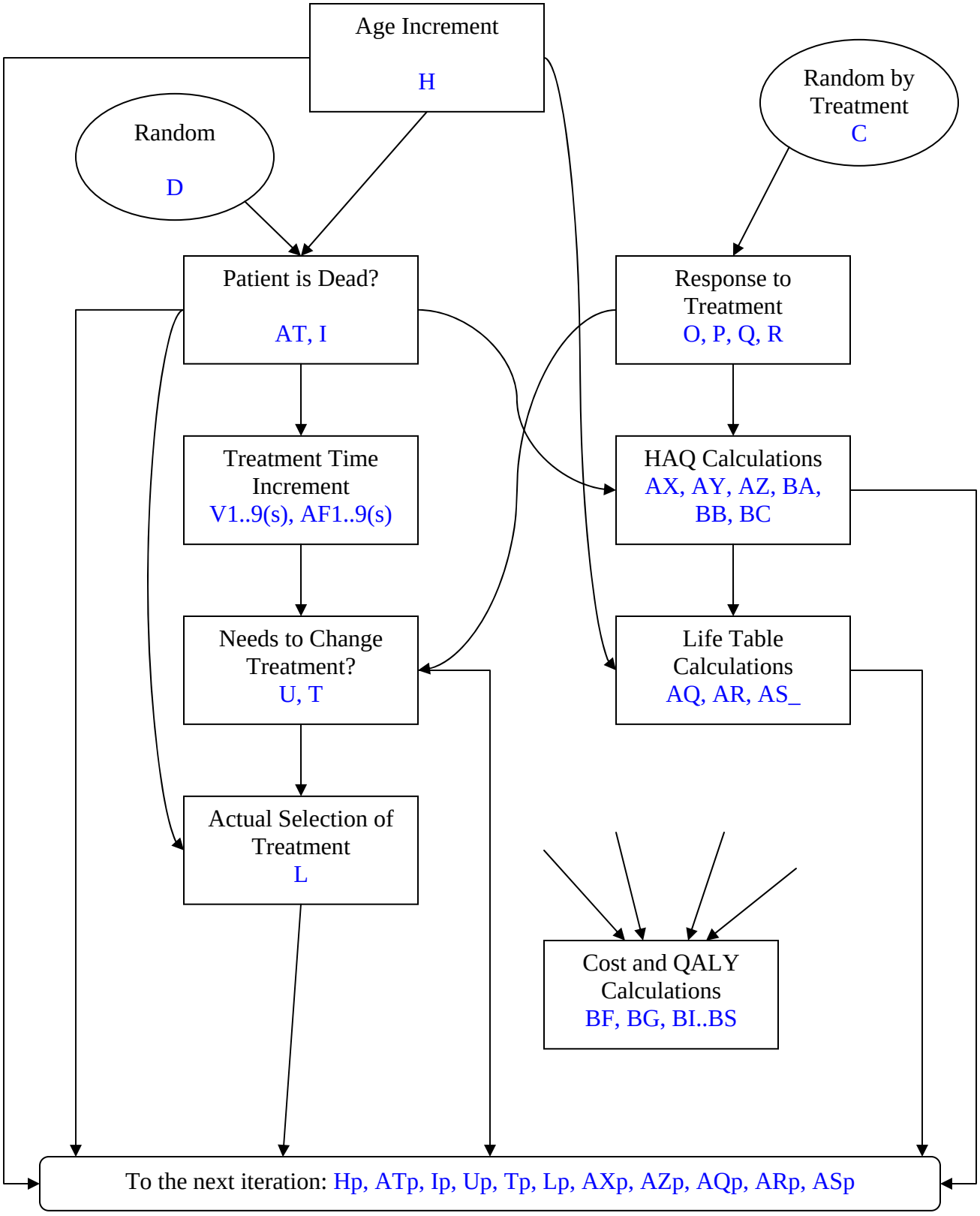
Name	***Description***
B	Stage Result
C	Random numbers pResponse
D	R pNatDeath
H, Hp	Age
I, Ip	Death
L, Lp	Treatment Line
O	Responses: ACR20
P	Responses: ACR50
Q	Responses: ACR70
R	Responses: Non responder
T, Tp	Change treatment: Tracker
U, Up	Change treatment: when complete time on treatment
V1-V9, V1s-V9s	Time on n-th treatment
AF1-AF9, AF1s-AF9s	Discounted time on n-th treatment
AQ, AQp	Adjust Age for Life Table
AR, ARp	Result of Life table
AS_, ASp	Adjust life table for model cohort (mortality)
AT, ATp	pNatDeath
AX, AXp	HAQ score
AY	HAQ score improvement from treatment effect
AZ, AZp	HAQ score rebound effect: track
BA	HAQ score rebound effect: unadjusted rebound
BB	HAQ rebound: adjusted rebound
BC	HAQ score deterioration while on treatment
BF	QALYs: Accumulated
BG, BGs	QALYs: Discounted
BI	Costs: Total Resource use
BJ	Costs: Treatment Cost
BK, BKs	Costs: Treatment Cost Discounted
BL	Costs: Total undiscounted
BM, BMs	Costs: Total Discounted
BN	Costs: Indirect Cost undiscounted
BO, BOs	Costs: IC Discounted

Note:

- suffix “p” denotes the value of the variable retained from the previous iteration
- suffix “s” denotes accumulated sum of the values of the variable throughout all the iterations up to the current iteration.

Calculation Flow Diagram - *mdlOptimization.runModel*

Note: the diagram shows the calculation flow of one iteration. Every box lists the names of variables being updated on the corresponding stage.



ar_slex01_rnd Listing - Withdrawals (All Populations, as Randomised)
Protocol(s): WA17042
Analysis: ALL PATIENTS Center: ALL CENTERS
Data Extracted on 07APR2005
Treatment: Placebo+MTX; N = 209
Week 24 Analysis

CRTN/Pt. No.	Age Yr	Sex	Weight kg	Race	Day of Last Dose	Reason for Withdrawal
36679/5826	66	F	115	BLACK	15	INSUFF. THERAPY
36679/7730	29	F	48	CAUCASIAN	16	INSUFF. THERAPY
36681/4523	32	M	120	CAUCASIAN	15	INSUFF. THERAPY
36682/4545	47	M	90	CAUCASIAN	15	INSUFF. THERAPY
36683/1275	45	F	83	CAUCASIAN	15	INSUFF. THERAPY
36684/1195	52	F	85	HISPANIC	15	INSUFF. THERAPY
36685/5175	54	F	123	CAUCASIAN	23	INSUFF. THERAPY
36685/5180	58	F	63	CAUCASIAN	15	INSUFF. THERAPY
36685/5181	41	F	58	CAUCASIAN	15	INSUFF. THERAPY
36686/5191	45	F	63	CAUCASIAN	16	INSUFF. THERAPY
36686/5196	44	F	98	CAUCASIAN	15	INSUFF. THERAPY
36687/1651	51	F	75	CAUCASIAN	15	INSUFF. THERAPY
36687/1660	51	F	89	CAUCASIAN	15	INSUFF. THERAPY
36688/5213	67	F	64	CAUCASIAN	16	AE/INT. ILLNESS
36689/4491	38	F	92	CAUCASIAN	15	INSUFF. THERAPY
36690/1714	60	F	93	CAUCASIAN	1	ADMIN/OTHER UNBLINDED DUE TO CLINICAL HOLD
36690/1716	29	F	80	CAUCASIAN	15	INSUFF. THERAPY
36691/1792	41	F	86	CAUCASIAN	15	INSUFF. THERAPY

CRTN = Clinical Research Task Number (center no.)
Including patients that withdrew into the rescue period
EX01 13APR2005:10:40:50 (1 of 10)

ar_slex01_rnd Listing - Withdrawals (All Populations, as Randomised)
Protocol(s): WA17042
Analysis: ALL PATIENTS Center: ALL CENTERS
Data Extracted on 07APR2005
Treatment: Placebo+MTX; N = 209
Week 24 Analysis

CRTN/Pt. No.	Age Yr	Sex	Weight kg	Race	Day of Last Dose	Reason for Withdrawal
36691/1795	64	F	84	CAUCASIAN	15	INSUFF. THERAPY
36691/1799	35	F	67	CAUCASIAN	15	INSUFF. THERAPY
36693/1283	49	M	83	CAUCASIAN	15	INSUFF. THERAPY
36697/1406	60	F	117	CAUCASIAN	14	INSUFF. THERAPY
36698/1411	62	F	69	CAUCASIAN	15	INSUFF. THERAPY
36699/4592	55	F	81	CAUCASIAN	14	INSUFF. THERAPY
36700/4614	49	F	82	CAUCASIAN	15	INSUFF. THERAPY
36700/4616	40	F	131	CAUCASIAN	15	INSUFF. THERAPY
36701/4648	41	F	97	CAUCASIAN	15	INSUFF. THERAPY
36701/4649	46	F	70	CAUCASIAN	23	INSUFF. THERAPY
36701/4654	20	F	86	AMERICAN INDIAN	15	INSUFF. THERAPY
36704/1112	50	F	71	CAUCASIAN	15	INSUFF. THERAPY
36717/1740	31	F	93	CAUCASIAN	15	INSUFF. THERAPY
36719/1612	58	F	80	CAUCASIAN	15	INSUFF. THERAPY
36719/1614	44	F	90	CAUCASIAN		AE/INT. ILLNESS
36719/1619	46	F	118	CAUCASIAN	1	INSUFF. THERAPY
36719/1620	51	F	109	CAUCASIAN	1	INSUFF. THERAPY

CRTN = Clinical Research Task Number (center no.)
Including patients that withdrew into the rescue period
EX01 13APR2005:10:40:50 (2 of 10)

ar_slex01_rnd Listing - Withdrawals (All Populations, as Randomised)
Protocol(s): WA17042
Analysis: ALL PATIENTS Center: ALL CENTERS
Data Extracted on 07APR2005
Treatment: Placebo+MTX; N = 209
Week 24 Analysis

CRTN/Pt. No.	Age Yr	Sex	Weight kg	Race	Day of Last Dose	Reason for Withdrawal
36721/1582	66	M	77	CAUCASIAN	1	ADMIN/OTHER PT WAS WITHDRAWN DUE TO CLINICAL HOLD
36724/5673	57	F	50	CAUCASIAN	14	INSUFF. THERAPY
36724/5675	67	F	46	CAUCASIAN	15	INSUFF. THERAPY
36724/5679	20	F	48	CAUCASIAN	15	INSUFF. THERAPY
36724/8076	51	F	87	CAUCASIAN	15	INSUFF. THERAPY
36732/1075	28	F	53	CAUCASIAN	1	REFUSED TREAT.
36732/1082	58	F	92	BLACK	15	INSUFF. THERAPY
36732/1087	45	F	74	CAUCASIAN	15	INSUFF. THERAPY
36734/1314	68	M	96	CAUCASIAN	15	INSUFF. THERAPY
36734/1319	36	F	50	BLACK	15	INSUFF. THERAPY
36735/5334	34	F	49	CAUCASIAN	15	REFUSED TREAT.
36738/5401	61	M	97	CAUCASIAN	16	ADMIN/OTHER RESCUE THERAPY
36741/4701	61	F	54	CAUCASIAN	16	INSUFF. THERAPY
36743/5255	62	F	67	CAUCASIAN	15	INSUFF. THERAPY
36743/5256	63	F	105	CAUCASIAN	16	INSUFF. THERAPY
36743/5260	49	F	96	CAUCASIAN	34	INSUFF. THERAPY
36745/5066	32	F	65	CAUCASIAN	15	INSUFF. THERAPY

CRTN = Clinical Research Task Number (center no.)
Including patients that withdrew into the rescue period
EX01 13APR2005:10:40:50 (3 of 10)

ar_slex01_rnd Listing - Withdrawals (All Populations, as Randomised)
Protocol(s): WA17042
Analysis: ALL PATIENTS Center: ALL CENTERS
Data Extracted on 07APR2005
Treatment: Placebo+MTX; N = 209
Week 24 Analysis

CRTN/Pt. No.	Age Yr	Sex	Weight kg	Race	Day of Last Dose	Reason for Withdrawal
36745/5068	28	F	73	CAUCASIAN	15	INSUFF. THERAPY
36745/5069	34	F	49	CAUCASIAN	15	INSUFF. THERAPY
36751/5582	54	F	46	CAUCASIAN	14	FAIL. TO RETURN
36752/5611	59	F	71	CAUCASIAN	15	INSUFF. THERAPY
36753/5132	54	M	80	CAUCASIAN	15	INSUFF. THERAPY
36756/1350	53	F	104	CAUCASIAN	15	FAIL. TO RETURN
36756/1362	49	F	67	HISPANIC	15	FAIL. TO RETURN
36757/4776	42	F	62	CAUCASIAN	15	INSUFF. THERAPY
36760/5641	33	M	91	CAUCASIAN	15	INSUFF. THERAPY
36760/5642	59	F	74	CAUCASIAN	15	INSUFF. THERAPY
36763/6401	60	F	105	CAUCASIAN	15	INSUFF. THERAPY
36764/6503	68	F	86	CAUCASIAN	15	INSUFF. THERAPY
36764/6504	60	M	79	CAUCASIAN	15	INSUFF. THERAPY
36765/6489	56	F	99	CAUCASIAN	17	INSUFF. THERAPY
36767/6544	67	F	74	CAUCASIAN	15	INSUFF. THERAPY
36767/6546	53	F	67	CAUCASIAN	15	INSUFF. THERAPY
36771/6562	66	F	94	CAUCASIAN	15	INSUFF. THERAPY
36772/6442	70	F	66	CAUCASIAN	15	INSUFF. THERAPY
36773/6767	58	F	88	CAUCASIAN	15	REFUSED TREAT.

CRTN = Clinical Research Task Number (center no.)
Including patients that withdrew into the rescue period
EX01 13APR2005:10:40:50 (4 of 10)

ar_slex01_rnd Listing - Withdrawals (All Populations, as Randomised)
Protocol(s): WA17042
Analysis: ALL PATIENTS Center: ALL CENTERS
Data Extracted on 07APR2005
Treatment: Placebo+MTX; N = 209
Week 24 Analysis

CRTN/Pt. No.	Age Yr	Sex	Weight kg	Race	Day of Last Dose	Reason for Withdrawal
36775/4852	46	F	43	CAUCASIAN	15	INSUFF. THERAPY
36776/4881	50	F	62	CAUCASIAN	15	INSUFF. THERAPY
36777/4925	54	F	58	CAUCASIAN	15	INSUFF. THERAPY
36777/4927	54	F	77	CAUCASIAN	15	INSUFF. THERAPY
36780/6622	44	F	69	CAUCASIAN	15	INSUFF. THERAPY
36781/6647	43	F	67	SOUTH AMERICAN NATIVE INDIAN	16	INSUFF. THERAPY
36782/6682	72	F	68	CAUCASIAN	1	ADMIN/OTHER PT UNBLINDED & WENT INTO RESUE THERAPY DUE TO CLINICAL HOLD
36782/6695	53	F	86	CAUCASIAN	15	INSUFF. THERAPY
36783/6732	56	F	78	CAUCASIAN	15	INSUFF. THERAPY
36788/4965	55	F	61	CAUCASIAN	15	INSUFF. THERAPY
36789/5001	50	M	103	CAUCASIAN	15	INSUFF. THERAPY
36797/6953	47	F	78	CAUCASIAN	15	ADMIN/OTHER RESCUE THERAPY
36797/6956	55	M	77	CAUCASIAN	15	INSUFF. THERAPY
36802/1895	65	F	89	CAUCASIAN	15	REFUSED TREAT.
36803/4493	55	M	66	CAUCASIAN	15	INSUFF. THERAPY
36804/7452	34	F	68	CAUCASIAN	15	INSUFF. THERAPY

CRTN = Clinical Research Task Number (center no.)
Including patients that withdrew into the rescue period
EX01 13APR2005:10:40:50 (5 of 10)

ar_slex01_rnd Listing - Withdrawals (All Populations, as Randomised)
Protocol(s): WA17042
Analysis: ALL PATIENTS Center: ALL CENTERS
Data Extracted on 07APR2005
Treatment: Placebo+MTX; N = 209
Week 24 Analysis

CRTN/Pt. No.	Age Yr	Sex	Weight kg	Race	Day of Last Dose	Reason for Withdrawal
36808/7491	62	F	71	CAUCASIAN	15	INSUFF. THERAPY
36809/7463	74	M	81	CAUCASIAN	17	INSUFF. THERAPY
36810/7431	56	F	88	CAUCASIAN	15	INSUFF. THERAPY
36816/7560	54	F	101	HISPANIC	15	INSUFF. THERAPY
42902/7796	59	F	64	CAUCASIAN	15	INSUFF. THERAPY
42907/7666	57	F	72	CAUCASIAN	15	INSUFF. THERAPY
42907/7670	55	F	78	HISPANIC	15	INSUFF. THERAPY
42907/7671	60	F	101	CAUCASIAN	15	INSUFF. THERAPY
42908/7679	72	F	72	HISPANIC	15	INSUFF. THERAPY
42915/7736	73	F	90	CAUCASIAN	15	REFUSED TREAT.

CRTN = Clinical Research Task Number (center no.)
Including patients that withdrew into the rescue period
EX01 13APR2005:10:40:50 (6 of 10)

ar_slex01_rnd Listing - Withdrawals (All Populations, as Randomised)
Protocol(s): WA17042
Analysis: ALL PATIENTS Center: ALL CENTERS
Data Extracted on 07APR2005
Treatment: Rituximab+MTX; N = 311
Week 24 Analysis

CRTN/Pt. No.	Age Yr	Sex	Weight kg	Race	Day of Last Dose	Reason for Withdrawal
36679/5823	41	F	86	BLACK	15	INSUFF. THERAPY
36679/5824	63	F	66	BLACK		REFUSED TREAT.
36679/5825	61	F	69	CAUCASIAN	15	INSUFF. THERAPY
36681/4524	61	F	93	CAUCASIAN	15	INSUFF. THERAPY
36681/4525	34	F	62	CAUCASIAN	16	AE/INT. ILLNESS
36686/5198	65	F	54	CAUCASIAN	15	INSUFF. THERAPY
36687/1658	54	F	106	CAUCASIAN	15	INSUFF. THERAPY
36689/4579	69	F	40	CAUCASIAN	15	INSUFF. THERAPY
36689/4588	50	F	97	CAUCASIAN	15	AE/INT. ILLNESS
36691/1793	39	F	66	CAUCASIAN	15	INSUFF. THERAPY
36696/1563	69	M	114	CAUCASIAN	15	INSUFF. THERAPY
36697/1404	76	F	87	CAUCASIAN	15	INSUFF. THERAPY
36699/4593	26	M	74	CAUCASIAN	14	INSUFF. THERAPY
36701/4650	48	F	83	CAUCASIAN	15	INSUFF. THERAPY
36707/1964	61	F	92	CAUCASIAN	21	INSUFF. THERAPY
36709/6431	55	F	100	BLACK		ADMIN/OTHER SUBJECT REFUSED TREATMENT
36724/5672	67	M	82	CAUCASIAN	15	INSUFF. THERAPY
36728/5733	46	F	84	CAUCASIAN	15	INSUFF. THERAPY

CRTN = Clinical Research Task Number (center no.)
Including patients that withdrew into the rescue period
EX01 13APR2005:10:40:50 (7 of 10)

ar_slex01_rnd Listing - Withdrawals (All Populations, as Randomised)
Protocol(s): WA17042
Analysis: ALL PATIENTS Center: ALL CENTERS
Data Extracted on 07APR2005
Treatment: Rituximab+MTX; N = 311
Week 24 Analysis

CRTN/Pt. No.	Age Yr	Sex	Weight kg	Race	Day of Last Dose	Reason for Withdrawal
36732/1080	44	F	128	HISPANIC	15	INSUFF. THERAPY
36732/1081	20	F	50	HISPANIC	15	AE/INT. ILLNESS
36732/1083	55	F	83	BLACK	1	FAIL. TO RETURN
36733/1105	43	F	84	CAUCASIAN	15	INSUFF. THERAPY
36734/1312	49	F	87	BLACK	15	INSUFF. THERAPY
36734/1325	63	F	50	CAUCASIAN	15	REFUSED TREAT.
36735/5332	59	F	65	CAUCASIAN	1	AE/INT. ILLNESS
36736/5354	32	F	48	CAUCASIAN	15	INSUFF. THERAPY
36740/4671	59	M	135	CAUCASIAN	15	INSUFF. THERAPY
36741/4704	54	F	54	EAST INDIAN	15	INSUFF. THERAPY
36745/5061	26	F	80	CAUCASIAN	15	REFUSED TREAT.
36746/5082	25	F	41	CAUCASIAN	15	INSUFF. THERAPY
36752/5612	54	F	69	CAUCASIAN	15	INSUFF. THERAPY
36755/1140	28	F	109	CAUCASIAN	15	INSUFF. THERAPY
36758/1950	71	F	73	HISPANIC	15	REFUSED TREAT.
36762/6601	55	F	49	CAUCASIAN	15	INSUFF. THERAPY
36773/6762	50	F	119	CAUCASIAN	15	INSUFF. THERAPY
36773/6768	57	F	82	CAUCASIAN	1	AE/INT. ILLNESS

CRTN = Clinical Research Task Number (center no.)
Including patients that withdrew into the rescue period
EX01 13APR2005:10:40:50 (8 of 10)

ar_slex01_rnd Listing - Withdrawals (All Populations, as Randomised)
Protocol(s): WA17042
Analysis: ALL PATIENTS Center: ALL CENTERS
Data Extracted on 07APR2005
Treatment: Rituximab+MTX; N = 311
Week 24 Analysis

CRTN/Pt. No.	Age Yr	Sex	Weight kg	Race	Day of Last Dose	Reason for Withdrawal
36773/6771	48	F	71	CAUCASIAN	1	INSUFF. THERAPY
36773/7962	61	F	60	CAUCASIAN	1	AE/INT. ILLNESS
36775/4854	43	F	47	CAUCASIAN	1	AE/INT. ILLNESS
36779/4941	52	F	51	CAUCASIAN	15	INSUFF. THERAPY
36782/6681	54	M	80	CAUCASIAN	15	INSUFF. THERAPY
36783/6721	34	F	90	CAUCASIAN	15	INSUFF. THERAPY
36785/6848	31	F	96	HISPANIC	14	REFUSED TREAT.
36789/5002	49	F	47	CAUCASIAN	15	FAIL. TO RETURN
36789/5004	67	F	71	CAUCASIAN	15	INSUFF. THERAPY
36792/6822	53	F	95	BLACK	15	INSUFF. THERAPY
36795/5427	30	M	82	CAUCASIAN	1	AE/INT. ILLNESS
36799/6944	44	M	110	CAUCASIAN	15	FAIL. TO RETURN
36804/7451	56	F	63	HISPANIC	15	INSUFF. THERAPY
36809/7467	35	F	72	CAUCASIAN	15	INSUFF. THERAPY
36810/7435	52	F	99	CAUCASIAN	15	ADMIN/OTHER EVALUATION BY PI CONFIRMED PATIENT DID NOT HAVE RA (TRUE DIAGNOSIS WAS OSTEOARTHRITIS) HENCE RA MEDICATIONS WERE DISCONTINUED
36816/7559	49	M	99	CAUCASIAN	15	FAIL. TO RETURN

CRTN = Clinical Research Task Number (center no.)
Including patients that withdrew into the rescue period
EX01 13APR2005:10:40:50 (9 of 10)

ar_slex01_rnd Listing - Withdrawals (All Populations, as Randomised)
Protocol(s): WA17042
Analysis: ALL PATIENTS Center: ALL CENTERS
Data Extracted on 07APR2005
Treatment: Rituximab+MTX; N = 311
Week 24 Analysis

CRTN/Pt. No.	Age Yr	Sex	Weight kg	Race	Day of Last Dose	Reason for Withdrawal
42902/7795	44	F	68	CAUCASIAN	18	INSUFF. THERAPY
42902/7799	69	F	50	CAUCASIAN	16	INSUFF. THERAPY
42907/7661	53	F	83	CAUCASIAN	15	REFUSED TREAT.
42907/7668	65	F	101	CAUCASIAN	15	INSUFF. THERAPY
42911/7729	34	F	118	CAUCASIAN	15	INSUFF. THERAPY

CRTN = Clinical Research Task Number (center no.)
Including patients that withdrew into the rescue period
EX01 13APR2005:10:40:50 (10 of 10)