Abatacept for the treatment of rheumatoid arthritis

ERG Report

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*
Sherrington Buildings
University of Liverpool
Ashton Street
Liverpool, UK
L69 3GE

Tel: +44 (0) 151 794 5682/5541/5067
Fax: +44 (0) 151 794 5585
Email: LRiG@liv.ac.uk

Members of the review team:

Adrian Bagust, Professor
Health Economics
The University of Liverpool Management School

Angela Boland, Ms
Research Fellow, *Health Economics*
Liverpool Reviews and Implementation Group

Yenal Dundar, Dr
Research Fellow, *Clinical Effectiveness*
Liverpool Reviews and Implementation Group

Rumona Dickson, Ms
Director
Liverpool Reviews and Implementation Group

Nigel Fleeman, Mr
Research Fellow, *Clinical Effectiveness*
Liverpool Reviews and Implementation Group

Juliet Hockenhull, Mrs
Research Fellow, *Clinical Effectiveness*
Liverpool Reviews and Implementation Group

Tom Kennedy, Dr
Consultant Rheumatologist
Royal Liverpool and Broadgreen University Hospital

Christine Proudlove, Ms
Director
North West Medicines Information Centre
Pharmacy Practice Unit, Liverpool

Sarah Donegan
Medical Statistician
Centre for Medical Statistics and Health Evaluation (CMSHE)
University of Liverpool

Correspondence to:

Ms. Rumona Dickson
Director, LRiG
Liverpool Reviews and Implementation Group
Sherrington Buildings
Ashton Street
University of Liverpool
Liverpool, UK
L69 3GE

Tel: [redacted]
Fax: [redacted]
Email: [redacted]

Date completed: May 31st 2007

Acknowledgements:

The authors are pleased to acknowledge Dr Andrew Walker (Robertson Centre for Biostatistics, University of Glasgow) who provided economic advice and commented on drafts of the report, Dr Andrew Hassell (Staffordshire Rheumatology Centre), Dr David Walker (The Freeman Hospital, Newcastle Upon Tyne) and Professor Robert Moots (Royal Liverpool and Broadgreen University Hospital) who provided clinical advice and commented on drafts of the report.

Contents:

1 SUMMARY 10

1.1 Scope of the submission 10

1.2 Summary of submitted clinical effectiveness evidence 10

1.3 Summary of submitted cost-effectiveness evidence 11

1.4 Commentary on the robustness of submitted evidence 12

1.5 Key issues 13

2 BACKGROUND 14

2.1 Critique of manufacturer's description of underlying health problem 14

2.2 Critique of manufacturer's overview of current service provision 16

3 CRITIQUE OF MANUFACTURER'S STATEMENT OF THE DECISION PROBLEM 20

3.1 Population 22

3.2 Intervention 22

3.3 Comparators 22

3.4 Outcomes 23

3.5 Time frame 24

3.6 Other relevant factors 24

4 CLINICAL EFFECTIVENESS 26

4.1 Critique of manufacturer's approach 26

4.2 Summary of submitted evidence 32

4.3 Summary of clinical evidence 46

5 ECONOMIC EVALUATION 48

5.1 Summary of published cost-effectiveness analyses identified in the manufacturer's submission 48

5.2 Overview of manufacturer's economic evaluation 50

5.3 Summary of cost-effectiveness and comment on validity of results presented with reference to methodology used 56

5.4 Critique of approach used 57

5.5 Additional comparison: abatacept + MTX versus cycled TNFi 78

5.6 ERG revised base-case economic results: abatacept + MTX versus MTX 80

5.7 ERG revised additional comparison results: abatacept + MTX versus cycled TNFi 84

5.8 Summary of cost-effectiveness evidence 87

6 DISCUSSION 88

6.1 Implications for future research 91

7 REFERENCES 92

Appendix 1 97

Tables:

Table 3-1: Statement of the decision problem 21

Table 3-2: Outcomes 24

Table 4-1: Quality assessment of the clinical effectiveness review 26

Table 4-2: Scope of the literature review 27

Table 4-3: Characteristics of the one included trial (ATTAIN) 28

Table 4-4: List of relevant non-RCT evidence 29

Table 4-5: Validity assessment of included RCT-ATTAIN 30

Table 4-6: Baseline patient characteristics from ATTAIN[1] 33

Table 4-7: Reported outcomes of ATTAIN trial (at six-months) 34

Table 4-8: Reported outcomes of TNFi use: sub-group analyses 34

Table 4-9: DAS28 results as presented in the CSR[a] 35

Table 4-10: Clinical outcomes of ATTAIN long-term extension analysis[a] 36

Table 4-11: Adverse events reported in ATTAIN 37

Table 4-12: Adverse events in double-blind controlled trial periods 38

Table 4-13: Malignant neoplasms in abatacept-treated patients 39

Table 4-14: Frequency of adverse events from day 1 through 169 in trial IM101064 (ARRIVE) 40

Table 4-15: Subjects with most frequently[a] reported drug related adverse events (ATTAIN) 41

Table 4-16: Efficacy results of trial IM101043 reproduced from the MS pg. 172 43

Table 4-17: Mean change in HAQ – BSRBR, NICE study 44

Table 4-18: Reflex trial versus ATTAIN mixed treatment comparison results 45

Table 5-1: Model parameters 51

Table 5-2: Characteristics of average RA patients used in the model 52

Table 5-3: Cost-effectiveness results for abatacept + MTX versus MTX 54

Table 5-4: Cost-effectiveness results for abatacept + MTX versus cycled TNFi 54

Table 5-5: ERG amendments to treatment costs 62

Table 5-6: ERG amendments to model disease-related hospital costs per patient year 63

Table 5-7: Percentage change in HAQ score from baseline to day 169 - submitted model and ERG analysis of ATTAIN IPD (difference between slopes for abatacept + MTX and placebo + MTX regression lines) 64

Table 5-8: Frequency distribution of annual changes in HAQ scores for 1246 early-stage patients followed for 5 years (Wiles et al)[51] 71

Table 5-9: Frequency distribution of changes in HAQ scores for ATTAIN patients 71

Table 5-10: Evidence of long-term HAQ progression in RA - studies used by Scott and Garrood[55] in estimating average annual progression rate 75

Table 5-11: ERG revised treatment costs 80

Table 5-12: Marginal effect of each of the identified corrections/amendments on cost-effectiveness results using the EXCEL-based model 81

Table 5-13: Cumulative effect of the identified corrections/amendments on cost-effectiveness results using the EXCEL-based model 81

Table 5-14: Three-way sensitivity analysis of results from the ERG-modified Excel-based model (including all identified changes except the ERG preferred progression rates) - ICER (£/QALY gained) 83

Table 5-15: Marginal effects of identified corrections/amendments and revised on alternative scenario cost-effectiveness results - using the Excel-based model 85

Table 5-16: Gender-specific cost-effectiveness results for abatacept + MTX versus three TNFi using the ERG revised Excel-based model of the alternative scenario (cost per QALY gained) 86

Table A-1: Characteristics of the five RCTS outside of the licensed indication 97

Figures:

Figure 4-1: Mean change in HAQ....44

Figure 5-1: Cost-effectiveness diagram....50

Figure 5-2: Scatter plot (1,000 simulations, 1,000 patients)....55

Figure 5-3: Cost-effectiveness acceptability curves....56

Figure 5-4: Change in HAQ score by day 169 by baseline HAQ score for all ATTAIN patients continuing on assigned therapy to the end of the trial....63

Figure 5-5: Alternative models used to predict utility from HAQ disability scores....65

Figure 5-6: Effects of changes in HAQ in MS....67

Figure 5-7: Sensitivity of the estimated ICER for abatacept to different values of the progression rates of HAQ scores (unmodified Excel-based model)....68

Figure 5-8: Variations on HAQ scores for individual patients reproduced from Scott and Garrood[55] 70

Figure 5-9: Illustration of non-linear trends in mean HAQ scores using NOAR[51] transition rates....72

Figure 5-10: Illustration of projected trend in mean HAQ using transition probabilities weighted strongly toward progression....73

Figure 5-11: Sensitivity of the estimated ICER for abatacept to different values of the progression rates of HAQ scores (modified Excel-based model)....82

Figure 5-12: Adjusted PSA scatterplot to correspond to revised base-case scenario using the modified Excel-based model....83

Abbreviations:

ACR	American College of Rheumatology
AE(s)	Adverse event(s)
AIM	Abatacept in inadequate methotrexate responders
ARRIVE	Abatacept researched in rheumatoid arthritis patients with an inadequate anti-TNF response to validate effectiveness
ASSURE	Abatacept study of safety in use with other RA therapies
ATTAIN	Abatacept in anti-TNF inadequate responders
BMS	Bristol-Myers Squibb
BRAM	Birmingham rheumatoid arthritis model
BSR	British Society for Rheumatology
BSRBR	British Society of Rheumatology Biologics Registry
CRP	C-reactive protein
CSR	Clinical study report
DAS	Disease activity score
DMARD(s)	Disease modifying anti-rheumatic drug(s)
EMEA	European Medicines Evaluation Agency
ERG	Evidence review group
ESR	Erythrocyte sedimentation rate
EULAR	European League Against Rheumatism
FDA	Food and Drug Administration
HAQ	Health Assessment Questionnaire (HAQ) Disability Index (DI)
HRQoL	Health related quality of life
ICER	Incremental cost-effectiveness ratio
ISS	Integrated safety study
ITT	Intention to treat
LY(s)	Life year(s)
MCID	Minimum clinically important difference
MS	Manufacturer's submission
MTC	Mixed treatment comparison
MTX	Methotrexate
NDRD	National Databank for Rheumatic Diseases
NICE	National Institute for Health and Clinical Excellence
NOAR	Norfolk Arthritis Register
NSAID(s)	Non-steroidal anti-inflammatory drug(s)
PSA	Probabilistic sensitivity analysis
QALY(s)	Quality adjusted life year(s)
RA	Rheumatoid arthritis
RCT(s)	Randomised controlled trial(s)
REFLEX	Randomized evaluation of long-term efficacy of rituximab in RA
RF	Rheumatoid factor
SA	Sensitivity analysis
SAE(s)	Serious adverse event(s)
SF-36	Short-form 36
SR	Systematic review
TNFi	Tumour necrosis factor alpha inhibitor(s)

Definition of terms:

ACR20/50/70	A specified percentage improvement (20, 50, 70%) 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 ($\leq$2.6), low disease activity ($\leq$ 3.2), moderate disease activity (> 3.2 and $\leq$ 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	A European organisation which represents the viewpoints of patients, health professionals and scientific societies 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
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 RA
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 (ERG) 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 National Institute of Health and Clinical Evidence (NICE) from Bristol-Myers Squibb (BMS) in support of the use of abatacept for the treatment of severe rheumatoid arthritis (RA) following failure of at least two previous drug modifying anti-rheumatic drugs (DMARDs) and at least one tumour necrosis factor alpha inhibitor (TNFi), when compared to current standards of care (i.e. treatment with conventional DMARDs). At the time of writing the MS, NICE had not yet issued guidance on the use of rituximab or the sequential use of TNFi for adult patients with RA.

The manufacturer's submission (MS) provides a detailed and generally accurate background to the underlying health problem and the current service provision for the relevant patient population.

1.2 Summary of submitted clinical effectiveness evidence

The MS provides clinical evidence from one randomised, placebo-controlled, double blind trial (ATTAIN) that compares the effects of abatacept plus DMARDs with placebo plus DMARDs, in a study population of 391 patients with moderate to severe RA. Data from the long-term extension (LTE) study of ATTAIN are provided to assess the clinical efficacy of abatacept over a period of up to two years. Data from other randomised controlled trials (RCTs) and LTEs are pooled to investigate the safety of abatacept. Evidence from an indirect comparison of rituximab versus abatacept is included in an appendix.

The ATTAIN trial showed that in patients with RA who have failed a TNFi, abatacept plus DMARDs is more effective than placebo plus DMARDs; 50.4% of patients in the abatacept group reached an ACR20 response at six months compared to 19.5% of patients in the placebo group ($p=<0.001$). Furthermore, the proportion of patients achieving an improvement of ≥ 0.3 on the Health Assessment Questionnaire scale, also achieved a statistically significant difference between the two study groups (47.3% versus 23.3% in the abatacept and placebo groups respectively, $p \leq 0.001$). At the end of the trial period of six months, all secondary clinical efficacy outcomes, including ACR50 and ACR70 responses, were significantly different between the two

groups ($p<0.05$) in favour of abatacept. Results from the long-term extension period of ATTAIN suggest that the clinical efficacy of abatacept was maintained with 56.2% of patients achieving an ACR20 after 18 months of open label period.

Pooled safety data from five RCTs showed that although patients receiving abatacept did report slightly more adverse events (AEs) than placebo, the differences were not statistically different. Malignancies were analysed separately and again whilst patients treated with abatacept reported more malignancies than would be expected in the general population, the rates were no higher than those reported in the placebo arm and the rates expected in the wider RA population.

1.3 Summary of submitted cost-effectiveness evidence

In the absence of UK-based economic evaluations of abatacept, the manufacturer conducted a de novo economic evaluation. The manufacturer built their economic model in R and also provided a simplified Microsoft Excel version of the model for validation. The principal analysis compares abatacept + methotrexate (MTX) versus MTX. An additional analysis compares abatacept versus a cycled TNFi. An economic model was developed to estimate the costs and outcomes of typical RA patients from the beginning of a specific treatment, after having failed a TNFi, until death. The model structure reflects the clinical outcomes of a phase III RCT of abatacept (ATTAIN),1 published economic evaluations, and expert opinion from clinicians, statisticians and health economists.

The manufacturer reports an incremental cost-effectiveness ratio (ICER) of £25,395 per quality adjusted life year (QALY) gained for the comparison of abatacept + MTX versus MTX. The manufacturer reports an ICER of £22,628 per QALY gained for the comparison of abatacept + MTX versus cycled TNFi. Results of the probabilistic sensitivity analysis (PSA) conducted by the manufacturer suggest that, based on the assumptions made and evidence available, abatacept has a high probability of being cost-effective at a willingness to pay (WTP) of £30,000 per QALY gained.

In the economic model, there are several logic errors (misunderstanding of discounting technique, application of incorrect formula, only including MTX costs in the MTX arm and no half-cycle correction) and uncertain parameter value estimates (use of clinical evidence base for abatacept annual discontinuation rate, choice of HAQ mortality multiplier, omission of non-steroidal anti-

inflammatory drug (NSAID) and corticosteroid use, drugs and hospital disease related costing methods and representation of abatacept and TNFi treatment effects).

The economic model also required some structural adjustments to be made. Firstly, the ERG constructed an overall mixed gender cohort for the comparisons, as is the norm. Secondly, the ERG had concerns about the use of utility values derived from the US model. Finally, an alternative consideration of evidence for progression rates for HAQ scores was applied in the model.

After model assumptions are corrected and/or adjusted, the ICER for the base-case comparison ranges from £47,503 per QALY gained to £72,865 per QALY gained. The ICER for the abatacept + MTX versus cycled TNFi ranges from £50,222 per QALY gained to £67,459 per QALY gained.

1.4 Commentary on the robustness of submitted evidence

1.4.1 Strengths

The company makes a convincing case, citing strong evidence from a high quality trial, of the clinical benefit of abatacept plus DMARDs compared to placebo plus DMARDs in the treatment of moderate to severe RA following failure of at least two conventional DMARDs and one TNFi. The evidence is particularly convincing given that this specific population is difficult to treat and has severe disabling disease with marked impairment of quality of life.

1.4.2 Weaknesses

Due to the lack of available clinical evidence, the MS could not provide direct evidence of the clinical effectiveness of abatacept compared to a second TNFi or rituximab. However, the main source of weakness in the MS is related to the economic model submitted by the manufacturer. The ERG has identified a number of different areas where it has been appropriate to correct or revise model assumptions.

1.4.3 Areas of uncertainty

As there are no published RCTs of abatacept versus any other relevant comparator (e.g. second or third TNFi, or rituximab) there is uncertainty around the clinical and cost effectiveness of abatacept in comparison to other relevant treatments for patients with severe RA who have failed therapy including a prior TNFi.

There is also uncertainty around the long-term progression of disease and its effect on HAQ scores. Due to the relatively recent introduction of abatacept in this patient population, there is also a paucity of long-term evidence for both the continued benefit of abatacept and its long-term comparative safety.

1.5 Key issues

Due to the lack of available evidence the MS could not provide direct verification of the clinical effectiveness of abatacept compared to a second TNFi or rituximab. The submission did not therefore examine the optimal sequencing of abatacept with conventional and biologic DMARDs. Furthermore, corrections and amendments to the economic model result in abatacept + DMARDs not being a cost-effective treatment option for patients with moderate to severe RA.

2 BACKGROUND

2.1 Critique of manufacturer's description of underlying health problem

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

A summary of the manufacturer's description of the underlying health problem is provided in Box 1 and Box 2.

Box 1: Summary of the manufacturer's description of underlying health problem (1)

1. RA affects between 0.5% and 1% of the population - approximately 400,000 people in England and Wales.[2-5]
2. Onset of RA is most common in individuals during their 40s.[2-5]
3. RA is a chronic systemic autoimmune disorder, which is primarily characterised by inflammation and swelling of multiple synovial joints.
4. The primary RA symptoms of pain, fatigue and disability are chronic and related to the underlying inflammatory disease process.
5. Co-morbid conditions such as hypertension, depression, gastroenterology diseases, and respiratory disease are common in patients with RA.[6]
6. Patients with RA have a reduced life expectancy.[7-11]
7. Quality of life for patients with RA has been shown to be as poor as in patients with congestive heart failure and advanced diabetes.[12]
8. The economic burden of RA is also substantial, especially since onset occurs most commonly in individuals during some of the most productive years of life and results in decades of disability and treatment for many patients.[13]
9. Caregivers and family members are also subject to the burden associated with caring for a chronically ill person.[14]

Box 2: Summary of the manufacturer's description of underlying health problem (2)

1. There is no cure for RA and so the therapeutic goals are a remission of symptoms involving the joints, a return of full function, and the maintenance of remission.
2. RA therapy entails the use of three classes of drugs, NSAIDs, corticosteroids and conventional and/or biologic DMARD which can reduce the number of painful and tender joints, the duration of morning stiffness, and inflammation.
3. While RA is a chronic condition, joint damage often occurs early and so there has been a shift in the management of patients with RA towards beginning disease modifying therapy immediately in order to gain early control of the disease[15]
4. Abatacept is a new biologic DMARD which was specifically developed for the treatment of autoimmune diseases, which although diverse in organ target and disease manifestation, have the same general T cell-mediated aetiopathology.
5. Abatacept targets T cells which are pivotal in the initiation of autoimmune diseases and so intercepts one of the causes of disease. This subsequently prevents the downstream events which lead to the joint damage and bone erosion associated with RA, including the activation of RF-producing B cells, macrophage activation and the production of inflammatory cytokines such as TNF, IL-1 and IL-6.
6. Because abatacept is a fully human protein co-stimulation modulator preventing the full activation (rather than a depletion or complete inhibition) of T cells, it lowers the risk of immunogenicity and the overall function of the immune system is not compromised.

The manufacturer's description of the underlying health problem is detailed and generally accurate. However, four issues require comment.

The manufacturer's discussion of context (p.23-34, MS) does not provide details of the proportion of patients in the population for whom abatacept might be appropriate, i.e. those with moderate to severe RA, who have failed a TNFi. Figures reported by NICE suggests that approximately 15% of people with RA have severe RA,[16] and 30% of these will go on to fail a TNFi.[17] Details of the number of patients eligible for abatacept are only discussed in the MS in the cost-effectiveness section (p.131 MS). In this section the MS estimates that 3,585 patients in the UK would be eligible for abatacept.

The MS does not discuss the role of rheumatoid factor (RF) in this population. RF is an antibody not usually found in the blood of the general population, which is of diagnostic and prognostic significance in patients with suspected RA. Approximately 80% of RA patients are RF positive. Results of an RF test alone 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,[18] RF may be one means of selecting patients for more aggressive therapy. The Bayesian mixed treatment comparison (MTC) of abatacept and rituximab included in the appendix of the MS reports that 79% of both the abatacept and placebo groups were RF positive.

The MS states that the age of onset is most common in individuals during their 40's (Point 2, Box 1). The studies cited to support this suggest that the number of people newly diagnosed with RA remains constant or even increases between the ages of 40 and 80 years.[2-5] These studies also suggest that there are nearly three times as many women as men with RA, a fact that is not explicitly stated in Section 4 of the MS, although it is recognised in the characteristics of an average RA patient used in the manufacturer's economic model.

Finally, the published literature suggests that within five years of diagnosis, a third of people with RA are unable to work.[19] Successful treatment of severe RA would enable these individuals to return to work. The MS fails to mention this important consequence in Section 4.

2.2 Critique of manufacturer's overview of current service provision

A summary of the manufacturer's overview of current service provision is provided in Box 3-5.

Box 3: Disease modifying anti-rheumatic drugs

1. The management of RA has undergone fundamental changes in the past 15 years, reflecting a growing number of available treatments and a shift in treatment strategies.
2. Most notably, the optimal use of DMARDs has evolved over the last 10 years.
3. It is now established that prompt diagnosis and early use of DMARDs accrues long-term benefits in terms of disease modification.[20]
4. Methotrexate (MTX) has been identified as the conventional DMARD that is most likely to induce a long-term response. MTX has demonstrated efficacy and durability, long-term use and low cost and so is the most frequently prescribed for the initial treatment of moderate to severe RA.

Box 4: Biologic disease modifying anti-rheumatic drugs

1. More recently, biologic DMARDs such as adalimumab, etanercept and infliximab which are TNFi have become available for patients who fail to respond to conventional DMARDs.
2. TNFi are typically given to patients who have failed at least two conventional DMARDs (one of which should be MTX), have active disease (DAS28>5.1) on two occasions a month apart and have no contraindication to TNFi – this is consistent with both NICE guidelines[21] regarding the use of etanercept and infliximab and the British Society for Rheumatology (BSR) guidelines[18] on standards of care for people with RA.
3. Injectable and infused biologic DMARD therapies have associated limitations and risks include infections (such as tuberculosis) and increased incidence of congestive heart failure and hepatotoxicity.
4. The BSR guidelines state that treatment response to a TNFi should be assessed after three months with treatment being withdrawn in the event of non-response (defined as an improvement in DAS28 of <1.2).[18] Responders may continue with treatment but should be monitored carefully for loss of response or AEs.
5. According to expert clinical opinion, there is no universally agreed treatment pathway and TNFi therapy varies significantly between clinicians and different parts of the country.
6. Approximately 30% of people given TNFi will not tolerate known side effects of TNFi therapy or will show evidence of therapy failure over time.[17]
7. Clinical evidence shows that the proportion of patients responding and the level of response to a second TNFi agent after failure to a first TNFi is substantially lower.
8. A second TNFi is not recommended by NICE as cost effective in instances of treatment failure[21] although clinical evidence suggests that in practice 46% of TNFi failures are given a second TNFi[17]
9. There is limited evidence for the clinical and cost effectiveness of conventional DMARDs after TNFi failure and considerable uncertainty about best practice.
10. A large unmet need exists for agents such as abatacept and rituximab, neither of which are yet a standard treatment of care in RA and both are currently being reviewed by NICE.
11. Rituximab binds to the CD20 antigen on the B lymphocyte causing B cell depletion. Because of concerns of using a T cell modulator (i.e. abatacept) in patients who are B cell depleted due to the adaptive arm of the immune system being severely compromised, it is suggested that abatacept may be more appropriate for sequential use before rituximab.

Box 5: Relevant guidelines

1. Abatacept has been approved for the treatment of RA in the US, Canada, Argentina, Peru, Macau and the European Union.
2. In addition to the NICE and BSR guidelines mentioned above, other relevant guidelines include:
   - NICE guidance[22] for the use of anakinra (another biologic DMARD) which is licensed for use in RA with MTX but only approved for patients taking part in long-term clinical and cost-effectiveness studies and/or patients who were already taking it at the time guidance was issued.
   - NICE guidelines for the use of adalimumab, etanercept and infliximab are expected in 2007 but are undergoing appeal. These do not differ significantly from the existing NICE guidelines on the use of etanercept and infliximab.[21]
   - American College of Rheumatology guidelines for the management of RA which reiterate the beneficial use of etanercept and infliximab for patients with active RA.[23]
   - Australian Rheumatology Association guidelines for the use of biological agents (adalimumab, etanercept, infliximab and anakinra) in the treatment of RA which do not significantly differ from NICE guidelines other than seemingly permitting the use of anakinra.[24]

The manufacturer's overview of current service provision is complete although some discussion around specific points is required.

In the MS it states that approximately 30% of patients receiving TNFi will discontinue TNFi therapy due to AEs or lack of efficacy (Point 6, Box 4), and 45% of these will receive a second TNFi (Point 8, Box 4). The ERG notes that this corresponds to approximately 13% of all patients receiving TNFi therapy. Furthermore, the ERG notes that 40%[17] of patients receiving a second TNFi would have discontinued their first TNFi therapy due to AEs, not lack of efficacy, and therefore a switch to a second TNFi may have been permitted under NICE guidance. In conclusion, the percentage of patients treated with a TNFi who go on to receive a second TNFi contrary to current NICE guidance is approximately only 7%.[17]

Furthermore, patients can receive abatacept after the failure of two DMARDs and one TNFi. However, in current clinical practice it is clear that patients with long standing disease have received more than two DMARDs. Whether or not the number of failed DMARDs will affect the likelihood of patient response to abatacept is unknown.

Whilst the MS states that the proportion of patients responding and the level of response to a second TNFi agent after failure of a first TNFi is substantially lower (Point 7, Box 4) than a response to the first TNFi, clinical evidence presented later in the MS (p. 77) cites the British Society of Rheumatology Biologics Registry (BSRBR) analyses.[25] This report shows that while a switch to a second TNFi results in a greater reduction in HAQ scores than if patients continued their first TNFi, these differences are not statistically significant. However, there was a significant difference in HAQ for patients switching to a second TNFi compared to those stopping TNFi treatment.

In conclusion, the MS provides a detailed and generally accurate background to the underlying health problem and the current service provision for patients with moderate to severe RA who have failed a TNFi.

3 CRITIQUE OF MANUFACTURER'S STATEMENT OF THE DECISION PROBLEM

The health care technology discussed in the MS is abatacept (ORENCIA®) for the treatment of patients with RA. At the time of submission, the proposed indication submitted with the Marketing Authorisation Application was as follows: 'ORENCIA® in combination with methotrexate (MTX) is indicated for the treatment of moderate to severe active RA in adult patients who have had an insufficient response or intolerance to other disease-modifying anti-rheumatic drugs including at least one tumour necrosis factor (TNF) inhibitor.' On 22 March 2007, the Committee for Medicinal Products for Human Use adopted a positive opinion, recommending that a marketing authorisation for the medicinal product abatacept be granted.

The manufacturer's statement of the decision problem is now discussed in relation to the MS.

Table 3-1: Statement of the decision problem

	Final scope issued by NICE	Decision problem(s) addressed in the submission
Intervention	Abatacept	Abatacept
Population	Adults with RA	Adults with RA
Comparator(s)	Management strategies without abatacept, for example: Alternative DMARDs, including TNFi agents (such as adalimumab, etanercept and infliximab) and rituximab	Management strategies without abatacept, for example: Alternative DMARDs, including TNFi agents (such as adalimumab, etanercept and infliximab) and rituximab
Outcomes	The outcome measures to be considered include: • disease activity • physical function • joint damage • pain • mortality • fatigue • Health related quality of life (HRQoL) • adverse effects of treatment	The outcome measures included: • disease activity • physical function • joint damage • pain • mortality • fatigue • HRQoL • adverse effects of treatment
Economic analysis	The reference case stipulates that the cost-effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year (QALY).	The cost-effectiveness of treatment with abatacept is expressed in terms of incremental cost per QALY.
	The time horizon for the economic evaluation should reflect the chronic nature of the condition.	Time horizon for economic evaluation reflects the chronic nature of RA: analyses are presented for 5 years and 20 years.
	Costs will be considered from an NHS and Personal Social Services (PSS) perspective.	Costs will be considered from an NHS and PSS perspective.
Special considerations and other issues	If the evidence allows, the appraisal will attempt to identify criteria for selecting patients for whom abatacept would be particularly appropriate. If the evidence allows, the appraisal will attempt to examine the optimal sequencing of abatacept with DMARDs, to include (if the licensed indication allows) the use of more than one TNFi agent in a sequence and a comparison with rituximab following the failure of other TNFi agents.	If the evidence allows, the submission will attempt to identify criteria for selecting patients for whom abatacept would be particularly appropriate. If the evidence allows, the submission will attempt to examine the optimal sequencing of abatacept with DMARDs, to include the use of more than one TNFi agent in a sequence and a comparison with rituximab following the failure of other TNFi agents.
	The intervention will be appraised according to its anticipated licensed indication. Guidance will only be issued in accordance with the marketing authorisation.	For this appraisal the intervention is as described in the licensed indication.

3.1 Population

The manufacturer's statement of the decision problem describes the relevant population i.e. adults with RA. However, the licensed indication only includes adult patients with moderate to severe active RA who have had an insufficient response or intolerance to other DMARDs including at least one TNFi. This narrower population is the only population considered in the MS.

3.2 Intervention

The health care technology outlined in the statement of the decision problem and considered in the MS is abatacept. Although the licensed indication is abatacept in combination with MTX the clinical evidence reported in the MS is in combination with a range of concomitant DMARDs, with 76% of abatacept patients receiving MTX and 24% a range of alternative non-biologic DMARDs. Due to the difference in the licensed indication and the evidence cited by the manufacturer, the ERG requested data on each of the relevant clinical outcomes sub-grouped for patients receiving MTX versus those receiving other DMARDs or anakinra. The results show that there are no differences in the co-primary outcomes of patients receiving MTX, other DMARDs or anakinra.

3.3 Comparators

The decision problem in the MS states that the comparators considered will include management strategies without abatacept. The MS assumes that abatacept + DMARDs is a final treatment option for patients as it is not placed in a sequence of treatments.

The MS accurately identifies three categories of comparators. Firstly, a return to non-biologic DMARDs, secondly, cycled TNFi therapy and lastly rituximab.

3.3.1 Return to non-biologic DMARDs

The main clinical evidence used to support the manufacturer's statement of the decision problem includes patients receiving a range of DMARDs after failure of a TNFi, thereby comparing abatacept to a return to non-biologic DMARDs as reflected in clinical practice. However, the clinical efficacy of abatacept versus any specific DMARD other than MTX (e.g. leflunomide) is not known.

3.3.2 Cycled TNFi

Although current NICE guidance does not recommend sequential treatment with TNFi, the MS provides evidence from the BSRBR,[25,] that in clinical practice, patients who fail a TNFi are often treated with a second TNFi. The ERG acknowledges that this is widely accepted as common practice (Tom Kennedy, personal communication, March 2007) and therefore a second TNFi can be considered a relevant comparator. However, the submission reports that no relevant clinical trial evidence is available to ascertain abatacept's clinical effectiveness over additional TNFi treatment. The economic model therefore considers a generic TNFi with clinical effectiveness assessed using the change in HAQ score at 12 months according to the BSRBR.[25] The manufacturer stresses that "This analysis should be considered speculative as the data from the BSRBR are not available in a way to make easy inclusion in the model and some assumptions were required."(p. 96 MS)

3.3.3 Rituximab

Whilst the MS acknowledges rituximab as a potentially relevant comparator to abatacept, it does not consider rituximab to be routinely prescribed in the UK and therefore does not include it as a formal comparator. However, for completeness, the MS does include a MTC of the clinical effectiveness of abatacept and rituximab as shown in Appendix 8.7 of the MS.

Thus in summary, whilst the MS does identify all the relevant comparators the limited evidence available does not allow all relevant analyses.

3.4 Outcomes

The decision problem outlines eight relevant clinical outcomes to be assessed and all of these are stated to have been addressed in the MS. The measures used to assess each of these clinical outcomes as described in the ATTAIN trial are shown in Table 3-2. The ATTAIN trial is the principal trial supporting the MS.

Table 3-2: Outcomes

Outcome in decision problem	Outcome measure used in ATTAIN trial
Disease activity	DAS28
Physical function	HAQ
Joint damage	Not reported
Pain	SF-36 subscale
Mortality	No. of deaths
Fatigue	100mm VAS[a]
HRQoL	SF-36
Adverse effects of treatment	Adverse events of treatment reported

[a] reported in clinical study report

Symptomatic relief, as measured by the set of American College of Rheumatology (ACR) response criteria, is frequently assessed in clinical trials of RA. Measurement of symptomatic relief was not identified for assessment in the manufacturer's statement of the decision problem but was the co-primary outcome measure evaluated in the ATTAIN trial. Joint damage was not assessed in the ATTAIN trial, however this outcome has been assessed in other longer-term trials of abatacept.

The manufacturer's statement of the decision problem appropriately measures the cost effectiveness of abatacept in terms of incremental cost per QALY gained.

3.5 Time frame

The manufacturer's time horizon for economic evaluation reflects the chronic nature of RA as the costs and benefits of treatment with abatacept are considered for 5 years and 20 years.

3.6 Other relevant factors

The statement of the decision problem proposes that, if the evidence allows, the submission will attempt to identify criteria for selecting patients for whom abatacept would be particularly appropriate. The MS states that no relevant sub-groups were identified and therefore no additional criterion for selecting patients was provided; the ERG proposes that sub-analysis based on presence of RF may have been useful. Primary clinical outcomes were analysed according to baseline history of TNFi treatment (current or prior) but these sub-groups were not powered to test for significant differences between groups. In the economic analysis ICERs are also reported for males.

The statement of the decision problem also states that if the evidence allows, the MS will attempt to examine the optimal sequencing of abatacept with DMARDs, to include the use of more than

one TNFi in a sequence and a comparison with rituximab following the failure of another TNFi. In the absence of head to head trials, the economic model considers abatacept plus MTX versus a second TNFi using evidence from the BSRBR.[35] No discussion of treatment options after abatacept was considered in the MS.

An indirect comparison with rituximab was possible. However, as the optimal strategy and timing of rituximab are currently unclear and, as yet there is no NICE guidance, this MTC of rituximab versus abatacept is only included as an appendix (Appendix 8.8 in MS). The analysis concludes that due to the different mechanisms of the drugs, abatacept is better placed before, rather than after, rituximab. This conclusion is reached because 'options for rituximab-inadequate responders are limited as treated patients are B cell depleted until repopulation occurs and may be in 'limbo' before other therapies can be safely used' (p.21 MS).

The manufacturer's statement of the decision problem reflects the final scope issued by NICE. However, the population identified therein is different from the relevant patient population as stated in the licensed indication and discussed in the MS. Abatacept plus MTX is a more accurate description of the health care technology. The available evidence is not adequate to provide answers to all of the questions raised by the manufacturer in their statement of the decision problem i.e. the clinical and cost effectiveness of abatacept versus DMARDs (non-MTX), abatacept versus rituximab.

4 CLINICAL EFFECTIVENESS

4.1 Critique of manufacturer's approach

The MS includes a systematic review (SR) of the clinical evidence available to assess the efficacy and safety of abatacept for the treatment of patients with moderate to severe RA who have failed a TNFi.

Key aspects of the methodological quality of the manufacturer's review of the clinical literature were assessed based on an accepted quality assessment tool[26] and the results are summarised in Table 4-1.

Table 4-1: Quality assessment of the clinical effectiveness review

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

N/A =not applicable

4.1.1 Description of manufacturers search strategy and comment on whether the search strategy was appropriate.

Two electronic databases were searched (Medline and EMBASE) covering the period 01/01/1990-22/08/2006. Internal manufacturer databases of clinical studies were also searched. In March 2007, an additional search of ongoing clinical trials databases was conducted.

The search strategies employed were comprehensively reported enabling replication. The ERG is confident that all relevant published clinical trials were identified by the manufacturer.

4.1.2 Statement of the inclusion/exclusion criteria used in the study selection and comment on whether they were appropriate.

Details of inclusion and exclusion criteria are provided in Table 4-2 and are considered appropriate and complete.

Table 4-2: Scope of the literature review

	Clinical effectiveness
Inclusion criteria	• RCTs published since 1990 where the full paper can be obtained. • Patients in at least one arm of the trial must receive abatacept as in the proposed indication. Comparators included any other DMARD agent or placebo (including the 'do nothing' option) or standard care. • Head-to-head trials were included. • The patients of interest are adults with moderate to severe RA. • Long-term extension studies of observational design were included. • Non-English (French, Spanish, Italian or German) publications were included.
Exclusion criteria	• Non-randomised or uncontrolled studies (unless these are long-term extensions of RCTs), observational studies, case series, letters to editor, studies with no abstracts, conference abstracts only. • Reviews were ordered for the purpose of checking the bibliographies but were excluded from the list of included studies. • Trials in diseases other than RA. • Patients with early RA were excluded as abatacept is not indicated for treatment of early RA and the scope of this submission focuses on more severe disease. • Studies reporting solely on laboratory measures aimed at investigating disease or treatment mechanisms, and which do not report relevant clinical outcomes.

4.1.3 Table of identified studies. What studies were included in the submission and what were excluded?

The search strategy resulted in the identification of 10 articles reporting six RCTs. The search strategy was then restricted to within-licence studies resulting in the inclusion of one trial[1] (ATTAIN). Trial characteristics of the one included trial are summarised in Table 4-3. Details of the five RCTs outside of the licensed indication are reported in Appendix 1.

Table 4-3: Characteristics of the one included trial (ATTAIN)

Study Name	Study intervention(s), comparator(s), drug, dose(s) and follow-up	Study design, location and enrolment	Study inclusion and exclusion criteria	Study outcomes
IM101029 (ATTAIN) Genovese 2005[1] Westhovens 2006[27]	Abatacept approximately 10mg per kg of body weight + DMARD (N=258) Placebo + DMARD (N=133) Medication was administered via a 30-minute intravenous infusion on days 1, 15, 29 and every 28 days thereafter for 6 months Study duration 26 weeks	Double-blind phase III RCT from 89 sites Dates of trial enrolment: 10/12/02-02/06/04 Patients completing the double-blind phase of the study were allowed to enter a long-term, open-label extension phase (during which all patients received an approximate dose of abatacept of 10mg per kg of body weight) Patients were stratified by current or prior use of TNFi at time of enrolment.	Patients must have/be: • Aged ≥18 years who have RA (according to ACR criteria) for ≥1 year • Taken an oral DMARD or anakinra for ≥3 months (at stable dose for ≥28 days) • Had an inadequate response to a TNFi at the approved dose after ≥3 months of treatment. • ≥10 SJC • ≥12 TJC • CRP levels of ≥1 mg/dL Patients were excluded if they had: • Active vasculitis of a major organ system • Uncontrolled renal, hepatic, haematological, gastrointestinal, pulmonary, cardiac, neurological, or cerebral disease, • History of cancer within the last 5 years, • Serious bacterial infection in the previous 3 months • Active tuberculosis or herpes zoster history • Had surgery on more than 5 joints • Women of child-bearing potential not using contraception	Primary outcomes (at 6 months): • ACR20 response • HAQ improvements of ≥0.3 • Secondary outcomes (at 6 months): • ACR50, ACR70 response • DAS28 (low level disease activity ≤3.2; remission <2.6) • HAQ mean improvement • SF-36 changes • AEs • SAEs • Changes in vital signs and laboratory tests • Immunogenicity testing on days 1, 29, 85, 169 and 85 days after last dose of abatacept

SJC=swollen joint count, TJC=tender joint count, MHAQ=modified HAQ, CRP=C-reactive protein, ESR= Erythrocyte sedimentation rate, AE = adverse event, SAE = serious adverse event, DAS28 = Disease activity score, SF-36 = Short form 36

As part of the SR, the MS also includes non-RCT evidence of (i) efficacy and safety of abatacept in the long-term, (ii) after a switch from a TNFi and (iii) assessment of efficacy of cycled TNFi, in patients with RA (Table 4-4).

Table 4-4: List of relevant non-RCT evidence

Trial no, (Acronym)	Drug dosages	Population (previous drug failure)	Objectives	Design/ duration	Justification for inclusion
029 (ATTAIN) LTE[1, 28, 29]	Abatacept	TNFi	Evaluate long-term efficacy/safety of abatacept for up to 2 years	Open label extension of ATTAIN	Provides long-term efficacy and safety data up to 2 years
064 (ARRIVE)[30]	Abatacept	TNFi	Compare safety in patients receiving abatacept with and without a washout period after a TNFi	Open label CT	Provides safety data on direct switching from a TNFi to abatacept
BSRBR[25]	TNFi	TNFi	Effect of a second course of TNFi on HAQ following lack of response to the first course	Observational	Provides assessment of efficacy of cycling TNFi

4.1.4 Details of any relevant studies that were not included in the submission?

All relevant studies were included in the MS and details of ongoing trials that are likely to be reporting additional evidence within 12 months were reported.

4.1.5 Description and critique of manufacturers approach to validity assessment

The MS includes a completed validity assessment and a JADAD score of five for the ATTAIN trial, the only RCT that met the review inclusion criteria. In addition, validity assessments for the additional five RCTs are included in Appendix 8 of the MS. The validity assessment tool used is not referenced but the questions are appropriate and complete.

The ERG agrees that the validity assessment tool used in the MS was appropriate and that all trials were of a good quality. The completed validity assessment tool for ATTAIN as reported in the MS is reproduced in Table 4-5.

Table 4-5: Validity assessment of included RCT-ATTAIN

Validity assessment		Study assessment
1	**How was allocation concealed?**	Randomisation schedules were kept sealed until unblinding of study Study was double-blind; appropriate dummy products were administered to those in placebo arm. Investigators were blinded to drug preparation in addition.
2	**Which randomisation technique was used?**	Computerised randomisation using a telephone-based system was used.
3	**Was follow up adequate?**	Follow up concurs with recommendations of FDA and EMEA.
4	**Were individuals undertaking the outcomes assessments aware of allocation?**	Protocol design ensured that those making measurements of outcome were unaware of allocation. A pharmacist was aware of allocation and performed study drug preparation independent of investigators.
5	**Was a justification of the sample size provided?**	Sample size is justified by power calculations included in the clinical study report and reproduced in brief in Section 5.3.5(of the MS). Hypothesised differences in outcomes were based on earlier studies by sponsor. The study had adequate power to test the primary hypothesis.
6	**Was the design parallel or cross over?** **Is there risk, for cross over designs, of carry-over effect?**	Parallel design.
7	**Was the RCT conducted in the UK?**	RCT was multinational, conducted in North America and Europe, but not in the UK.
8	**How patients included in the RCT compare with patients likely to receive the intervention in the UK?**	RCT populations had similar disease severity and demographic mix to patients in the UK. The setting for the RCT cited is the same as that for treatment of similar patients in UK.
9	**Are dosage regimens within those cited in the summary of product characteristics?**	Recommended dose in the draft summary of product characteristics was used in all cases.
10	**Where study groups comparable?**	Groups had similar demographic and clinical profiles.
11	**Were statistical analyses performed appropriate?**	Abatacept treatment group was compared with the placebo group. Sequential primary analyses were completed (ACR 20, physical function change HAQ) and compared using CMH Chi-square tests.

EMEA = European Medicines Evaluation Agency, FDA = Food and Drug Administration

4.1.6 Description and critique of manufacturers outcome selection

The clinical outcomes measured and reported in the ATTAIN trial included the following: ACR20, 50 and 70 at six-months, over time and sub-grouped by baseline TNFi treatment (current or prior use); HAQ scores (both the mean change and the proportion of patients achieving a clinically significant change of at least 0.3); DAS, SF-36 at six-months and AEs.

The co-primary outcomes were ACR20 and the proportion of patients achieving a clinically significant change of at least 0.3 on the HAQ scale. The MS contains extensive information regarding the validity of the clinical outcome selection, quoting both the US Food and Drug

Administration (FDA)[31] and the European Medicines Evaluation Agency (EMEA)[32] in relation to the primary outcomes (ACR20 and HAQ). However, the FDA reference[31] suggests that HAQ should be measured at two years rather than six-months as measured in the ATTAIN trial or 12 months as cited in the submission.

In addition to the ATTAIN trial, the five RCTs cited use a range of clinical outcomes, including radiographic endpoints designed to measure joint structural damage. Radiographic endpoints were not collected in ATTAIN probably due to the short trial period. Measures of fatigue were measured in the ATTAIN trial[33] but were not reported in the MS. The ERG considers all clinical outcomes relevant and appropriate.

4.1.7 Describe and critique the statistical approach used

The ATTAIN trial was powered to 96% to detect a 20% change for the primary outcome of ACR20 and 87% to detect an 18% change in HAQ scores. For binary measures, Cochran-Mantel-Haenszel chi-square tests with stratification based on baseline history of TNFi treatment (current or prior use) were used. For continuous measures, an analysis of covariance was used, with treatment as the main factor and baseline measures as the covariate. All statistical tests and confidence intervals were two sided. Subgroup analyses were not sufficiently powered to detect a difference. ***CIC removed*** All statistical methods were fully reported for each of the trials.

4.1.8 Summary statement

Although the majority of the MS reflects the licensed indication for the use of abatacept, in patients with RA who have failed a TNFi, it does not reflect the broader population outlined in the manufacturer's decision problem (patients with RA).

In relation to the licensed indication the SR reported in the MS was complete and of a high standard. The search strategy was adequately reported enabling replication. All relevant clinical trials were identified and quality assessed.

The clinical outcomes reported in the single relevant RCT identified cover all appropriate outcomes, excluding joint damage. The appropriateness of HAQ measures at six-months may be an issue for concern. Statistical methods were described in full and appropriately applied.

4.2 Summary of submitted evidence

The one relevant RCT included in the manufacturer's SR is the ATTAIN trial. ATTAIN is a phase III, multi-centre, randomised, double-blind, placebo-controlled trial to evaluate the efficacy and safety of abatacept in subjects with active RA on background DMARDs who have failed a TNFi (Table 4-3).

Of the 738 patients screened, 393 underwent randomization on a 2:1 ratio, to either receive abatacept or placebo; all patients also continued to receive background DMARDs. Groups were stratified for patients who had failed a TNFi more than two months prior to screening (prior users) or within two months of screening (current users). Two patients who were randomised did not meet eligibility criteria and did not receive treatment; they were therefore excluded from the clinical efficacy analyses. A total of 258 patients received abatacept plus DMARDs and 133 received placebo plus DMARDs. Medication of approximately 10mg per kg of body weight was administered via a 30-minute intravenous infusion on days 1, 15, 29 and every 28 days thereafter for six months.

Patients completing the double-blind phase of the trial were allowed to enter a long-term, open-label extension phase (during which all patients received a fixed dose of abatacept of approximately 10mg per kg of body weight every 28 days).

4.2.1 Summary of clinical results

Data presented in this report have been extracted from the MS, the primary published peer-reviewed clinical paper[1] and the clinical study report (CSR[33]) as provided electronically by the manufacturer. Additional information was provided by the manufacturer in clarification of questions raised by the ERG.

Details of patient characteristics are presented in Table 4-6.

Table 4-6: Baseline patient characteristics from ATTAIN[1]

		Abatacept (N=258)	Placebo (N=133)
Age - years		53.4±12.4	52.7±11.3
Weight - kg		78.2±19.0	78.2±21.0
Female		77.1%	79.7%
Race[a]	White	248 (96.1%)	124 (93.2%)
	Black	9 (3.5%)	5 (3.8%)
	Other	1 (0.4%)	4 (3.0%)
Geographic region	North America	189 (73.3%)	99 (74.4%)
	Europe	69 (26.7%)	34 (25.6%)
Duration of disease - years		12.2±8.5	11.4±8.9
No. of tender joints (68 joints)		31.2±13.0	32.8±13.4
No. of swollen joints (66 joints)		22.3±10.2	22.0±10.0
Pain score (VAS 100mm)		70.8±19.8	69.9±19.0
Physical-function score (HAQ)		1.8±0.6	1.8±0.6
Global assessment of disease activity (VAS 100mm)	Patient	69.2±19.7	69.7±20.3
	Physician	68.8±17.7	67.3±16.8
DAS28		6.5±0.9	6.5±0.8
C-reactive protein — mg/dl		4.6±4.0	4.0±3.6
Positive for rheumatoid factor		189 (73.3%)	97 (72.9%)
Use of TNFi[b]	Current	98 (38.0%)	55 (41.4%)
	Former	160 (62.0%)	78 (58.6%)
TNFi	Etanercept	83 (32.2%)	53 (39.8%)
	Infliximab	175 (67.8%)	80 (60.2%)
	Adalimumab	6 (2.3%)	2 (1.5%)
Medications at randomisation	Methotrexate	195 (75.6%)	109 (82.0%)
	Azathioprine	7 (2.7%)	3 (2.3%)
	Penicillamine	1 (0.4%)	0 (0.0%)
	Gold	0 (0.0%)	1 (0.8%)
	Hydroxychloroquine	23 (8.9%)	12 (9.0%)
	Chloroquine	0 (0%)	1 (0.8%)
	Leflunomide	23 (8.9%)	11 (8.3%)
	Sulfasalazine	18 (7.0%)	13 (9.8%)
	Anakinra	7 (2.7%)	3 (2.3%)
	NSAIDs	181 (70.2%)	95 (71.4%)
	Corticosteroids	181 (70.2%)	86 (64.7%)
Methotrexate dose at baseline — mg/wk		15.2±5.3	14.4±6.1
Median corticosteroid dose at baseline — mg/day		5	5

[a] Race was self-reported, [b] Current users i.e. patients who received TNFi therapy within 2 months of screening, without a clinical response, or had an insignificant response and a persistent DAS28 (CRP≥5.6) Prior users i.e. patients who discontinued TNFi therapy ≥2 months before screening due to lack of clinical response or an inadequate response and persistent disease activity

Table 4-7 shows the key results of the ATTAIN trial and Table 4-8 displays the results of the post-hoc subgroup analyses.

Table 4-7: Reported outcomes of ATTAIN trial (at six-months)

	Abatacept (N=256)		Placebo (N=133)		Difference	95% CI	p-value
	N	%	N	%			
Primary outcomes							
ACR 20 response*	129	50.4	26	19.5	30.8	20.6, 41.1	<0.001
HAQ improved > 0.3*	121	47.3	31	23.3	24	13.8, 34.2	<0.001
Mean HAQ change [Mean, SD]	-0.45	***CIC***	-0.11	***CIC***	-0.34	***CIC***	<0.001
Secondary outcomes							
ACR	**N**	**%**	**N**	**%**	**Difference**	**95% CI**	**p-value**
ACR 50	*CIC*	20.3	*CIC*	3.8	16.6	8.5, 20.6,	<0.001
ACR 70	*CIC*	10.2	*CIC*	1.5	8.7	2.7, 14.6	0.003
DAS28 *CIC*	**N**	**%**	**N**	**%**	**Difference**	**95% CI**	**p-value**
DAS28 MCID(≥1.2)	NR	71.0	NR	32.0	NR	NR	NR
DAS28 Low disease (≤3.2)	NR	17.1	NR	4.0	NR	NR	<0.001
DAS28 Remission (<2.6)	NR	10.0	NR	0.8	NR	NR	<0.001
DAS28 mean change [Mean, SD]	-1.98	-0.10	-0.71	-0.14	-1.27	-1.62, -0.93	<0.001
HRQoL (SF-36) summary scores [a]	**Mean**	**SD**	**Mean**	**SD**	**Difference**	**95% CI**	**p-value**
Physical component	*CIC*	*CIC*	*CIC*	*CIC*			
Mental component	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*
HRQoL (SF-36) subscale scores[a]	**Mean**	**SD**	**Mean**	**SD**	**Difference**	**95% CI**	**p-value**
Physical function	*CIC*	*CIC*	*CIC*	*CIC*	NR	NR	*CIC*
Physical role	*CIC*	*CIC*	*CIC*	*CIC*	NR	NR	*CIC*
Pain	*CIC*	*CIC*	*CIC*	*CIC*	NR	NR	*CIC*
General health	*CIC*	*CIC*	*CIC*	*CIC*	NR	NR	*CIC*
Vitality	*CIC*	*CIC*	*CIC*	*CIC*	NR	NR	*CIC*
Social function	*CIC*	*CIC*	*CIC*	*CIC*	NR	NR	*CIC*
Emotional role	*CIC*	*CIC*	*CIC*	*CIC*	NR	NR	*CIC*
Mental health	*CIC*	*CIC*	*CIC*	*CIC*	NR	NR	*CIC*

a Results taken from CSR, MCID = minimum clinically important difference, NR = not reported, ESR= Erythrocyte sedimentation rate, * clarification from the company showed there was no difference in the co-primary outcomes of patients receiving MTX, other DMARDs or anakinra

Table 4-8: Reported outcomes of TNFi use: sub-group analyses

Outcome	TNFi Sub-group	Abatacept (N=256)		Placebo (N=133)		Difference (95% CI)	p-value
		n	%	N	%		
ACR20	Current user	*CIC*	45	*CIC*	15	*CIC*	<0.001
	Prior user	*CIC*	54	*CIC*	23	*CIC*	<0.001
HAQ improvement of >0.3	Current user	*CIC*	43	*CIC*	22	*CIC*	0.013
	Prior user	*CIC*	50	*CIC*	24	*CIC*	<0.001

[a] = Extracted from CSR, Current users = patients who received TNFi therapy within 2 months of screening, without a clinical response, or had an insignificant response and a persistent DAS28 (CRP≥5.6) Prior users = patients who discontinued TNFi therapy ≥2 months before screening due to lack of clinical response or an inadequate response and persistent disease activity

As Table 4-7 and Table 4-8 show, all reported clinical outcomes showed a statistical difference in favour of abatacept. All of these outcomes are likely to be clinically significant for this group of patients, who have severe, difficult to control disease, and have failed a TNFi.

Critique of data reported

Data checking the MS with the CSR highlighted that not all results in the MS are from the intention to treat (ITT) population. The mean change in HAQ reported in the MS is the same

as that reported in the CSR; however in the CSR, the N values are given as CIC removed and CIC removed for the abatacept and placebo groups respectively. The ITT population should have N values of 256 and 133. Whilst this difference is likely to have little impact on the outcome, similar discrepancies regarding the DAS28 measures may.

The CSR reports that DAS28 scores were calculated in two different ways; CIC removed The values reported in the MS are those from CIC removed However, the N values are not consistent with those reported in the CSR. A comparison of the different analyses reported in the CSR is shown in Table 4-9. As can be seen from this table, the percentage values reported in the MS are misleading. If an ITT analysis was conducted then CIC removed of patients treated with abatacept achieved a minimum clinically important difference (MCID) compared to the 71% reported in the MS. In the placebo arm the difference is CIC removed to 32%. The effects these discrepancies may have on the results are unclear as no *p* values are reported.

Table 4-9: DAS28 results as presented in the CSR[a]

	Abatacept			Placebo			Mean Diff	(95% CI)	p-value
	CIC		N=256	*CIC*		N=133			
DAS28 *CIC*	**n**	**%**	**ITT[c] %**	**n**	**%**	**ITT[c] %**			
DAS28 MCID(≥1.2)	*CIC*	71	*CIC*	*CIC*	32	*CIC*			
DAS28 Low disease (≤3.2)	*CIC*	16	*CIC*	*CIC*	4	*CIC*			
DAS28 Remission (<2.6)	*CIC*	10	*CIC*	*CIC*	1	*CIC*			
DAS28 mean change [Mean, (SD)]	-1.98, (-0.10)			-0.71, (-0.14)			-1.27	-1.62, -0.93	<0.001
DAS28 *CIC*									
DAS28 MCID(>=1.2)	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*			
DAS28 Low disease (<=3.2)	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*			
DAS28 Remission (<2.6)	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*			
DAS28 mean change [Mean, (SD)]			*CIC*			*CIC*	*CIC*	*CIC*	*CIC*

a Results taken from CSR, b Results taken from MS, c percentages were calculated by the ERG, MCID = minimum clinically important difference, NR = not reported

Further evidence of the clinical effectiveness of abatacept was provided in the MS from the open-label extension of the ATTAIN trial (ATTAIN LTE).

Results from the ATTAIN LTE are shown in Table 4-10. After two years of treatment with abatacept all improvements in clinical outcomes reported at six-months were maintained. Patients who received placebo in the original six-month double-blind trial but received abatacept in the open label trial showed CIC removed in clinical outcomes by six-months.

Table 4-10: Clinical outcomes of ATTAIN long-term extension analysis[a]

	6 months in LTE (Day 365)				12 months in LTE (Day 533)				18 months in LTE(Day 730)			
	Abatacept		Previously placebo		Abatacept		Previously placebo		Abatacept		Previously placebo	
	N= *CIC*		N=99		N= *CIC*		N=99		N= *CIC*		N=99	
	n	%	N	%	n	%	n	%	n	%	n	%
ACR 20	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	56.2	NR	NR
ACR 50	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	NR	NR
ACR 70	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	NR	NR
HAQ improve ≥ 0.3	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	47.9	NR	NR
	Change[b] N= *CIC*		Change[b] N= *CIC*		Change[b] N= *CIC*		Change[b] N= *CIC*		Change[b] N=NR		Change[b] N=NR	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Pain 100mm VAS (percentage improvement)	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	37.2	2.4	NR	NR
	N= *CIC*		N= *CIC*		N= *CIC*		N= *CIC*		N=NR		N=NR	
MOS-SPI	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	-12.7	1.4	NR	NR
Fatigue 100mm VAS	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	-28.2	2.1	NR	NR
SF-36 mean changes	N= *CIC*		N= *CIC*		N= *CIC*		N= *CIC*		N=NR		N=NR	
Physical component	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	10.3	NR	NR	NR
Mental component	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	6.2	NR	NR	NR
Post-hoc analysis					N= *CIC*		N= *CIC*		N=NR		N=NR	
DAS28 Low disease activity	NR	NR	NR	NR	*CIC*	*CIC*	*CIC*	*CIC*	NR	*CIC*	NR	NR
DAS28 Remission	NR	NR	NR	NR	*CIC*	*CIC*	*CIC*	*CIC*	NR	*CIC*	NR	NR

[a] = data extracted from CSR, [b] = change from the double blind period baseline, NR = Not reported, MOS-SPI = Medical Outcomes Study Sleep Problem Index

4.2.2 Safety analyses

The MS reports safety data from the ATTAIN trial and an integrated safety study (ISS). The ISS comprises data from five RCTs and was presented to the FDA as part of the registration process. It is noted in the MS that additional safety data analysis exist from open label extensions of the five trials and two additional studies (IM101043 and ARRIVE) but as these full data have yet to be tabulated, they are not included in the pooled safety analysis within the MS. However, some comment is included in the MS about malignant neoplasms in these open-label extension periods. In the additional safety analysis AEs from the ARRIVE open-label trial are discussed.

ATTAIN and ATTAIN-LTE safety data

The ATTAIN safety analysis of 391 patients includes two patients in the abatacept group who were excluded from the clinical efficacy analysis due to protocol violations.

The main findings presented in the MS from the ATTAIN trial show that slightly more patients reported any AE, infection or infusion reactions in the abatacept group; however, none of these differences was statistically significant (Table 4-11).

Table 4-11: Adverse events reported in ATTAIN

Adverse events	Abatacept (N=258)		Placebo (N=133)		p-value
	n	%	n	%	
Any adverse event (AE) ‡	205	79.5	95	71.4	0.08
Discontinuation due to AE	9	3.5	5	3.8	0.89
Serious adverse event (SAE)	27	10.5	15	11.3	0.81
Discontinuations due to SAE	7	2.7	2	1.5	NR
Any infection	97	37.6	43	32.3	0.30
Serious infections†	6	2.3	3	2.3	0.97
Discontinuations due to infections	NR	0.8	NR	1.5	0.61
Acute infusion reactions	13	5.0	4	3.0	0.35
Deaths	1	0.4*	0	0	1.00

‡ Events were defined as any new or worsening illness, sign, symptom or clinically significant abnormality in a laboratory test noted by the investigator during the course of the trial, regardless of the cause, † One patient in the abatacept group and two patients in the placebo group had two serious infections, * One patient died of myocardial infarction and congestive heart failure, an event considered by the investigator to be unrelated to the study drug, NR= Not reported

The only statistically significant difference in AEs described was in the proportion of patients reporting headaches (12.4% and 5.3% of patients in the abatacept and placebo groups respectively, p=0.03). There was no difference in the proportion of patients with infections and most infections were considered mild to moderate; infusion reactions were also reported as mild to moderate with the most common being dizziness (1.6%) and headache (1.2%). The

only death in the trial was in the abatacept group; however, this was not deemed to be related to abatacept.

The ATTAIN-LTE open-label trial showed that rates of serious adverse events (SAEs) were consistent between the six-month double-blind period and the eighteen months open label period, with SAE rates reported as 34.5 and 29.4 /100 person–years, respectively.

ISS safety data

The ISS safety analysis included data from five clinical trials of abatacept (ATTAIN, AIM, ASSURE, IM101100, and IM101101) details of which are shown in Table 4-3. Of the 2,994 patients included in the ISS analysis, 1,955 patients had received abatacept (1,697 for one-year and 258 for six-months) and 989 had received placebo (856 for one-year and 133 for six-months).

Table 4-12: Adverse events in double-blind controlled trial periods

	Abatacept N=1955		**Placebo N=989**	
	n	**%**	**n**	**%**
Any AE	1,736	88.8	840	84.9
Related AE	1,013	51.8	456	46.1
Discontinuation due to AE	107	5.5	39	3.9
SAE	266	13.6	122	12.3
Related SAE	58	3	17	1.7
Discontinuations due to SAE	53	2.7	16	1.6
Any infection	1,051	53.8	478	48.3
Serious infections	58	3	19	1.9
Discontinuations due to infections	24	1.2	10	1.0
Deaths*	9	0.5	6	0.6

* One additional death was reported after the database lock for the double-blind period and is not reflected in the table

Whilst the majority of patients experienced some kind of AE, few patients discontinued treatment as a result (5.5% and 3.9% in the abatacept and placebo arms of the pooled trials respectively). The most commonly reported AEs were headaches (18% and 13%) and nasopharyngitis (12% and 9% in both the abatacept and placebo arms respectively).

The most common infections were upper respiratory tract infection (12.7% and 12.0%), and nasopharyngitis (11.5% and 9.1%). Serious infections were reported in 3.0% of patients treated with abatacept and 1.9% of patients treated with placebo; not one particular infection accounted for 1% of all serious infections.

Reported AEs rates in the ISS (Table 4-12) are higher than those reported in the ATTAIN trial (Table 4-11 and Table 4-15). This is possibly due to the inclusion of trials with differing doses of abatacept and concomitant medications in the ISS analysis.

As malignancy is a known risk of immunosuppressive therapy, the MS pays particular attention to the rates of malignancies both in the double-blind controlled periods of the five trials and their open-label extensions (Table 4-13).

Table 4-13: Malignant neoplasms in abatacept-treated patients

	Double-blind period				Cumulative with open label period	
	Abatacept N=1955		Placebo N=989		Abatacept N=2688	
	n (per 100 person-years)	%	n (per 100 person-years)	%	n (per 100 person-years)	%
All malignant neoplasms	**25**	**1.3**	**11**	**1.1**	**50**	**1.9**
Non-melanoma skin cancers	**15**	**0.8**	**6**	**0.6**	**24**	**0.9**
Basal cell carcinoma	10	0.5	4	0.4	16	0.6
Squamous cell carcinoma	6	0.3	2	0.2	9	0.3
Other (Neoplasm skin)	0	0	0	0	1	0.0
Solid cancers	**9**	**0.5**	**5**	**0.5**	**21**	**0.8**
Lung neoplasm malignant	4	0.2	0	0	8	0.3
Thyroid	2	0.1	0	0	2	0.1
Breast	1	<0.1	2	0.2	2	0.1
Prostate	1	<0.1	0	0	2	0.1
Bladder	1	<0.1	0	0	1	0.0
Renal	1	<0.1	0	0	1	0.0
Endometrial	0	0	2	0.2	2	0.1
Melanoma	0	0	1	0.1	1	0.0
Cervix	0	0	0	0	1	0.0
Haematological/ lymphatic cancers	**2**	**<0.1**	**0**	**0**	**5**	**0.2**
Lymphoma	1	<0.1	0	0	4	0.1
Myelodysplastic syndrome	1	<0.1	0	0	1	0.0

To determine whether the rates of malignancies differed from the rates found in the general population and from a wider RA population (RA patients who had not been treated with a biologic DMARD), retrospective studies were conducted using data from six RA cohorts, two of which were UK based registers. These analyses concluded that whilst the rates of lymphoma and lung cancer were higher than expected compared to the general population they were in the expected range for patients with RA who had been treated with non-biologic DMARDs. Overall, the number of malignancies was consistent with the number found in both the general population and the wider RA population.

In addition to these analyses, a SR was conducted by the manufacturer to assess the risk of site-specific malignancies in patients with RA. The review concluded that patients with RA are at a greater risk of lung cancer and lymphoma but at a reduced risk of colorectal cancers.[34] This further supports the evidence, that the rates of lung cancer and lymphoma, seen in the

trials with abatacept, are no higher than would be expected in a cohort of RA patients. However, the safety of abatacept, as with many new biologics, continues to be monitored.

Additional safety data

Further safety analysis is provided from the open-label trial ARRIVE that was conducted to assess the safety and tolerability of abatacept with or without a washout period after TNFi; thus reflecting clinical practice where physicians would need to switch patients from TNFi directly to abatacept.

The ARRIVE trial was an international, six-month, open-label phase IIIb trial of 1,285 patients with active RA. A sub-group analysis of 842 US patients who were treated and evaluated for safety has been submitted for presentation at the European League Against Rheumatism (EULAR) conference in 2007. The results of this analysis were included in the MS and are discussed here. The sub-group analysis included 370 patients who had discontinued TNFi at least two months prior to screening and 472 patients who had failed but not discontinued TNFi in the two months before screening. The trial determined that at six-months, the frequency of AEs, SAEs, discontinuations due to AEs/SAEs, infections, neoplasms and deaths were similar in patients whether they were prior or current users (Table 4-14).

Table 4-14: Frequency of adverse events from day 1 through 169 in trial IM101064 (ARRIVE)

	Prior users N=370		**Current users N=472**	
	n	**%**	**n**	**%**
AEs	284	76.8	363	76.9
Discontinuations due to AEs	15	4.1	19	4.0
SAEs	34	9.2	36	7.6
Discontinuations due to SAEs	8	2.2	4	0.8
Infections	149	40.3	181	38.3
Serious infections	8	2.2	11	2.3
Neoplasms*	3	0.8	2	0.4
Deaths	1	0.3	0	0.0

* SAEs (benign, malignant and unspecified)

Critique of safety data

The number of drug related AEs in the ATTAIN trial are not reported in the MS but are shown in the CSR and summarised in Table 4-15. Although no *p* values are reported it would appear that there were CIC removed drug related AEs reported in the abatacept group compared to the placebo group CIC removed versus CIC removed respectively). This

difference is primarily due to the number of abatacept treated patients reporting CIC removed CIC removed compared to CIC removed in the placebo group.

Table 4-15: Subjects with most frequently[a] reported drug related adverse events (ATTAIN)

	Abatacept	N=258	Placebo	N=133
	n	%	n	%
Total subjects with drug related AEs	CIC	CIC	CIC	CIC
CIC	CIC	CIC	CIC	CIC
CIC	CIC	CIC	CIC	CIC
CIC	CIC	CIC	CIC	CIC
CIC	CIC	CIC	CIC	CIC
CIC	CIC	CIC	CIC	CIC
CIC	CIC	CIC	CIC	CIC
CIC	CIC	CIC	CIC	CIC
CIC	CIC	CIC	CIC	CIC
Total patients with drug related SAEs	CIC	CIC	CIC	CIC

a=at least 2 % of subjects in any treatment group

Long-term safety data are of high importance when evaluating the safety of new drugs. At present, the longest available published safety data for abatacept are two years and these are derived from one trial. While the SAE incidence rate was lower at two years than at six-months, there remain questions about long-term safety, especially with regard to lung cancer and lymphoma where evidence[34] suggests these rates may be higher than in the general population although comparable to rates in the general RA population. However, the manufacturer has made available summary 5-year safety data from the longer term extension study of trial IM10100 CIC removed This shows that in most cases incidence rates were CIC removed in the cumulative period compared with the double-blind period CIC removed. The incidence of only CIC removed in the cumulative period compared with the double-blind period. A total of CIC removed patients reported autoimmune events and CIC removed reported malignancies during the cumulative period.

4.2.3 Critique of submitted evidence syntheses

Whilst the MS states that no direct or indirect comparisons of abatacept versus TNFi or rituximab were possible, it does include clinical evidence from three distinct sources. One trial is outside of the licensed indication and considers abatacept versus infliximab (IM101043). Another is an analysis of BSRBR[25] data regarding the efficacy of sequential TNFi. Finally, the results of the MTC of abatacept and rituximab are included in Appendix 8. The evidence from these sources are summarised in the following sections.

Clinical efficacy of abatacept versus infliximab (IM101043)

Details of trial IM101043 are shown in Appendix 1. The MS states that the data presented in Table 4-16 show that over the six-month placebo-controlled trial period abatacept and infliximab demonstrated similar efficacy but that over one year, treatment with abatacept was associated with a sustained clinical response whereas the clinical response with infliximab was less durable. However, it is not possible to ascertain how robust these differences are.

Table 4-16: Efficacy results of trial IM101043 reproduced from the MS pg. 172

	Abatacept + MTX N = 156	**Infliximab + MTX N = 165**	**Placebo + MTX* N= 110**	**Difference (95% CI)**	**P value**
DAS28	**Mean (SE)**	**Mean (SE)**	**Mean (SE)**		
Adjusted mean change from baseline	-2.53 (0.12)	-	-1.48 (0.15)	-1.04 (-1.42, -0.67)	p<0.001
Adjusted mean change from baseline	-	-2.25 (0.12)	-1.48 (0.15)	-0.77 (-1.14, -0.39)	p<0.001
Difference in DAS28 area under the curve in infliximab and abatacept groups at 1 year.	N=150, 1631.6 (31.6)		N=156 1664.3 (16.0)	-32.7 (-119.9, 54.6)	NR
HAQ- DI	**Mean**	**Mean**	**Mean**	**Difference from infliximab 95% CI**	
Adjusted mean change from baseline (6 months)	-0.69	-0.61	-0.31	NR	NR
Change from baseline (12 months)	-0.67	-0.59	-0.56	-0.08 (-0.22,0.06)	NR
ACR (6 months)	**%**	**%**	**%**	**Difference from infliximab 95% CI**	
ACR20	67	59	42	7.3 (-3.9,18.5)	NR
ACR50	40	37	20	3.4 (-7.9,14.7)	NR
ACR70	21	24	9	-3.7 (-13.5,6.0)	NR
ACR (12 months)	**%**	**%**	**%**	**Difference from infliximab 95% CI**	
ACR20	72	56	68	16.7 (5.5,27.8)	NR
ACR50	46	36	51	9.1 (-2.2,20.5)	NR
ACR70	26	21	29	5.7 (-4.2,15.6)	NR
EULAR (12 months)	**%**	**%**	**%**	**Difference from infliximab 95% CI**	
Good	32	19	28	NR	NR
Moderate	41	45	49	NR	NR
Non responders	27	36	24	NR	NR
Adjusted mean change in HRQoL, SF-36 (6 months)	**%**	**%**	**%**	**Difference from infliximab 95% CI**	
Physical	8.36	7.66	4.34	0.7 (-1.19,2.58)	NR
Mental	5.14	4.32	1.64	0.83 (-1.33,2.98)	NR
Adjusted mean change in HRQoL, SF-36 (12 months)	**%**	**%**	**%**	**Difference from infliximab 95% CI**	
Physical	9.52	7.59	8.00	1.93 (0.02,3.84)	NR
Mental	5.96	4.03	5.85	1.92 (-0.3,4.15)	NR

* After 6 months patients receiving placebo were reallocated to receive abatacept, NR= not reported

Clinical efficacy of cycled TNFi (BSRBR[25])

The MS includes a summary of a data analysis conducted in 2006 by the BSR, using the data available in the BSRBR.[25] The analyses were conducted to provide additional information for the Birmingham Rheumatoid Arthritis Model (BRAM), which evaluated the efficacy of a second course of TNFi on HAQ following a lack of response to the first course.[25, 35] At the time of the analysis the database contained information on 12,615 patients of whom 808 were included in the analysis. The MS emphasizes that "the database is observational and therefore all the inherent potential for bias in that study design applies." (p.77 MS). The results of the analyses (Table 4-17 and Figure 4-1) showed that while a switch to a second TNFi resulted in a greater reduction in HAQ than if patients continued their first TNFi, these differences were not statistically significant. However, there was a significant difference in HAQ for patients switching to a second TNFi compared to those stopping TNFi.

The ERG notes that in relation to the economic model comparing abatacept and cycled TNFi, the MS refers to the unadjusted mean change in treatment effect. However, in the BSRBR report[35] the authors go on to calculate an adjusted mean change in treatment effect removing bias due to covariates (see Section 5.5.1).

Table 4-17: Mean change in HAQ – BSRBR, NICE study

Treatment group	Mean change in HAQ (95% CI)
Same TNFi continued	–0.07 (–0.03, –0.11)
TNFi stopped	–0.01 (–0.07, 0.06)
Different TNFi used	–0.12 (–0.17, –0.07)
Different TNFi used (>6 months)	–0.15 (–0.23, –0.07)

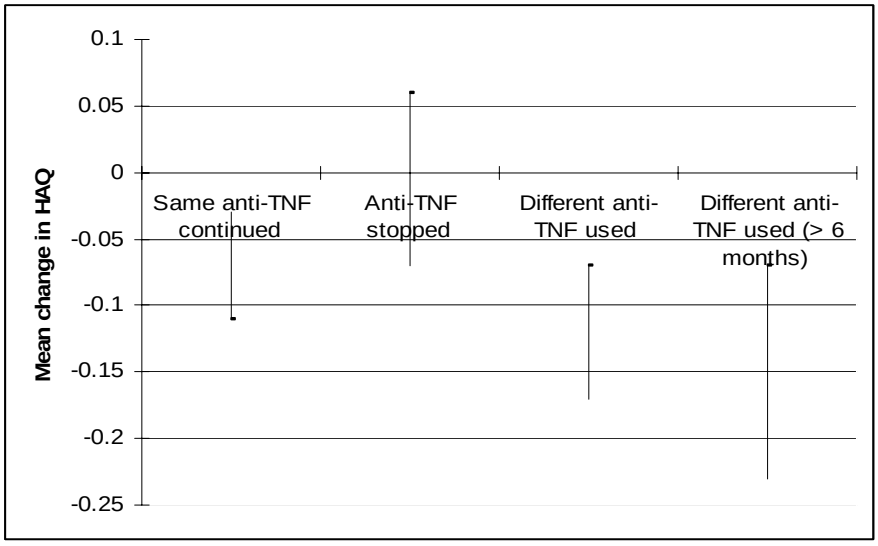


Figure 4-1: Mean change in HAQ

Clinical efficacy compared to rituximab

A MTC of abatacept and rituximab was only included in an appendix of the MS. For the comparison of abatacept versus rituximab the Cohen paper[36] reporting the REFLEX trial of rituximab and the Genovese paper[1] reporting the results of the ATTAIN trial of abatacept were used.

Table 4-18: Reflex trial versus ATTAIN mixed treatment comparison results

	Comparison or outcome	2.5%	Point estimate	97.5%	More effective treatment
ACR20	**Placebo vs. abatacept**	2.65	4.24	7.05	Abatacept
	Placebo vs. rituximab	3.06	4.62	7.20	Rituximab
	Rituximab vs. abatacept	0.48	0.92	1.79	No sig difference
ACR50	**Placebo vs. abatacept**	2.90	7.03	20.12	Abatacept
	Placebo vs. rituximab	3.72	7.21	15.40	Rituximab
	Rituximab vs. abatacept	0.30	0.98	3.45	No sig difference
ACR70	**Placebo vs. abatacept**	2.42	8.79	52.68	Abatacept
	Placebo vs. rituximab	4.32	14.99	80.38	Rituximab
	Rituximab vs. abatacept	0.07	0.58	5.10	No sig difference
HAQ	**Placebo vs. abatacept**	-0.47	-0.34	-0.22	Abatacept
	Placebo vs. rituximab	-0.38	-0.30	-0.22	Rituximab
	Rituximab vs. abatacept	-0.19	-0.04	0.11	No sig difference
Withdrawals due to AEs	**Placebo vs. abatacept**	0.33	0.97	3.20	No sig difference
	Placebo vs. rituximab	0.70	3.20	23.47	No sig difference
	Rituximab vs. abatacept	0.03	0.30	2.01	No sig difference
Withdrawals for any reason	**Placebo vs. abatacept**	0.27	0.46	0.78	Abatacept
	Placebo vs. rituximab	0.17	0.26	0.38	Rituximab
	Rituximab vs. abatacept	0.93	1.81	3.49	No sig difference
SAEs	**Placebo vs. abatacept**	0.48	0.94	1.88	No sig difference
	Placebo vs.rituximab	0.31	0.58	1.09	No sig difference
	Rituximab vs. abatacept	0.63	1.61	4.08	No sig difference

Sig= significant at p=0.05

As Table 4-18 shows, the MTC found no significant differences between abatacept and rituximab at six months for either the clinical effectiveness outcomes (e.g. ACR20 or HAQ) or the reasons for withdrawals from treatment.

4.3 Summary of clinical evidence

4.3.1 Clinical results

The ATTAIN trial showed that in patients with RA who have failed TNFi treatment, abatacept plus a DMARD is more effective than placebo plus a DMARD; 50.4% of patients in the abatacept group reached an ACR20 response at six months compared to 19.5% of patients in the placebo group (p=<0.001).

The proportion of patients achieving an improvement of ≥0.3 on the HAQ scale, also achieved a statistically significant difference between the two study groups (47.3% versus 23.3% in the abatacept and placebo groups respectively, p≤0.001).

At the end of the trial period of six months, all secondary clinical efficacy outcomes, including ACR50 and ACR70 responses, were significantly different between the two groups (p<0.05) in favour of abatacept.

Results from the long-term extension period of ATTAIN suggest that the clinical efficacy of abatacept was maintained with 56.2% of patients achieving ACR20 after 18 months of open label period.

Pooled safety data from five RCTs showed that although patients receiving abatacept did report slightly more AEs than placebo, the differences were not statistically different. Malignancies were analysed separately and again whilst patients treated with abatacept reported more malignancies than would be expected in the general population, the rates were no higher than those reported in the placebo arm and the rates expected in the wider RA population.

4.3.2 Clinical issues

Due to the lack of available evidence, the MS could not provide direct evidence of the clinical effectiveness of abatacept compared to a second TNFi or rituximab. The submission did not therefore examine the optimal sequencing of abatacept with conventional and biologic DMARDs.

As the ATTAIN trial did not measure radiographic endpoints, the evidence for the effectiveness of abatacept at reducing joint damage was reported from an additional trial of abatacept that was conducted outside the terms of the licence.

The co-primary endpoint of a clinically meaningful improvement in HAQ is reported at six months. The appropriateness and validity of this measure at six months is not discussed in the MS.

As with all biologics, the safety of these drugs is unknown, long-term data are required to accurately assess the long-term health risks of such medication. At present, the longest term data available for abatacept is limited to five years.

5 ECONOMIC EVALUATION

5.1 Summary of published cost-effectiveness analyses identified in the manufacturer's submission

A SR was conducted by the manufacturer to identify published economic models, information on costs, cost effectiveness and quality of life impact of biologic DMARDs, specifically abatacept, adalimumab, anakinra, etanercept, infliximab and rituximab.

The results of the SR were presented for (i) review of economic analyses and (ii) review of quality of life studies.

5.1.1 Identification and description of studies

The MS included full details of the electronic search strategy used in the review. The ERG could therefore replicate the electronic searching undertaken by the manufacturer. 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 type**

Cost-consequence analysis, cost-benefit analysis, cost-effectiveness analysis, cost-utility analysis, cost study, QoL study (for QoL review)

- **Condition**

Rheumatoid arthritis only. Other types of arthritis and autoimmune disease were excluded

- **Treatment**

Etanercept, infliximab, adalimumab, anakinra, abatacept, biologic DMARDs, TNF blockers

- **Populations**

Adults with RA. Studies on children and adolescents were excluded

- **Outcomes**

Cost estimates (including unit costs, resource utilization), cost effectiveness/utility measures, QoL, utility measures (the last two for the QoL review)

- **Time horizon**

Unlimited

- **Language**

Only English language publications were considered

Identified studies

Using these criteria the manufacturer identified 18 studies for inclusion in the review of economic analyses and six studies for inclusion in the review of quality of life studies (conducted in July 2006). Only the Davis study[37] included abatacept as a treatment option; this study was presented as a poster and the manufacturer stated that there was not enough information to fully appraise the study.

After the SR was conducted, three further studies were identified for inclusion. Two[38, 39] of these were economic evaluations; the third was guidance[40] on rituximab from the Scottish Medicines Consortium. Only the Thorne[39] economic evaluation included abatacept as a comparator; this study was presented as a poster and the manufacturer stated that there was not enough information to fully appraise the study.

Data extraction

The manufacturer presented summary details (Table 6.1, p.91 of MS) of the cost-effectiveness studies (n=10) which described (i) abatacept in any country context or (ii) any other biologic used in the UK setting. All of the economic analyses and quality of life studies are also summarised (including details of study, aims, methods, results and comments/relevance) in tables in Appendix B of the MS.

Data were extracted into pre-specified tables by one reviewer. A second reviewer conducted independent data abstraction and any discrepancies were discussed.

Quality assessment

The results of each of the studies were discussed in light of the critical appraisal of its methodology. The specific critical appraisal tool employed was not stated.

5.1.2 Summary and conclusions

The SR conducted by the manufacturer appears to be well conducted. The MS included summaries of relevant studies, highlighted key issues and discussed the relevance of the studies to decision-making in the UK. Only two of the identified studies included abatacept as a comparator; neither of which could be critically appraised due to lack of data. In addition, the MS rightly states that direct or meaningful comparison of the included studies was not possible due to the fact that the economic analyses were very different. In particular, the studies were heterogeneous in terms of the modelling approaches employed, time horizons and country of origin. All of the studies were limited by the lack of available long-term clinical effectiveness data.

5.2 Overview of manufacturer's economic evaluation

In the absence of UK-based economic evaluations of abatacept, the manufacturer conducted a de novo economic evaluation. The principal analysis compares abatacept + MTX versus MTX. An additional analysis compares abatacept versus a cycled TNFi. An economic model was developed to estimate the costs and outcomes of typical RA patients from the beginning of a specific treatment, after having failed a TNFi, until death. The model structure reflects the clinical outcomes of a phase III RCT of abatacept (ATTAIN),[1] published economic evaluations, and expert opinion from clinicians, statisticians and health economists.

The manufacturer built their economic model in R and also provided a simplified Microsoft Excel version of the model for validation. As agreed with NICE, the ERG has carried out an assessment of the cost-effectiveness evidence presented by the manufacturer using the Excel-based version of the model - not the R version.

The manufacturer constructed a patient-level state simulation model (Figure 5-1) which focuses on a hypothetical cohort of 10,000 patients. Patient disability is simulated over time using six-monthly cycles. Each patient in the hypothetical cohort is "run through" the model, one at a time, to estimate outcomes for the cohort as a whole. The nature of RA is modelled at the patient level in terms of changes in HAQ scores over time. The model estimates the worsening of HAQ scores due to underlying disease progression and treatment discontinuation. The model can be run for different durations up to lifetime duration.

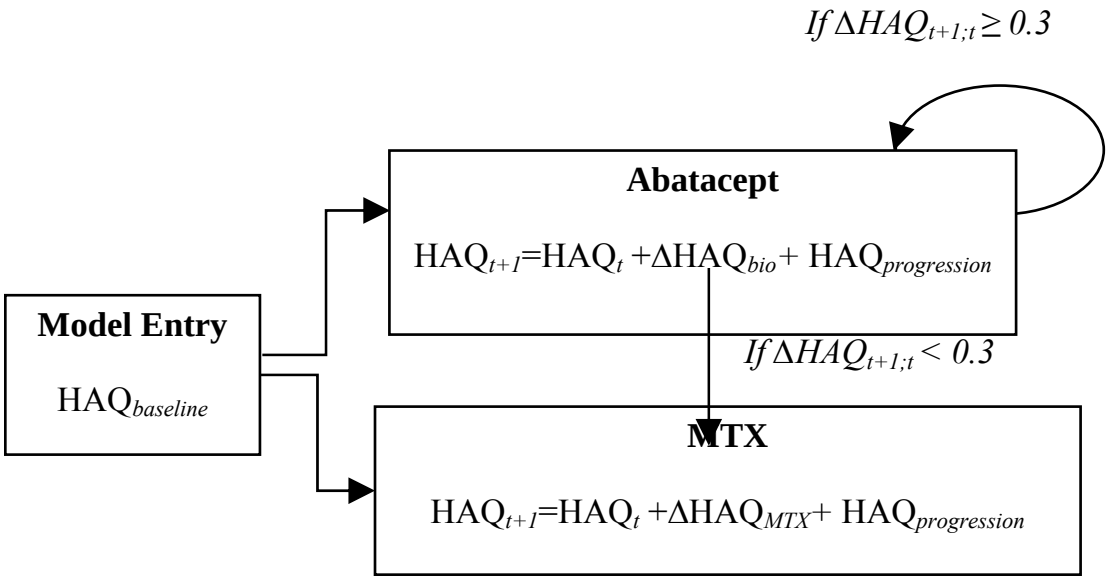


Figure 5-1: Cost-effectiveness diagram

Table 5-1: Model parameters

Treatment-specific model parameters				
Treatment	Mean % change in HAQ at 6-mths (SD)	Annual change in HAQ on treatment (%)	Annual rate of treatment failure[a] (%)	Source
Abatacept	CIC removed	0.015	8.2	ATTAIN CSR[33]
TNFi	-7.32 (11.89)[b]	0.015	8.2	
MTX	0[c]	0.06	0[d]	
Treatment-independent model parameters				
Variable/input	Value	Range/distribution		Source
Annual mortality rate multiplier for each unit of HAQ increase	1.8	Range tested in SA		Mean value of those identified via SR
Discount rate	3.5% (costs) 3.5% (outcomes)	6% and 1.5% tested in SA		NICE reference case[41]
HAQ threshold required to continue treatment	0.3	0.22-0.75 tested in SA		Co-primary endpoint in ATTAIN trial
Baseline HAQ score	1.82	Range tested in SA, including 2.05 to reflect baseline HAQ of patients in BSRBR		ATTAIN trial
Utility values and costs other than treatment costs, by HAQ category				
HAQ category	Mean utility (SD	Other costs (£, per patients, per 6-mths)[e]		
0-<0.25	0.857 (0.16)	148.07		
0.25-<0.5	0.803 (0.13)	310.20		
0.5-<0.75	0.762 (0.14)	310.20		
0.75-<1.00	0.713 (0.15)	310.20		
1.00-<1.25	0.657 (0.15)	656.86		
1.25-<1.5	0.590 (0.18)	656.86		
1.5-<1.75	0.511 (0.19)	656.86		
1.75-<2.00	0.427 (0.21)	656.86		
2.00-<2.25	0.333 (0.24)	1843.08		
2.25-<2.5	0.229 (0.25)	1843.08		
2.5-<2.75	0.120 (0.27)	1843.08		
2.75-<3.00	0.034 (0.33)	1843.08		

SD standard deviation, SA sensitivity analysis, CSR clinical study report, HAQ health assessment questionnaire, MTX methotrexate, TNFi tumour necrosis factor alpha inhibitor

[a] Discontinuation due to reasons other than lack of efficacy
[b] 12-mth data for HAQ change were used as an estimate for 6-mth value for TNFi. As SD unavailable, assumption was made to use the same ratio of mean and SD as that observed in the ATTAIN trial
[c] Rates are adjusted for placebo (MTX) response
[d] Patients are assumed to remain on MTX through the analysis
[e] Update of study by Barbieri et al [48]

5.2.1 Population

The manufacturer states that the modelled population is reflective of the population in (i) the licensed indication and (ii) their statement of the decision problem. However, the primary source of evidence used in the economic model is from the ATTAIN trial. It is noted that 24%

of the patients in the abatacept arm of the ATTAIN trial received abatacept outside of its licensed indication as they did not receive abatacept + MTX.

The characteristics of an average biologic-treated RA patient used in the base-case analysis are detailed in Table 5-2. In the base-case, all of the patients in the model are assumed to be female.

Table 5-2: Characteristics of average RA patients used in the model

	Value	Source
Age	53	029 (ATTAIN) trial
Sex	Female	Majority of RA patients are female. An analysis for male patients is provided in sensitivity analysis
Baseline HAQ	1.82	029 (ATTAIN) trial. Baseline HAQ of 2.05 was reported by BSR which is provided in sensitivity analysis
Weight	70kg	Mean weight of a female RA patient in the UK (GPRD)[42]

The manufacturer performed an additional analysis for males to consider the generalisability of the cost-effectiveness results beyond the base-case.

5.2.2 Perspective and time horizon

An NHS perspective is adopted, in line with current NICE guidance.[41] In the base-case, the model is run for 20 years in order to reflect the health benefits and costs of RA over a long period. In the sensitivity analysis (SA), the model is run for five years.

5.2.3 Comparator

In their statement of the decision problem, the manufacturer suggests that abatacept can be compared to three different comparators.

In the base-case, abatacept + MTX is compared to MTX after failure of an initial TNFi.

In the additional analysis, the manufacturer compares abatacept to a cycled TNFi after failure of an initial TNFi. It is inappropriate to combine the three TNFi as a single 'cycled' drug as there are differences between these drugs (e.g. costs and duration of treatment effect). Also, as acknowledged by the manufacturer, a second TNFi is not currently recommended by NICE. The manufacturer repeatedly advises that this comparison should be considered as speculative due to data constraints.

The manufacturer also identified rituximab as a potential comparator. However, as the optimal strategy and timing of rituximab are currently unclear, the manufacturer did not conduct a full cost-effectiveness analysis of abatacept versus rituximab.

5.2.4 Treatment effectiveness within the submission

The primary measures of clinical efficacy used in the ATTAIN trial are ACR20 response rates and a clinically important improvement (≥0.3) in physical function as assessed by the HAQ between treatment groups at six-months. It is noted that efficacy data from all trial participants is used to furnish the model, yet 24% of the patients in the abatacept arm received abatacept outside of its licensed indication.

In the economic model, the measure of treatment efficacy used is the mean percentage change in HAQ score versus baseline at six months between groups. Clinical outcomes based on HAQ scores were extrapolated beyond the trial with an assumption of continued benefit whilst receiving therapy. In the model, treatment is stopped if a patient either does not respond initially (measured as an improvement in HAQ score at six months of ≥0.3 versus baseline), loses efficacy or experiences significant adverse events; the last two categories are accounted for in the annual treatment discontinuation rate.

In the base-case, HAQ scores at baseline and six months as well as AEs were derived from the ATTAIN trial. In the additional analysis, HAQ scores for the TNFi were obtained from the BSRBR submission[25] to NICE and were measured at 12 months instead of six months. The ERG notes that a MTC of REFLEX and ATTAIN trials was conducted but was not used to inform any of the cost-effectiveness analyses discussed in the submission.

5.2.5 Health related quality of life

The model links HAQ scores (intermediate outcomes) to final outcomes (QALYs). Patients are assigned EQ-5D health-state utilities according to their estimated HAQ scores, on a 6-monthly basis. In the base-case, six-month HAQ scores from the ATTAIN trial were used. In the additional analysis, 12-month HAQ scores were used as reported in the BSRBR[25] submission to NICE.

Estimates of health state utilities by HAQ score interval were derived from analyses of US data from the National Databank for Rheumatic Diseases (NDRD).[43] The estimated health state utilities by HAQ score interval are presented in Table 6.8 in the MS (p.109). Sensitivity analysis was conducted using the BRAM/Hurst utility estimate.[44]

The manufacturer acknowledges that their approach deviates from the reference case analysis requirement for societal values recommended by NICE as the NDRD[43] is made up of patient generated utility values.

5.2.6 Resources and costs

In the model, resources are split into three different cost categories: drug acquisition costs, administration costs and other medical costs. Unit costs and sources were stated for all of these costs (Table 6.9, p.111, MS). Medical costs other than treatment costs were linked to HAQ scores. UK data on average 6-monthly use of hospitalization, outpatient visits and joint replacement according to HAQ scores were taken from Barbieri et al,[45] who derived them from the Norfolk Arthritis Register (NOAR) Database. The unit costs applied to resource use were taken from NHS reference costs 2005/6.[46]

Patients in the model may suffer AEs, some of which may be serious. However, there are no costs associated with having an AE included in the economic model.

In the additional analysis, the manufacturer sets the price of the cycled TNFi equal to a weighted average of the three common TNFi based on published data.[17]

The manufacturer provides a detailed list of resource use assumptions in Section 6.2.9.9 of the MS (p.113).

5.2.7 Discounting

Health benefits and costs were discounted at 3.5% in line with current NICE guidance.[41]

5.2.8 Results included in manufacturer's submission

The manufacturer provides aggregate and disaggregate results for the base-case (abatacept + MTX versus MTX) and the additional analysis (abatacept + MTX versus TNFi).

Table 5-3: Cost-effectiveness results for abatacept + MTX versus MTX

Treatment	Abatacept + MTX	MTX	Difference
Total costs (£, 2006)	84,679.11	44,307.25	40,371.86
Total QALYs	4.7501	3.1604	1.5897
ICER	£25,395/QALY		

Table 5-4: Cost-effectiveness results for abatacept + MTX versus cycled TNFi

Treatment	Abatacept + MTX	TNFi	Difference
Total costs (£, 2006)	84,679.11	59,417.79	25,261.32
Total QALYs	4.7501	3.6337	1.1164
ICER	£22,628/QALY		

Subgroup analyses were conducted for males. In the abatacept + MTX versus MTX comparison, the ICER is £26,160 per QALY. In the abatacept + MTX versus TNFi comparison, the ICER is £ 23,155 per QALY.

5.2.9 Sensitivity analyses

Univariate SA and probabilistic sensitivity analysis (PSA) were conducted by the manufacturer. Univariate SA was performed on a range of key parameters and the results are presented in Table 6.19 and Table 6.20 in the MS (p.123-124). In the base-case and additional analyses, the cost-effectiveness results appear to be most sensitive to the following parameters: time horizon, discount rate, annual treatment cost of abatacept and assumption on rebound following treatment discontinuation. In addition, the cost-effectiveness results in the abatacept + MTX versus MTX comparison appear to be sensitive to the annual rate of HAQ progression on MTX.

For the PSA, scatter plots and cost-effectiveness acceptability curves were calculated and are shown in Figure 5-2 and Figure 5-3.

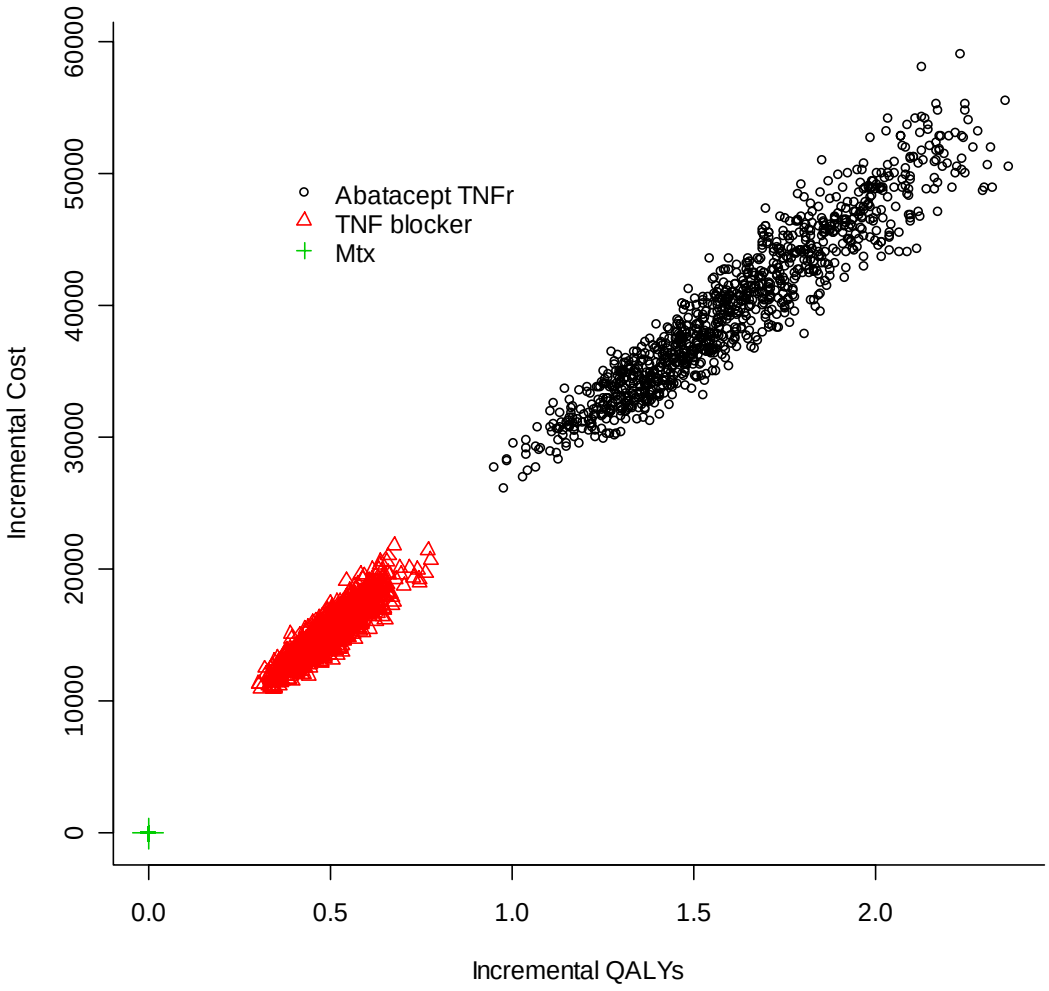


Figure 5-2: Scatter plot (1,000 simulations, 1,000 patients)

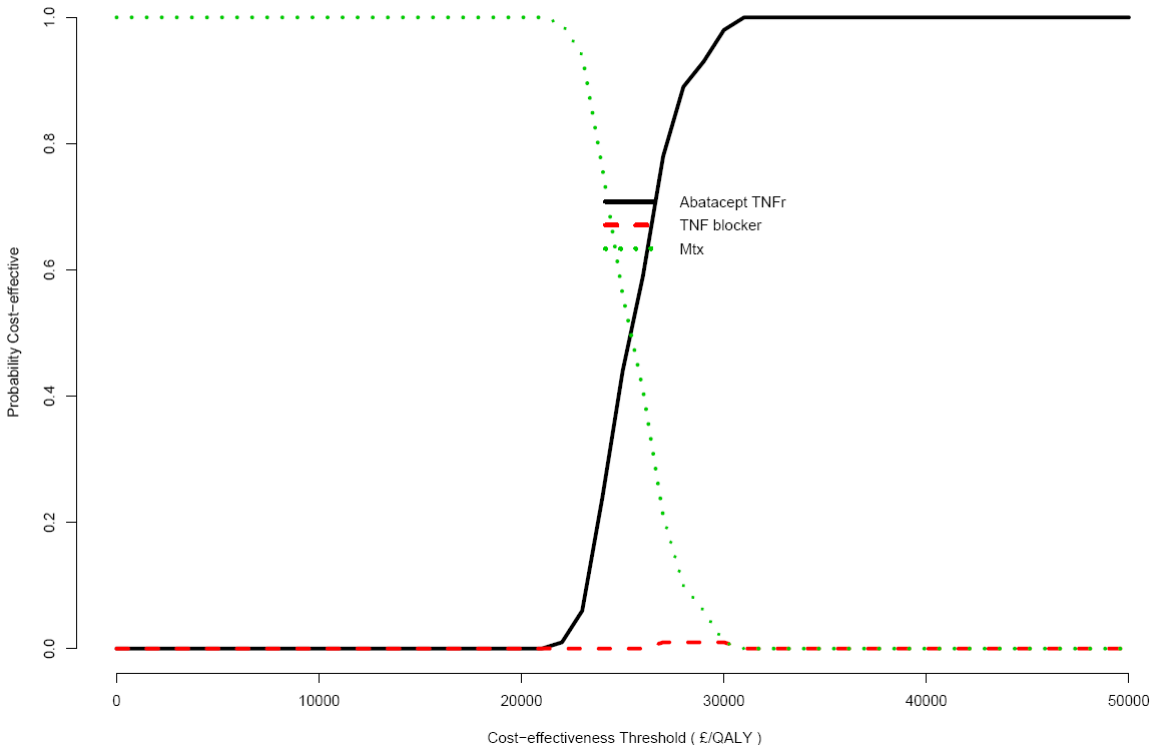


Figure 5-3: Cost-effectiveness acceptability curves

5.2.10 Model validation

The ERG cannot comment definitively on the reliability or validity of the R model. However, having examined the Excel-based model in detail, the ERG is confident that all of the essential features of the R-based model are in fact present. The MS also states that a number of steps were carried out to check and validate results from the model including: advice from clinical, economic and modelling experts, SR of published economic evaluations and a review of the model by an independent statistician and modeller.

5.3 Summary of cost-effectiveness and comment on validity of results presented with reference to methodology used

The ERG has undertaken a careful examination of the submitted Excel-based model (confirmed where possible by examination of the R language program code). In addition, the ERG has been able to make use of the extensive data in the CSR, made available in electronic format by the manufacturer, to test some of the clinical assumptions employed within the economic model.

Several logic errors, parameter value amendments and uncertain structural assumptions were identified in the economic model.

5.3.1 Base-case: abatacept + MTX versus MTX

The impact of each ERG modification on the cost-effectiveness of abatacept in the base-case is presented in Table 5-12. The individual ICER estimates vary from £25,072 per QALY gained to £30,479 per QALY gained. The cumulative effect of these changes is shown in Table 5-13 and the ICER then increases to £47,503 per QALY gained.

In addition, the ERG has identified one specific issue of major importance concerning assumptions about the progression of HAQ disability scores over time, and this is discussed in some detail. In summary, use of ERG estimates of HAQ progression rates instead of those described in the MS means that the ICER may increase further to £72,865 per QALY gained.

Due to limitations of time, the narrative results given in this section relate solely to the submitted base-case, which compares combination treatment using abatacept + MTX with MTX monotherapy. However, since almost all of the ERG modifications would also apply to the alternative scenario (involving a second TNFi), it can be assumed that their impact on cost-effectiveness estimates will be similar.

5.3.2 Additional analysis: abatacept + MTX versus cycled TNFi

The impact of each ERG modification in the additional analysis is displayed in Table 5-15. The ICER estimates for individual factors vary from £22,314 per QALY gained to £26,744 per QALY gained. The ERG had identified that choice of treatment effect for cycled TNFi (adjusted or unadjusted) has a significant impact on the magnitude of the ICER. The combined effects of all the identified amendments result in a revised ICER of £50,222 per QALY gained. If the ERG estimated HAQ progression rates are also applied the ICER increases still further to £67,459 per QALY gained.

A full discussion of the most influential issues in the abatacept + MTX versus cycled TNFi is presented in Section 5.6.

5.4 Critique of approach used

5.4.1 Model logic errors

Discounting

The Excel-based model calculates discounted costs and outcomes by applying discount rates to each half-year period, starting from the first treatment period. This is contrary to normal conventions which require that in the first year costs and outcomes are not discounted, and that all costs incurred or outcomes accumulated during any subsequent year are assigned the

same discounting factor (rather than discounting by different amounts for within-year sub-periods). This error appears to apply also to the R-based version of the model.

Correcting the discounting error has very little effect on the resulting ICER, since both incremental costs and QALY gains are closely linked to the duration of treatment with abatacept, and in the base-case are both discounted using the same appropriate discount rate (3.5%). The ICER with this correction is reduced marginally from £25,473 to £25,446 per QALY gained.

Sampling treatment duration

The submitted Excel-based model uses random sampling to assign an estimated duration of abatacept treatment to each patient. It is assumed that the annual rate of treatment discontinuation is constant for all years, and therefore requires random draws from an exponential survival distribution. The formula used in the model to convert a simple random number between 0 and 1 (uniform distribution) to its equivalent exponential survival distribution is incorrect, leading to a systematic over-estimation of mean survival time (by 0.55 years for the submitted baseline model value of 8.2% loss per year). The error also appears to apply to the R-based version of the model.

Correcting the sampling error has very little effect on the resulting ICER, since both incremental costs and QALY gains are closely linked to the duration of treatment with abatacept. The ICER rises slightly when this correction is applied to £25,530 per QALY gained.

Methotrexate costs in abatacept treatment

The calculation of treatment costs in the Excel-based model omits the costs of MTX acquisition and monitoring from the estimates for the abatacept treatment arm. This understates the annual cost of abatacept therapy by over £600 per patient-year, and therefore biases the analysis in favour of abatacept. Correction of this error alone leads to the estimated ICER increasing to £27,194 per QALY gained.

Half-cycle correction

The MS states that no half-cycle correction was applied to the first period so as not to understate initial costs. It does not indicate whether half-cycle corrections were used for later periods. Examination of the Excel-based model reveals that no half-cycle corrections were used (the ERG has not been able to verify if this is the same in the R-based model).

In amending the Excel-based model to correct for this omission, the ERG has assumed that formal assessment of treatment efficacy occurs at six-monthly reviews, but that failure of effect can occur at any time during the preceding six months. This means that treatment-related costs should be calculated for the whole of each period, but that efficacy is lost after an average of three months during the last period of treatment. Thus no half-cycle correction is required for treatment costs, but half-cycle corrections should be used for HAQ estimates. Since other costs as well as utility estimates are determined by the HAQ value, it is not necessary to use a separate half-cycle correction for these other variables.

Applying a half-cycle correction to the submitted model results in slight reductions in incremental life expectancy and QALYs, but a small increase in incremental costs, so that the ICER increases to £26,177 per QALY gained.

5.4.2 Parameter value adjustments

Abatacept discontinuation rate

The authors of the model make an important assumption that discontinuation of abatacept treatment for lack of efficacy only occurs during the initial six-month period. Thereafter, they consider that response is sustained indefinitely and patients only cease treatment due to AEs, reactions or other non-treatment related factors. They estimate an annual rate of 8.2% for these non-efficacy effects, based on clinical trial experience in the CIC removed. However, according to the CSR, 8.2% is the rate in just CIC removed so that the correct figure to apply should be approximately CIC removed that used in the MS. However, there is also evidence from the detailed appendix tables in the open-label extension for the ATTAIN trial that many of the subsequent discontinuations are assigned to 'lack of efficacy', thus contradicting the basis for the modeller's assumption.

A more reliable estimate for the true overall long-term withdrawal rate can be based on an analysis reported in the CSR for the ATTAIN trial open-label extension:

> "Among HAQ responders at Day 169 who entered the open-label period, CIC removed at Day 365 CIC removed and at Day 533 CIC removed (LT Table 10.2B)."

This suggests that CIC removed of patients who initially responded to abatacept at six months failed to retain that response one year later. This direct observation is to be preferred over the CIC removed figure (8.2%) which relies on assuming that only a minority effect, observed during the CIC removed, will be continued indefinitely.

Substituting this value for the long-term discontinuation rate in the Excel-based model produces important reductions in expected survival and utility gains, as well as in treatment

costs due to treatment with abatacept. The net effect is to increase the ICER by about £4,500 per QALY gained (to £30,038).

HAQ mortality multiplier

It is widely accepted that the HAQ is related to increased mortality risk, and this is commonly incorporated into models by means of a 'mortality multiplier' applied to the normal population mortality age/sex risk for each one point increase in HAQ score. Various values have been employed by previous modellers. In the MS (Appendix E) six estimates ranging from 1.3 to 2.73 observed in published models are averaged as the basis for using 1.8 as the multiplier in the submitted model. The averaging of risk ratios is technically invalid so that a parameter value of 1.8 lacks inherent validity. By contrast the authors of the BRAM model used 1.33, based on a 35-year longitudinal US study.[11] It appears that the value used in the submitted model is unduly influenced by a recent Finnish study[47] which appears to be a statistical outlier.

The SA presented in the submission suggests that the ICER is not very sensitive to changes in this parameter. The ERG has chosen to substitute an intermediate value of 1.5 in place of 1.8. This results in extended survival, increased costs and improved utility gains, but with only a small reduction in the estimated ICER (to £25,072).

NSAID use

In the MS, it is assumed that:

> "NSAID usage is equal in all patient groups and costs, benefits and side effects are therefore analytically ignorable." (p.103, MS)

This argument is only valid if there is no possibility of differences in patient survival arising as a result of the treatment given. Because mortality rates in the model are modified by HAQ scores - the primary treatment benefit claimed for abatacept - it is inevitable that survival gains will be generated by the model. Omitting consideration of the costs of widespread NSAID and corticosteroid use (in about CIC removed of patients respectively in the ATTAIN trial) will lead to an understatement of the cost consequences following treatment with abatacept and therefore may bias the analysis in favour of abatacept. To estimate the annual average cost of these additional medications, we have used the proportions of ATTAIN patients recorded as receiving any of the following drugs:

- NSAIDs (and one common opioid) (CIC removed of patients)

- treatment/prophylaxis for peptic ulcer related to regular NSAID use (CIC removed of patients)

- corticosteroids (CIC removed of patients)

- treatment/prophylaxis for bone disease related to regular corticosteroid use (CIC removed of patients).

A total additional cost of £261.33 per patient-year was calculated using British National Formulary 52[48] prices and recommended doses. Only the main agents cited were included in this estimate, so that the overall estimated cost per patient is deliberately conservative. The easiest way to incorporate this additional feature into the model is to increase the cost of regular MTX medication by this amount. Since this change only adds to the incremental cost during the few months of additional life expectancy, its impact on the cost-effectiveness results is very small (increasing the ICER to £25,534 per QALY gained)..

Treatment, administration and monitoring costs

The annual acquisition cost of abatacept is estimated very simply on the assumption that the 'average' patient will use three vials per infusion, rather than the more accurate estimate of 2.85 vials based on the GPRD[42] weight distribution for UK patients suffering from RA; this is therefore a conservative estimate by the manufacturer. In addition, it is wrongly assumed that two additional loading doses are required in the first period when the product licence requires only a single extra dose after two weeks. Once again this serves to inflate both the acquisition cost and the administration cost (one extra infusion) of abatacept. To summarise, the manufacturer overestimates the annual acquisition and administration costs of abatacept.

No attempt is made by the manufacturer to distinguish treatment costs between males and females, though the different distributions of body weight for males and females indicates that treatment costs will vary considerably.

The model uses the unit cost of an out-patient visit (£133) for each administration, rather than the day case cost generally charged in NHS hospitals of £222 (TRDNA RDH 98).[49] No attempt is made to include monitoring costs related to the use of MTX as indicated in clinical guidelines[50], which apply to both evaluation arms at the rate of six out-patient visits per year costing £124 per visit (TOPS FUA 410F).[49]

Taking all of these factors into account, the ERG has re-estimated the cost of abatacept + MTX and MTX only therapy per patient-year as shown in Table 5-5.

Table 5-5: ERG amendments to treatment costs

		BMS	ERG		
		Submitted model	**Males**	**Females**	**Overall**
Abatacept + MTX	1st period	£7,113	£7,836	£7,291	£7,450
	Annually	£11,558	£13,687	£12,743	£13,018
MTX only	Annually	£662	£786	£786	£786

The incremental cost of abatacept + MTX compared to MTX is increased by about 8% when the ERG overall treatment costs are used. These revised costs are fully implemented together with gender reconciliation costs as described below (see Section 5.4.3).

Hospital disease-related costs

The estimation of disease-related hospital costs (in-patient and out-patient) is based on a paper by Barbieri et al,[45] which quotes summary results from the NOAR database relating information on annual resource use to HAQ scores over a five year period for a cohort of early arthritis patients. There are serious questions about the relevance of these data to the patients in the ATTAIN trial who had a mean duration of disease of about 12 years, and generally greater disability. However, in the absence of any alternative source of UK resource evidence, there appears to be no more credible basis for estimating these important costs.

The ERG identified several problems with the use of NOAR information:

- the Barbieri et al[45] paper does not provide full details of the results from the database (as available in a research report prepared by Wiles et al[51])

- the use of in-patient bed days as well as NHS Reference Costs for surgical episodes for joint replacement involves double counting of hotel costs

- an average surgery cost is calculated using the full range of all replacement Health Related Groups weighted by the national episode volumes, instead of the arthritis specific case-mix recorded by NOAR[51]

- the calculation of this weighted average includes an important transcription error leading to a substantial over-estimate of surgery costs.

In view of these problems, Table 5-6 shows the ERG re-estimated disease-related hospital costs, based on the full NOAR[51] results, and using the NOAR[51] surgical case-mix, adjusting for the double-counting of hotel costs.

Table 5-6: ERG amendments to model disease-related hospital costs per patient year

HAQ scores	0	0.001-0.5	0.5-1.0	1.0-1.5	1.5-2.0	2.0-2.5	2.5-3.0
BMS original estimate	£296	£620	£620	£1,314	£1,314	£3,686	£3,686
ERG revised estimate	£187	£187	£190	£383	£507	£975	£1,879

The revised costs lead to substantial reductions in estimated costs in both arms, but rather less for abatacept than for the MTX arm. As a result of this single amendment, the estimated ICER increases to £27,695 per QALY gained.

Estimated benefit of abatacept therapy on HAQ scores

There are two concerns about the manner in which abatacept treatment effects are represented in the model:

1) A single proportional multiplier is used to reduce all baseline HAQ scores for patients responding to abatacept. This is a strong assumption for which no explanation is offered in the submission: it implies that efficacy increases in direct proportion to the baseline HAQ score across the full range of experience. To test the validity of this formulation, the ERG used the individual patient data (IPD) included in the appendices to the ATTAIN CSR to plot the change in HAQ between baseline and day 169 against the baseline HAQ (Figure 5-4). It can be seen that the fitted trend line CIC removed by the manufacturer.

CIC removed

Figure 5-4: Change in HAQ score by day 169 by baseline HAQ score for all ATTAIN patients continuing on assigned therapy to the end of the trial

2) The method of estimating the size of this treatment effect appears to have involved calculating a percentage change in score for each case, and then calculating the arithmetic average of these values. If this is a correct interpretation of the statistics included in Appendix E of the MS, then it constitutes an erroneous approach, potentially liable to substantial and unpredictable bias (this is seen clearly if we consider that a reduction in HAQ from 2.5 to 1.25 is given the same weight as an improvement from 0.5 to 0.25). In general, it is mathematically meaningless to average proportions. The correct approach in this case is to apply linear regression to the IPD to obtain separate linear models for patients treated with abatacept + MTX and placebo + MTX: the efficacy gain is then given by the difference between the two estimated slopes.

Table 5-7 shows the results of carrying out this procedure on the ATTAIN IPD for all patients, and for males and females separately (since exploratory regression analysis CIC

removed). Compared to the model parameters, the IPD estimates suggest that the general effectiveness of abatacept has been CIC removed. There also appears to be CIC removed, though this would need confirmation as the sample sizes are much smaller for males.

Replacing the effectiveness parameter in the submitted model with the regression estimate leads to a larger proportionate reduction in the incremental QALY gain than the corresponding reduction in incremental costs, so that the ICER increases to £27,517 per QALY gained.

Table 5-7: Percentage change in HAQ score from baseline to day 169 - submitted model and ERG analysis of ATTAIN IPD (difference between slopes for abatacept + MTX and placebo + MTX regression lines)

		Mean change	Standard error	Standard deviation
Submitted model	All patients	CIC	CIC	CIC
ATTAIN IPD regression estimates	All patients	CIC	CIC	CIC
	Males	CIC	CIC	CIC
	Females	CIC	CIC	CIC

5.4.3 Structural adjustments

Gender-specific estimation

The base-case model results submitted by BMS relate to a female cohort using standard population age-specific mortality rates. Though this is the larger gender group, it is normal to present a base-case for the whole treated group, and therefore the ERG has adopted an overall mixed-gender cohort for the base-case analysis. This requires the calculation of a set of overall mortality rates, with the gender mix varying over time. A minor modification of the model has been introduced for this purpose.

In the submitted model, the female population is combined with single mixed-gender treatment cost estimates, and a single mixed-gender effectiveness estimate. The model has been modified by the ERG with gender-specific cost and effectiveness parameter estimates (Table 5-5 and Table 5-7), to allow fully consistent gender-specific or overall population calculation of cost-effectiveness results.

The effect of changing from the submitted inconsistent female analysis to the ERG base-case (consistent mixed-gender cohort) is to reduce the incremental benefit while increasing the incremental cost, so that the ICER rises to £28,744 per QALY gained.

Choice of utility model

The BMS model uses a relationship between HAQ scores and utility derived from analysis of data relating to patients recorded in the US NDRD.[43] Utility estimates were estimated using the EQ-5D instrument calibrated for US population preferences, a SA was undertaken using the Hurst[52] model. There are two other HAQ-utility models that have been used in economic evaluations of RA treatments (Hawthorne[53] and Bansback[54]), which should also be considered. All four options are plotted in Figure 5-5.

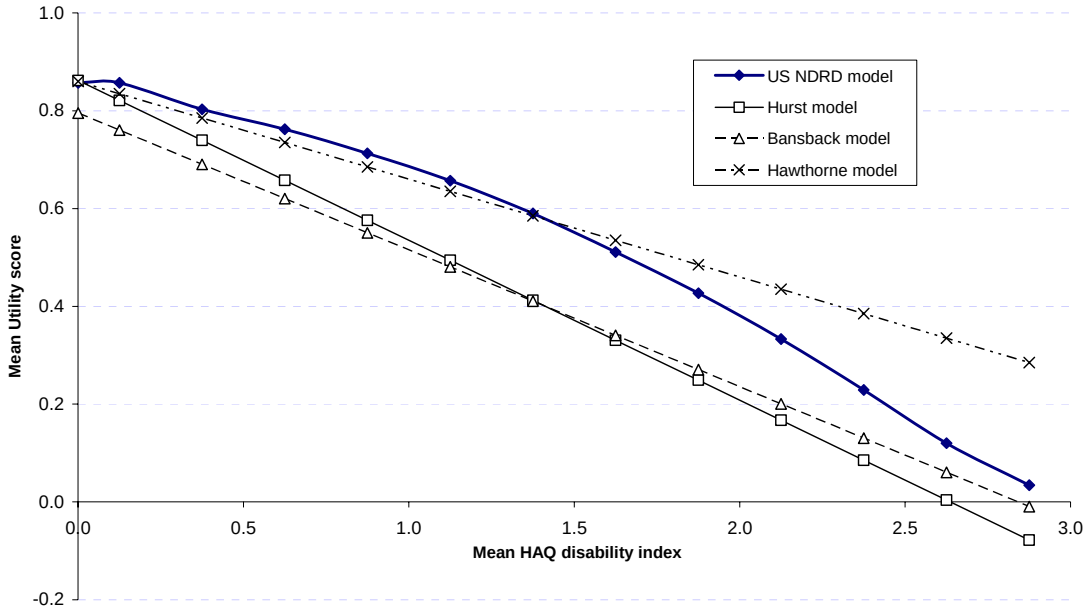


Figure 5-5: Alternative models used to predict utility from HAQ disability scores

The most important characteristic of such models is the gradient of the curve in the region occupied most frequently by most of the population being considered. In the base-case (with an initial HAQ score of 1.82) this is the part of the curves between about HAQ = 1.5 and HAQ = 3.0, and it is noticeable that the NDRD[43] model exhibits the strongest downward gradient of all the available models in this region. The MS includes a SA using the Hurst[52] model (with the next steepest gradient). However, it is probably more appropriate to consider the Bansback[54] model as an alternative, since it is based on a much larger patient sample (about 2,000 cases) than either the Hurst[52] or Hawthorne[53] models. Since the NDRD[43] model appears to be substantially different from the other studies, relates to a non-UK population and has not yet appeared in a peer-reviewed publication, the ERG prefers to adopt the Bansback[54] model for its base-case scenario.

The Excel-based model has been modified to allow utility estimation by any of these four models. Using the Bansback[54] model substantially reduces the estimated incremental utility gain, leading to an increase in the estimated ICER to £30,479 per QALY gained.

5.4.4 Progression of functional disability

Importance of HAQ to model logic

In the submitted model a direct modification is applied to HAQ scores to represent the impact of response to treatment. This is assumed to be operative in each period during which the treatment continues to be effective. The proportion of patients considered to have achieved a response (defined as a reduction in HAQ from baseline of at least 0.3) is estimated from trial data. The logical structure of the submitted model mediates the impact of different treatments through the estimation of changes in HAQ scores. These changes impact on mortality/survival, patient utility (quality of life) and direct medical costs (both treatment-related and disease-related), as illustrated in Figure 5-6. Thus, we can expect that assumptions and parameter values governing HAQ scores will be highly influential on model results.

This is borne out by the one-way SA included in the MS which shows that the cost-effectiveness ratio is most sensitive to assumptions about the rate of progression in the comparator arm, which can lead to MTX dominating abatacept + MTX (i.e. being both less costly and more effective).

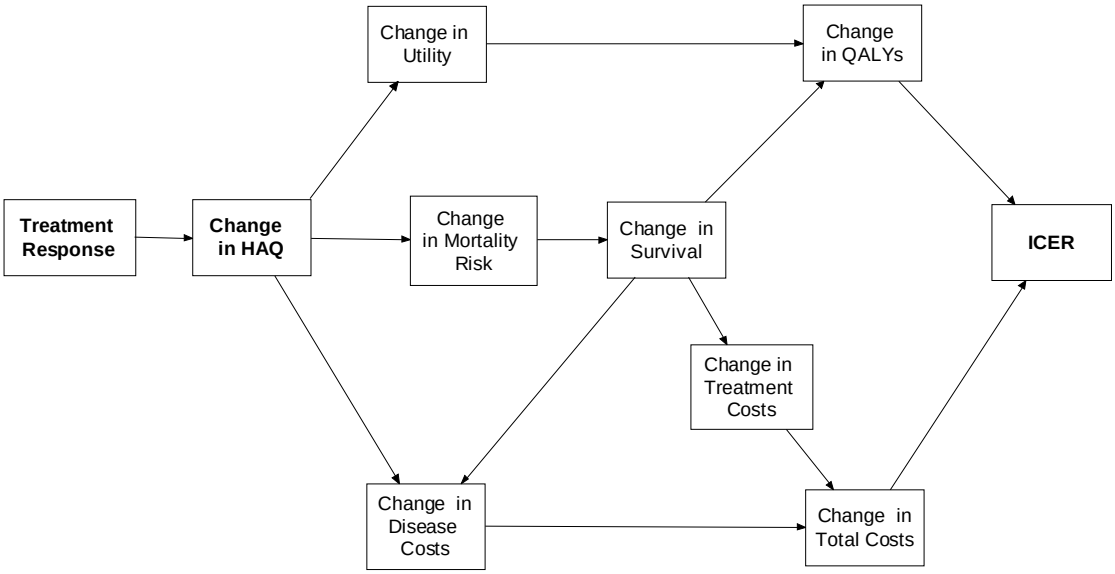


Figure 5-6: Effects of changes in HAQ in MS

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 0.015 points per year whilst undergoing any active treatment with a DMARD, and a greater rate of 0.06 points per year when all active treatment options have been exhausted and the patient is deemed to receive only palliative therapies (in this case including MTX). 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 abatacept. Figure 5-7 presents a 2-way SA from the submitted 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 year 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 Excel-based model version.

The manufacturer's base-case scenario assumes that progression on palliative care is 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 abatacept could be considered cost effective under any assumptions. The threshold ratio for cost effectiveness (if this were considered to be about £30,000 per QALY gained) using the manufacturer's assumption of 0.015 increase in HAQ per year whilst on active treatment is

about 1.75, corresponding to a long-term progression rate of at least 0.026 per year. 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 5-7 is the impact of HAQ progression on incremental utility, since incremental costs are not very sensitive 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 5-7).

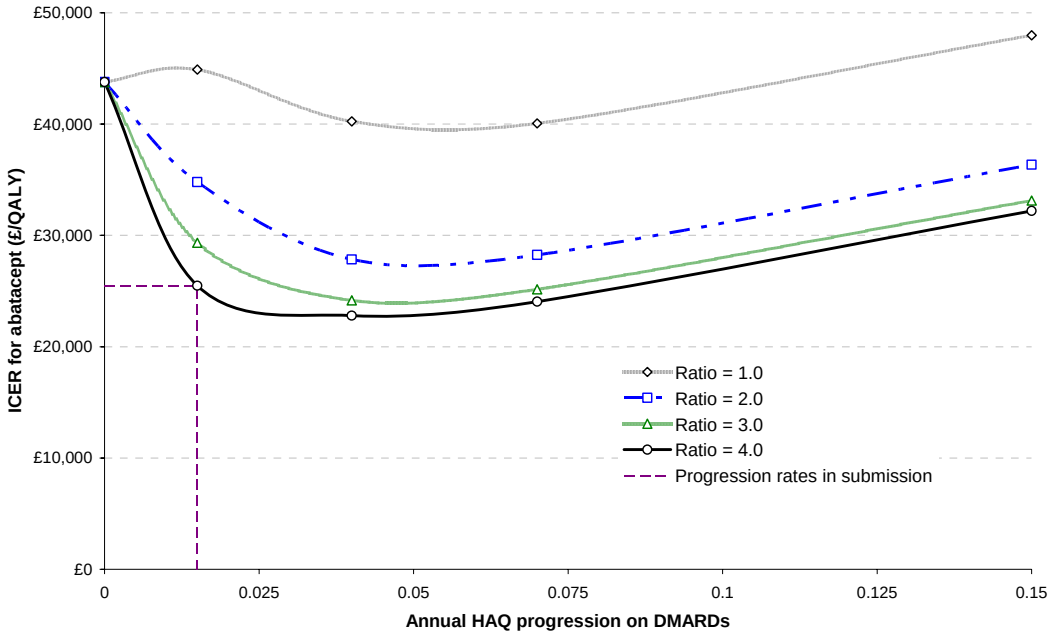


Figure 5-7: Sensitivity of the estimated ICER for abatacept to different values of the progression rates of HAQ scores (unmodified Excel-based model)

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) gives rise to anomalies in the representation of treatment effects. The submitted model randomly assigns a decrement in HAQ score to each patient which can (and from time to time does) exceed the upper limit or fall below the lower limit of the HAQ scale. Conversely, patients with high initial HAQ scores 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 the method of calculating HAQ scores in the model is not consistent with real-life, 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. Variation in baseline scores is not incorporated within the model, and is only considered via a one-sided sensitivity test. 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 essentially variable nature of the entities being measured (disease variability). The extent of such changes is clearly seen in Figure 5-8 (reproduced from Scott[55]). Although the mean score appears to increase slowly and steadily over time, the extent of individual fluctuations 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.

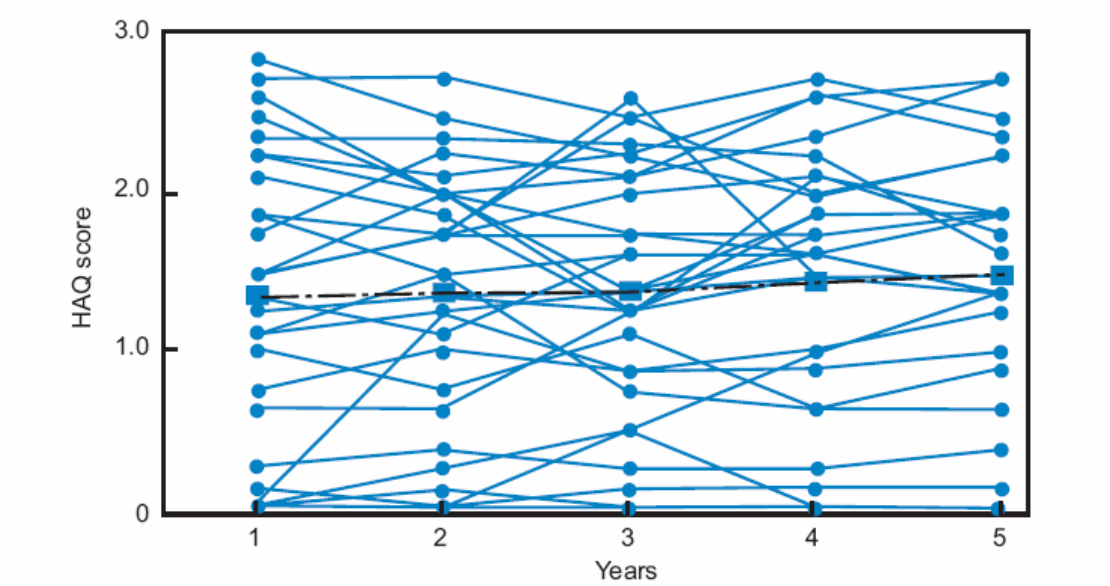


Figure 2. Individual (solid lines) and mean (broken lines) changes in Health Assessment Questionnaire (HAQ) score in 30 rheumatoid arthritis patients followed for 4 years in a single unit.

Figure 5-8: Variations on HAQ scores for individual patients reproduced from Scott and Garrood[55]

Detailed analysis of the NOAR database reported by Wiles et al[51] is valuable in offering greater insight into how HAQ scores change over time. Table 5-8 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 5-8: Frequency distribution of annual changes in HAQ scores for 1246 early-stage patients followed for 5 years (Wiles et al)[51]

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

A similar analysis of changes in HAQ score over six months for placebo patients in the ATTAIN trial shows a similar pattern (albeit with a much smaller volume of observations). Table 5-9 indicates that CIC removed of patients with the most severe disability index scores at baseline showed CIC removed after six months. Of those in the next severity class (2.00-2.375), CIC removed. Thus, a similar pattern appears to apply to both early and established RA patients with active disease.

Table 5-9: Frequency distribution of changes in HAQ scores for ATTAIN patients

		HAQ band at 6 months					
		0.00 – 0.375	0.50 - 0.875	1.00 – 1.375	1.50 - 1.875	2.00 - 2.375	2.50 - 3.00
Baseline HAQ band	0.00 - 0.375	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*
	0.50 - 0.875	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*
	1.00 - 1.375	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*
	1.50 - 1.875	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*
	2.00 - 2.375	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*
	2.50 - 3.00	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*	*CIC*

The impact of this degree of variability can be gauged by repeatedly applying the transition rates reported from the NOAR[51] database to a specified cohort of patients to simulate trends over several years; Figure 5-9 shows the effects for the NOAR[51] early-stage RA cohort, and also for a more severely affected illustrative cohort with initial scores drawn from the ATTAIN trial population and a starting mean HAQ of 1.83. The estimated score for the NOAR[51] 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 five-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 1.83) shows a downward non-linear convergence to the same

steady-state value (which is wholly determined by the NOAR[51] probabilities). Clearly this is not realistic either, 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.

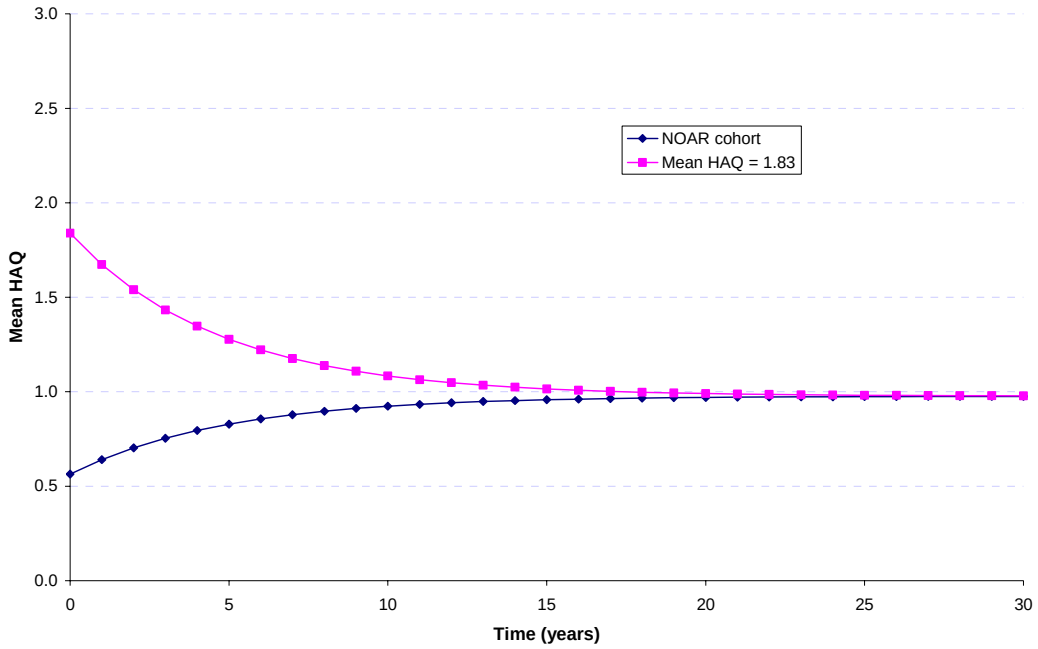


Figure 5-9: Illustration of non-linear trends in mean HAQ scores using NOAR[51] transition rates

In Figure 5-10 we show a scenario much closer to that implied by the submitted model - with an initial mean HAQ of 1.83 using the ATTAIN baseline score profile, 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;

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

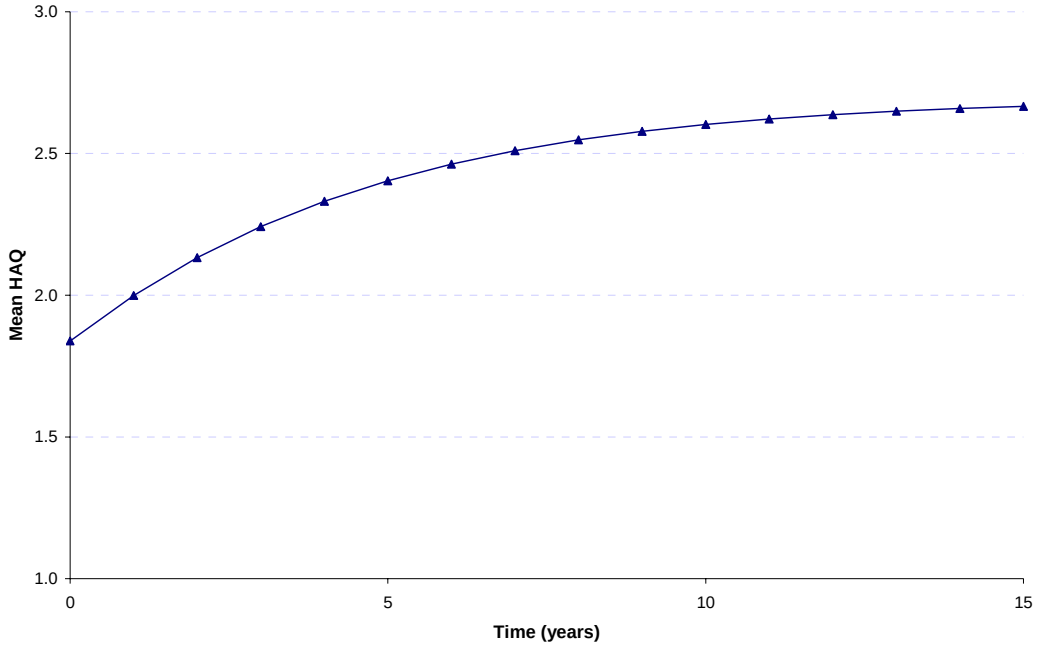


Figure 5-10: Illustration of projected trend in mean HAQ using transition probabilities weighted strongly toward progression

Evidence for HAQ progression rates

The MS cites a single source by Brennan et al[56] for the assumed long-term HAQ progression rates of 0.015 per annum when treated with biologics and 0.06 per annum on 'palliative therapy'. However, these parameter values rest on questionable assumptions and analyses as described in Scott and Garrood[55]and Genovese et al[57]:

- that in the absence of biologic therapy HAQ scores progress at a mean rate of 0.034 per year, obtained from Table 6 in Scott and Garrood's review paper[55] published in 2000

- that treatment with a biologic reduces HAQ progression to 44% of the non-biologic rate, based on an etanercept trial reported by Genovese et al[57] in 2002

- that, for patients no longer considered suitable for active therapy, HAQ progression occurs at a higher rate derived from the ERAS[58] study of early RA.

We examine each of these propositions in turn.

HAQ progression on non-biologic DMARDs: Table 6 in Scott and Garrood's review paper[55] mentions results from nine observational studies of different types and durations. The authors combined the 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 5-10 summarises our findings which differ in important respects from those of Scott and Garrood.[55]

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.[55] Moreover, the cross-sectional studies yielded an estimate twice the size of the longitudinal studies.

HAQ progression on biologic treatments: Brennan et al[56] cite a review paper by Scott[59] published in 2000 which considers joint damage and disability in RA. This paper demonstrates strong correlations between different radiologic measures of damage and disability; it does not attempt to quantify a direct relationship between changes in radiologic damage scores and increases in mean HAQ scores. This is principally because they recognise that HAQ scores are confounded by other factors:

> "In addition to the effects of joint damage, RA disability is also influenced by demographic factors (including age, sex, socio-economic and educational status, and income), measures of disease activity and inflammation (including pain, fatigue, ESR and CRP levels), and therapy with slow-acting and other anti-rheumatic drugs. These relationships can readily confound the link between disability and joint damage."

Notwithstanding these evident uncertainties, Brennan et al[56] derive a parameter value from the extension phase of a clinical trial of etanercept versus MTX in a population of patients with early RA (mean duration at baseline 12 months) reported by Genovese et al[57] in 2002. Only a minority of these patients had previously received any DMARD therapy (46% in the MTX arm and 40% in the etanercept arm). No mean values were reported for this study, but the mean Sharp radiologic score was found to have increased less in the etanercept arm over two years than in the MTX arm (+1.3 units versus +3.2 units). On this basis Brennan et al[56] assumed that a simple pro-rata estimate of long-term progression could be estimated as 44% (i.e. 1.3/3.2) of the non-biologic progression rate described by Scott and Garrood.[55] This procedure is inappropriate since both the source RCT related to a very different patient group, and it is most unlikely that a ratio of radiologic scores will translate directly into the same ratio of HAQ scores, given the extent of confounding involved. Indeed, it may be presumed that if such a significant and clinically important difference in HAQ had occurred at 2 years, it would have been included in the paper.

Table 5-10: Evidence of long-term HAQ progression in RA - studies used by Scott and Garrood[55] in estimating average annual progression rate

Study	Data period	Cases	Study Type	Rates quoted	Rates calculated	Comments
Wolfe[60]	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[61]	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[62]	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

Study	Data period	Cases	Study Type	Rates quoted	Rates calculated	Comments
Ward (1)[63]	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[64]	1984-89	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[65]	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[66]	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)[67]	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[68]	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

HAQ progression on palliative care: Tracing back the reference given in support of the submitted model's assumed long-term progression rate on palliative care (0.06 per year), we find that the source is the paper by Brennan et al[56] reporting results of a five year observational study of early RA patients presenting at nine UK rheumatology hospitals (the ERAS[58] study). In the original paper by Young[58] it is notable that results are not presented for mean HAQ for any time point nor for any sub-group of the cohort. The figure used in the MS submission (0.06 per annum) appears to have been obtained as a personal communication from Young and as such is included in the Brennan et al study,[56] and relates to 145 patients (20% of the cohort) who failed on at least two DMARDs during the study, and concerns increases over just two years.

The derivation of this parameter value appears to be somewhat tenuous, relying as it does on limited experience of an unreported subset of an observational cohort study. Even if these concerns are set aside, the application of the value in the model to patients receiving only palliative treatments is questionable. The ERAS paper[58] 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 five years, nor how many (if any) of those were included in the 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 six years at the end of the study) could reasonably be considered a suitable source for projecting the experience of patients with RA of duration 10-20 years or more.

A recent study[69] examined the effectiveness of MTX used in patients failing all other DMARDs, despite having previously failed on MTX therapy. They concluded that

> "Re-employment of MTX despite prior inefficacy, but not re-employment of other DMARDs, is an effective therapeutic option, especially in those patients in whom the methotrexate dose of the original dose was low."

This suggests that even if palliative care generally were to be associated with accelerated progression of HAQ scores, it is likely that MTX is an exception and may lead to rates closer to those on active DMARD therapies.

5.4.5 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 to be anticipated from the use of abatacept.

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 ERG's view is that HAQ progression would be best estimated by use of a non-linear trend line, consistent with a long-term stable maximum mean value a little below the scale maximum (approximately 2.75) to reflect the inherent variability in HAQ measurement. However, this cannot be implemented within the structure of the submitted model, which obliges us to consider only simple linear progression rates.

The analysis of observational studies cited[55] to support a progression rate of 0.034 per annum whilst on non-biologic DMARD 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 non-biologic DMARD progression rate of 0.012 per annum.**

The contention that progression on biologic treatments is only 44% of that for other DMARDs is based on a very generous and largely unsupported assumption of the direct proportionality of radiologic progression and progression in disability, which the cited authors do not claim. **The ERG considers that a more reasonable assumption might be that biologics reduce the progression rate to around 75% of the non-biologic rate, approximately 0.009 per annum.**

The ERAS[51] paper 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. No direct evidence has been provided to support the notion that a substantially different long-term progression rate should apply when all DMARD options are exhausted. In particular, there is published evidence[69] that repeated MTX treatment may have continued efficacy. **Therefore the ERG considers that only the non-biologic DMARD progression rate (0.012 per annum) is supportable from the evidence cited for use when all other DMARDs have failed, and that a modest sensitivity analysis (up to twice this estimate) would be an appropriate test of uncertainty.**

5.5 Additional comparison: abatacept + MTX versus cycled TNFi

In the MS, results are presented for an additional comparison in which it is assumed that patients may receive treatment with a second TNFi despite the current NICE guidelines.[21] This requires incorporating new parameter values into the manufacturer's economic model to represent the clinical effectiveness of a second TNFi on HAQ scores, as well as the appropriate treatment costs.

5.5.1 Treatment effectiveness

The model authors make use of information on the effects of using a second TNFi contained in an analysis of BSRBR data commissioned by NICE.[35] In particular, the unadjusted mean change in HAQ score (-0.15 equivalent to -7.32%) for patients who switched treatments shown in Table 1 of the BSRBR report[35] is used as the basis for populating the model. This ignores the main part of the report which describes a series of adjustments required to correct for imbalances likely to bias the size of the effect. The most appropriate HAQ effect for this purpose is found in the last column of Table 4 of the BSRBR report[35] which shows a value for the reduction in HAQ of 0.2146 (0.0823, 0.3470) after adjusting for all potential confounders.

In Appendix E of the MS, it is further assumed that the standard deviation of the TNFi effect will be proportionate to the size of that effect. This is erroneous, since in the ATTAIN trial the standard deviations of proportionate effects for both trial arms and for the difference between the arms are all of similar magnitude, despite very different effect sizes. The ERG therefore estimates that appropriate model parameters should be a mean effect size of -10.47% and standard deviation of 20.57% (based on the BSRBR report). Rerunning the additional comparison with these new effect values substituted, markedly improves clinical and economic outcomes for the comparator so that incremental QALYs fall by more than 40% of the previous estimate while incremental costs fall by 30%; as a consequence, the ICER increases from £22,601 to £26,744 per QALY gained (Table 5-15).

5.5.2 Treatment costs

The treatment costs for TNFi used in the submitted model do not include treatment with MTX which is generally required for patients receiving adalimumab, etanercept or infliximab, nor for regular monitoring as required for all patients receiving biologic therapies. Revised costs are shown in Table 5-11, separated by gender where dosing is by body weight to reflect gender-specific distributions of weight.

Table 5-11: ERG revised treatment costs

Treatment	Gender	First 6 months	Continuing annual cost
Abatacept + MTX	Males	£7,835.64	£13,686.59
	Females	£7,291.44	£12,743.31
	Overall	£7,450.18	£13,018.47
MTX	All	£765.40	£786.10
Etanercept + MTX	All	£5,429.12	£10,113.54
Adalimumab + MTX	All	£5,428.86	£10,113.01
Infliximab + MTX	Males	£7,027.47	£10,371.09
	Females	£6,377.92	£9,376.87
	Overall	£6,567.39	£9,666.88
Weighted average TNF □ inhibitor	Males	£6,149.40	£10,229.54
	Females	£5,856.67	£9,781.48
	Overall	£5,942.06	£9,912.18
Submitted TNFi cost	Overall	£5,687.61	£9,183.44

5.6 ERG revised base-case economic results: abatacept + MTX versus MTX

The consequences for the cost-effectiveness of abatacept + MTX versus MTX only of each of the logic, parameter value and structural changes to the submitted model described above (excluding HAQ progression rates) can be seen in Table 5-12. The most influential changes are the long-term withdrawal rate from abatacept therapy, and the choice of utility model.

The cumulative effects as each change is combined with preceding changes are displayed in Table 5-13, leading to the revised ERG base-case cost-effectiveness results in row 12. Although both incremental costs and benefits are substantially reduced, the larger reduction for outcomes leads to the ICER increasing from £25,473 to £47,503 per QALY gained.

If in addition the ERG-preferred HAQ progression rates are also employed (row 13), the incremental cost is quite stable but the QALY gain is reduced by about a third - thus the ICER for abatacept + MTX versus MTX increases to £72,865 per QALY gained. The effects of a wider SA of cost effectiveness in the ERG-modified model with respect of HAQ progression rates are displayed in Figure 5-11.

Table 5-12: Marginal effect of each of the identified corrections/amendments on cost-effectiveness results using the EXCEL-based model

		Abatacept + MTX				MTX				Incremental				ICER
		LYs	QALYs	Costs	Final HAQ	LYs	QALYs	Costs	Final HAQ	LYs	QALYs	Costs	Final HAQ	Cost /QALY
1	**Submitted base-case**	**16.254**	**4.742**	**£84,631**	**2.440**	**15.660**	**3.157**	**£44,270**	**2.730**	**0.594**	**1.584**	**£40,361**	**-0.289**	**£25,473**
2	Discounting correction	16.254	4.865	£86,794	2.440	15.660	3.239	£45,421	2.730	0.594	1.626	£41,373	-0.289	£25,446
3	Sampling treatment duration	16.240	4.708	£83,864	2.450	15.660	3.157	£44,270	2.730	0.580	1.551	£39,594	-0.279	£25,530
4	Abatacept discontinuation rate	15.871	3.776	£62,862	2.665	15.660	3.157	£44,270	2.730	0.211	0.619	£18,592	-0.065	£30,038
5	HAQ mortality multiplier	17.305	4.855	£87,793	2.504	16.984	3.258	£47,745	2.809	0.321	1.597	£40,049	-0.305	£25,072
6	Disease related hospital costs	16.254	4.742	£64,466	2.440	15.660	3.157	£20,583	2.730	0.594	1.584	£43,883	-0.289	£27,695
7	NSAIDs use	16.254	4.742	£87,785	2.440	15.660	3.157	£47,328	2.730	0.594	1.584	£40,458	-0.289	£25,534
8	Missing MTX costs	16.254	4.742	£87,359	2.440	15.660	3.157	£44,270	2.730	0.594	1.584	£43,089	-0.289	£27,194
9	Half-cycle correction	16.282	4.802	£84,140	2.443	15.692	3.243	£43,340	2.732	0.589	1.559	£40,801	-0.289	£26,177
10	Bansback utility	16.254	3.200	£84,631	2.440	15.660	1.876	£44,270	2.730	0.594	1.324	£40,361	-0.289	£30,479
11	Estimated benefit on HAQ scores	16.192	4.572	£83,197	2.466	15.660	3.157	£44,270	2.730	0.532	1.415	£38,927	-0.264	£27,517
12	Gender specific estimation	15.751	4.672	£89,450	2.408	15.093	3.109	£44,525	2.696	0.658	1.563	£44,925	-0.287	£28,744

Table 5-13: Cumulative effect of the identified corrections/amendments on cost-effectiveness results using the EXCEL-based model

		Abatacept + MTX				MTX				Incremental				ICER
		LYs	QALYs	Costs	Final HAQ	LYs	QALYs	Costs	Final HAQ	LYs	QALYs	Costs	Final HAQ	Cost /QALY
1	**Submitted base-case**	**16.254**	**4.742**	**£84,631**	**2.440**	**15.660**	**3.157**	**£44,270**	**2.730**	**0.594**	**1.584**	**£40,361**	**-0.289**	**£25,473**
+2	Discounting correction	16.254	4.865	£86,794	2.440	15.660	3.239	£45,421	2.730	0.594	1.626	£41,373	-0.289	£25,446
+3	Sampling treatment duration	16.240	4.830	£86,007	2.450	15.660	3.239	£45,421	2.730	0.580	1.591	£40,585	-0.279	£25,503
+4	Abatacept discontinuation rate	15.834	3.776	£62,319	2.678	15.660	3.239	£45,421	2.730	0.174	0.537	£16,898	-0.052	£31,456
+5	HAQ mortality multiplier	17.075	3.883	£65,682	2.753	16.984	3.342	£48,988	2.809	0.091	0.541	£16,694	-0.056	£30,849
+6	Disease related hospital costs	17.075	3.883	£40,637	2.753	16.984	3.342	£23,005	2.809	0.091	0.541	£17,631	-0.056	£32,582
+7	NSAIDs use	17.075	3.883	£44,006	2.753	16.984	3.342	£26,359	2.809	0.091	0.541	£17,647	-0.056	£32,611
+8	Missing MTX costs	17.075	3.883	£44,694	2.753	16.984	3.342	£26,359	2.809	0.091	0.541	£18,335	-0.056	£33,882
+9	Half-cycle correction	17.095	3.974	£44,374	2.755	17.001	3.435	£25,918	2.810	0.094	0.538	£18,456	-0.055	£34,278
+10	Bansback utility	17.095	2.473	£44,374	2.755	17.001	2.023	£25,918	2.810	0.094	0.449	£18,456	-0.055	£41,069
+11	Estimated benefit on HAQ scores	17.087	2.417	£43,871	2.758	17.001	2.023	£25,918	2.810	0.086	0.394	£17,953	-0.052	£45,554
+12	Gender specific estimation	**16.664**	**2.396**	**£45,727**	**2.731**	**16.567**	**2.005**	**£27,165**	**2.784**	**0.097**	**0.391**	**£18,562**	**-0.053**	**£47,503**
-	+ERG HAQ progression rates	**17.089**	**3.603**	**£40,203**	**2.008**	**17.040**	**3.342**	**£21,143**	**2.018**	**0.050**	**0.262**	**£19,060**	**-0.011**	**£72,865**

N.B. All results based on simulation of 10,000 patients using a common set of random numbers throughout

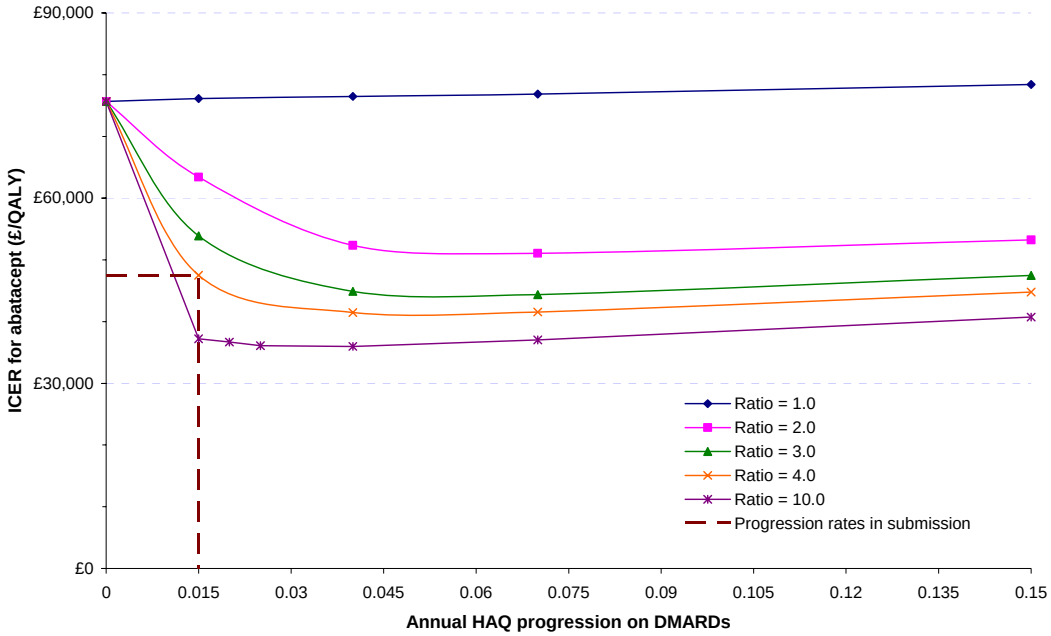


Figure 5-11: Sensitivity of the estimated ICER for abatacept to different values of the progression rates of HAQ scores (modified Excel-based model)

5.6.1 Sensitivity analysis: base-case comparison

The Excel-based version of the submitted model did not include program codes to allow PSA to be undertaken, and it was not feasible within the time available to attempt to implement all the ERG changes to the R-based model to carry out revised PSA calculations.

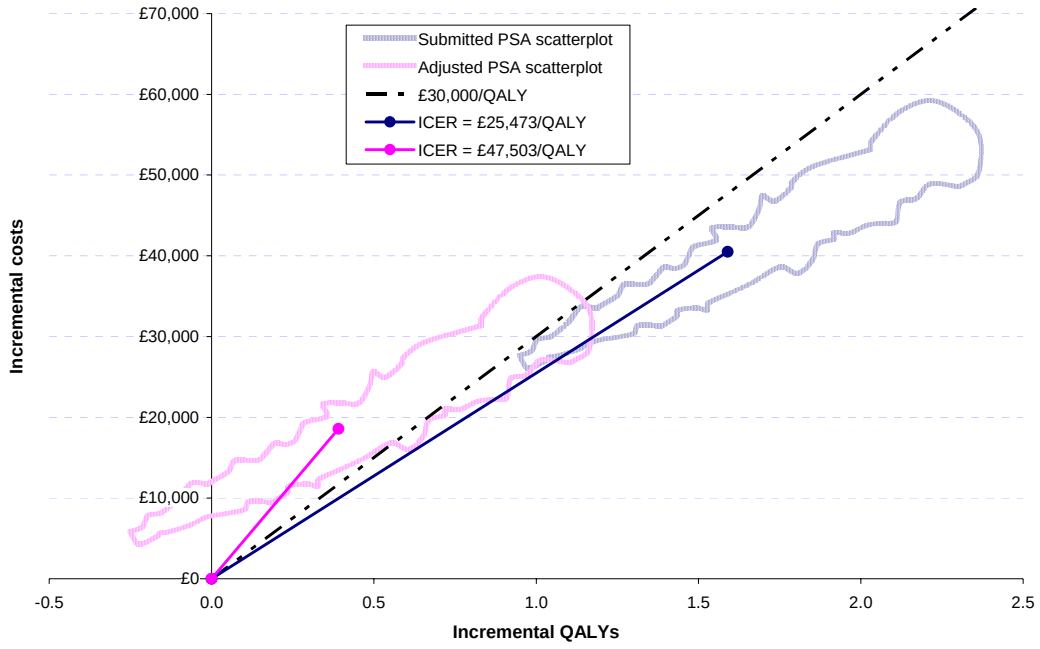
The ERG used the modified Excel-based model to explore sensitivity jointly for the combined effects of initial age, patient gender and initial HAQ score. The results are shown in Table 5-14: younger age, female gender, and more severe baseline HAQ are all associated with better cost-effectiveness ratios. However, in all cases the ICER remains outside the range normally considered cost-effective.

The PSA reported in the MS, using the R-based model version, included uncertainty estimates for most of the available model variables and assumptions. On the basis that in most cases these estimates would continue to be appropriate following the ERG amendments, an approximation to the effect of PSA using the amended model can be obtained simply by relocating the position of the original scatterplot to correspond to the revised mean values of incremental cost and utility for the base case comparison. This is illustrated in Figure 5-12; visual inspection then suggests that the probability of cost-effectiveness may be no more than 7% for a willingness-to-pay threshold of £30,000 per QALY gained.

Table 5-14: Three-way sensitivity analysis of results from the ERG-modified Excel-based model (including all identified changes except the ERG preferred progression rates) - ICER (£/QALY gained)

		Baseline HAQ		
Gender	**Baseline age**	1.50	1.83	2.15
Male	40	£53,878	£48,196	£45,639
	53	£56,234	£50,965	£47,803
	65	£61,913	£56,989	£52,648
Mixed	40	£50,662	£45,247	£42,847
	53	£52,351	**£47,503**	£44,426
	65	£57,059	£52,314	£48,327
Female	40	£49,375	£44,032	£41,710
	53	£50,776	£45,727	£43,023
	65	£55,064	£50,394	£46,592

Figure 5-12: Adjusted PSA scatterplot to correspond to revised base-case scenario using the modified Excel-based model



5.7 ERG revised additional comparison results: abatacept + MTX versus cycled TNFi

The impact of the various corrections and amendments described in Section 5.4 and 5.6 on the additional comparison can be judged from Table 5-15. Two changes are especially influential (i) the use of the adjusted TNFi clinical effectiveness estimate (ii) reconciling effectiveness, mortality rates and drug costing for the overall population. However, the cumulative importance of the other changes should not be underestimated.

The revised economic results incorporating all alterations reveal strong reductions in both incremental health outcome gains and in incremental costs, so that the ICER for abatacept + MTX versus cycled TNFi increases from £22,602 to £50,222 per QALY gained. If, in addition, the ERG-preferred HAQ progression rates are used, the ICER increases further to £67,459 per QALY gained.

Table 5-15: Marginal effects of identified corrections/amendments and revised on alternative scenario cost-effectiveness results - using the Excel-based model

		Abatacept+MTX				**TNFi treatments**				**Incremental**				**ICER**
		LYs	QALYs	Costs	Final HAQ	Lys	QALYs	Costs	Final HAQ	LYs	QALYs	Costs	Final HAQ	Cost /QALY
1	**Submitted alternative scenario**	**16.239**	**4.706**	**£84,678**	**2.449**	**15.833**	**3.610**	**£59,907**	**2.638**	**0.405**	**1.096**	**£24,771**	**-0.189**	**£22,601**
2	Discounting correction	16.239	4.828	£86,842	2.449	15.833	3.703	£61,426	2.638	0.405	1.125	£25,416	-0.189	£22,599
3	Sampling treatment correction	16.224	4.672	£83,910	2.459	15.829	3.600	£59,675	2.641	0.396	1.072	£24,235	-0.182	£22,602
4	Abatacept discontinuation rate	15.854	3.737	£62,914	2.673	15.714	3.326	£53,343	2.712	0.140	0.411	£9,571	-0.039	£23,276
5	HAQ mortality multiplier	17.297	4.819	£87,867	2.513	17.076	3.721	£63,366	2.712	0.221	1.098	£24,502	-0.199	£22,314
6	Disease related hospital costs	16.239	4.706	£64,497	2.449	15.833	3.610	£37,508	2.638	0.405	1.096	£26,989	-0.189	£24,626
7	NSAIDs use	16.239	4.706	£87,830	2.449	15.833	3.610	£62,993	2.638	0.405	1.096	£24,837	-0.189	£22,662
8	Missing MTX costs	16.239	4.706	£86,205	2.449	15.833	3.610	£59,907	2.638	0.405	1.096	£26,298	-0.189	£23,995
9	Half-cycle correction	16.266	4.769	£84,209	2.452	15.863	3.693	£59,062	2.640	0.403	1.076	£25,147	-0.189	£23,380
10	Bansback utility	16.239	3.170	£84,678	2.449	15.833	2.211	£59,907	2.638	0.405	0.959	£24,771	-0.189	£25,830
11	Estimated benefit on HAQ score	16.178	4.536	£83,287	2.474	15.833	3.610	£59,907	2.638	0.344	0.927	£23,379	-0.164	£25,230
12	Gender reconciliation	15.732	4.636	£89,487	2.416	15.280	3.552	£60,773	2.605	0.452	1.084	£28,714	-0.189	£26,482
13	TNFi effectiveness	16.239	4.706	£84,678	2.449	16.005	4.056	£67,316	2.557	0.233	0.649	£17,362	-0.108	£26,744
1-13	**Revised alternative scenario**	**16.653**	**2.362**	**£45,535**	**2.741**	**16.623**	**2.226**	**£38,736**	**2.755**	**0.029**	**0.135**	**£6,798**	**-0.014**	**£50,222**
-	**+ERG HAQ progression rates**	**17.078**	**3.567**	**£40,011**	**2.017**	**17.061**	**3.464**	**£33,080**	**2.021**	**0.018**	**0.103**	**£6,931**	**-0.004**	**£67,459**

5.7.1 Sensitivity analysis: additional analysis

The interaction of differential parameter values (mortality rates, treatment effectiveness and treatment costs) by gender and their impact on cost effectiveness is shown in summary in Table 5-16. In all cases, abatacept performs poorly in comparison to the three available TNFi. However, the results are generally more favourable for male patients, and compared to treatment with infliximab, though the best ICER obtained exceeds £33,000 per QALY gained.

Table 5-16: Gender-specific cost-effectiveness results for abatacept + MTX versus three TNFi using the ERG revised Excel-based model of the alternative scenario (cost per QALY gained)

Comparator	Males	Females	Overall
Adalimumab + MTX	£41,865	£74,687	£61,299
Etanercept + MTX	£41,862	£74,676	£61,292
Infliximab + MTX	£33,285	£47,231	£42,833
Weighted average TNF □ inhibitor	£37,339	£56,886	£50,222

5.8 Summary of cost-effectiveness evidence

5.8.1 Economic evaluation results

Base-case: Manufacturer

- The manufacturer reports an ICER of £25,395 per QALY gained for the comparison of abatacept + MTX versus MTX. The manufacturer reports an ICER of £22,628 per QALY gained for the comparison of abatacept + MTX versus cycled TNFi
- Results of the probabilistic analysis conducted by the manufacturer suggest that, based on the assumptions made and evidence available, abatacept has a high probability of being cost-effective at a WTP of £30,000/QALY gained

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 corrected and/or adjusted, the ICER for the base-case comparison ranges from £47,503 per QALY gained to £72,865 per QALY gained. The ICER for the abatacept + MTX versus cycled TNFi ranges from £50,222 to £67,459 per QALY gained

5.8.2 Economic issues and uncertainties

- The manufacturer built their economic model in R rather than in one of the standard modelling packages. The ERG carried out an assessment of the cost-effectiveness evidence presented by the manufacturer using the Excel-based version of the model
- In the model, there are several logic errors (misunderstanding of discounting technique, application of incorrect formula, only including MTX costs in the MTX arm and no half-cycle correction) and uncertain parameter value estimates (use of clinical evidence base for abatacept annual discontinuation rate, choice of HAQ mortality multiplier, omission of NSAID and corticosteroid use, drugs and hospital disease related costing methods and representation of abatacept and TNFi treatment effects)
- The ERG also made some structural adjustments to the economic model. Firstly, the ERG constructed an overall mixed gender cohort for the comparisons, as is the norm. Secondly, the ERG had concerns about the use of utility values derived from the US model. Finally, an alternative consideration of evidence for progression rates for HAQ scores was applied in the model

6 DISCUSSION

The manufacturer presents a case for the use of abatacept + DMARDs versus placebo + DMARDS in adult patients with severe RA. In their statement of the decision problem, the manufacturer states that abatacept should be compared with management strategies without abatacept, for example, alternative DMARDs including TNFi and rituximab. The clinical effectiveness section of the MS concentrates primarily on the comparison of abatacept + DMARDs versus placebo + DMARDs as conducted in the ATTAIN trial.

The systematic literature review conducted by the manufacturer was designed to identify the clinical evidence available for the assessment of the efficacy and safety of abatacept for the treatment of RA patients who had failed a TNFi. The ERG is confident that all relevant published clinical trials were identified by the manufacturer. The literature search identified a single RCT (ATTAIN) conducted by the manufacturer comparing abatacept + DMARDs versus placebo + DMARDs. The MS also included details of the characteristics, results and quality assessment of five additional RCTs; these RCTs do include abatacept as a comparator but do not meet the SR inclusion criteria.

Results from the ATTAIN trial furnish the principal clinical evidence presented in the MS. The manufacturer provided the CSR from the ATTAIN trial which enabled the ERG to conduct some analyses on the IPD therein.

The ATTAIN trial appears to be a well-conducted RCT, the results of which seem to demonstrate that abatacept + DMARDs is more clinically effective than placebo + DMARDs. At 24 weeks, ACR20/50/70 responses are greater in the abatacept arm than in the placebo arm. Mean change in HAQ score is also higher in the abatacept group than in the placebo group. As the patients who would be eligible to receive abatacept are difficult to treat, having severe disabling disease with marked impairment of quality of life, the results of the ATTAIN trial are convincing for this patient population.

Unfortunately, the clinical evidence from the ATTAIN trial does not allow the manufacturer to answer the questions raised in their statement of the decision problem. Evidence from the ATTAIN trial does not provide any answers to the question of whether or not abatacept is more clinically effective when compared to another DMARD (e.g. leflunomide, second or third TNFi) or rituximab. However, the manufacturer does attempt to compare abatacept with a second TNFi using BSRBR data[25,35] and also abatacept with rituximab using a MTC. The manufacturer concludes that both of these comparisons should be interpreted with caution due to inherent data limitations. The MS therefore is not able to examine the optimal sequencing of abatacept with conventional and biologic DMARDS.

The cost-effectiveness section of the MS considers two relevant comparisons using clinical evidence from the ATTAIN trial and a recent BSRBR report.[35] The base-case comparison is abatacept + MTX versus MTX. The manufacturer also carries out an additional comparison of abatacept + MTX versus cycled TNFi. The manufacturer repeatedly states that this comparison is speculative due to data constraints.

In their economic analysis, the manufacturer concludes that abatacept + MTX should be considered cost-effective when compared to both MTX and cycled TNFi.

In the base-case, abatacept + MTX is compared with MTX after failure of at least one TNFi. In the MS, there are no treatment options considered for patients who fail on abatacept i.e. abatacept is not placed within a sequence of treatments; it appears to be the final treatment option for patients. Whether or not this is a true reflection of abatacept's future position in clinical practice is debatable as there may be other options (e.g. leflunomide) that could benefit some patients. If placed in a sequence of treatments it is likely that abatacept + MTX would be considered to be even less cost-effective than the ERG currently demonstrates.

The appropriateness of the treatment pathways selected for modelling by the manufacturer is likely to be controversial within the medical community. The manufacturer describes as the base-case a comparison between abatacept + MTX and MTX (MTX being considered representative of a range of DMARDs). A critical variable in the economic model is the assumed annual rate of disease progression (measured by HAQ score) whilst on MTX monotherapy, which is set to a high value of 0.06. This rate is intended to reflect the rate of treatment progression during periods of non-response to treatment in patients who have failed at least two DMARDs; i.e. when all active treatment options have been exhausted and the patient is deemed to receive only palliative therapies (in this case including MTX). If this is not the case, then the rate of disease progression would be expected to be somewhat lower, with important consequences for estimates of cost-effectiveness.

In the additional comparison performed by the manufacturer, abatacept + MTX is compared to cycled TNFi. The estimate of the clinical effectiveness of cycled TNFi used in the model is taken directly from the BSRBR report.[25,35] As there are no RCTs directly comparing abatacept with TNFi, use of this source of clinical evidence is appropriate. However, the manufacturer chose to ignore the adjusted estimate extensively described in the BSRBR report[25,35] in favour of the unadjusted estimate. In the MS the manufacturer does not provide sufficient justification for this choice. This is unfortunate as this parameter value estimate has a significant effect on the magnitude of the ICER.

On detailed examination of the model, the ERG identified significant errors. In addition, the ERG was not confident that the most appropriate methods of estimating specific costs and

benefits had been employed in the model. The ERG also included amendments that favour the manufacturer's case for the use of abatacept: e.g. lower acquisition costs due to more accurate vial estimation and reduced number of loading doses. As previously described, there are a number of clinical and economic issues and assumptions that call into question the validity of the manufacturer's claims of cost-effectiveness and the credibility of the ICERs generated.

Using alternative ERG assumptions and parameters, and correcting for errors in the model, has the effect of generating substantially worsened cost-effectiveness results for the two comparisons described in the MS. The ICER for the abatacept + MTX versus MTX ranges from £47,503 per QALY gained to £72,865 per QALY gained. Sensitivity analysis demonstrates that younger patients, females and more severe baseline HAQ are all associated with better cost-effectiveness ratios. However, in all cases the ICER remains outside the range normally considered cost-effective.

The ICER for the abatacept + MTX versus cycled TNFi ranges from £50,222 to £67,459. Sensitivity analysis reveals that the cost-effectiveness results are generally more favourable for male patients, and for treatment with infliximab, though the best ICER obtained exceeds £33,200 per QALY gained.

In summary, the consequence of the corrections and adjustments made to the manufacturer's model is that the economic results for use of abatacept no longer appear to support the claims of cost-effectiveness made in the MS.

6.1 Implications for future research

There are no published RCTs of abatacept versus any other relevant comparator (e.g. second or third TNFi, or rituximab) to inform current clinical decision-making. Future trials are therefore necessary in order to undertake comprehensive comparisons of abatacept with all relevant treatments 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 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 relatively recent introduction of abatacept in this patient population, there is a paucity of long-term evidence for both the continued benefit of abatacept and its long-term comparative safety. Close monitoring and surveillance of patients in receipt of abatacept are therefore necessary.

7 REFERENCES

[1] Genovese MC, Becker J-C, 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. September 15, 2005;353(11):1114-23.

[2] Doran MF, Pond GR, Crowson CS, O'Fallon WM, SE. G. Trends in incidence and mortality in rheumatoid arthritis in Rochester, Minnesota, over a forty-year period. Arthritis Rheum. 2002;46(3):625-31.

[3] Gabriel S, Crowson C, O'Fallon W. Mortality in rheumatoid arthritis: have we made an impact in 4 decades? Journal of Rheumatology. 1999;26(12):2529-33.

[4] Symmons D, Barrett E, Bankhead C, Scott D, Silman A. The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. British Journal of Rheumatology. 1994;33(8):735-9.

[5] Uhlig T, Kvien T, Glennas A, Smedstad L, O. F. The incidence and severity of rheumatoid arthritis, results from a county register in Oslo, Norway. Journal of Rheumatology. 1998;25(6):1078-84.

[6] Hyrich K, Symmons D, Watson K, Silman A. Baseline comorbidity levels in biologic and standard DMARD treated patients with rheumatoid arthritis: results from a national patient register. Ann Rheum Dis. 2006 Jul;65(7):895-8.

[7] Kvien T. Epidemiology and Burden of Illness of Rheumatoid Arthritis. Pharmacoeconomics 2004;22(Suppl 1):1-12.

[8] Mitchell D, Spitz P, Young D, Bloch D, McShane D, Fries J. Survival, prognosis, and causes of death in rheumatoid arthritis. Arthritis Rheum. 1986;29(6):706-14.

[9] Symmons D, Jones M, Scott D, Prior P. Longterm mortality outcome in patients with rheumatoid arthritis: early presenters continue to do well. . J Rheumatol 1998;25(6):1072-7.

[10] Maradit-Kremers H, Crowson CS, Nicola PJ, Ballman KV, Roger VL, Jacobsen SJ Iuchdasdiraap-bcsAR-.

[11] Wolfe F, Mitchell D, Sibley J, Fries J, Bloch D, Williams C, et al. The mortality of rheumatoid arthritis. Arthritis & Rheumatism. 1994;37(4):481-94.

[12] Kosinski M, Kujawski S, Martin R, Wanke L, Buatti M, Ware Jr J, et al. Health-related quality of life in early rheumatoid arthritis: impact of disease and treatment response. The American Journal of Managed Care. 2002;8(3):231-40.

[13] Cooper NJ. Economic burden of rheumatoid arthritis: a systematic review. Rheumatology. 2000 January 1, 2000;39(1):28-33.

[14] Brouwer W, van Exel N, van de Berg B, Dinant H, Koopmanschap M, van den Bos G. Burden of caregiving: evidence of objective burden, subjective burden, and quality of life impacts on informal caregivers of patients with rheumatoid arthritis. Arthritis & Rheumatism. 2004;51(4):570-7.

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

[16] National Institute for Health and Clinical Excellence. Final scope for the appraisal of abatacept for rheumatoid arthritis. 26 October 2006. 2006 [cited 20 April 2007]; Available from: http://www.nice.org.uk/page.aspx?o=380240

[17] Hyrich K, Lunt M, Watson K, Symmons D, Silman A. Outcomes after switching from one anti-tumor necrosis factor alpha agent to a second anti-tumor necrosis factor alpha agent in patients with rheumatoid arthritis: results from a large UK national cohort study. Arthritis & Rheumatism. 2007;56(1):143-50.

[18] 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.

[19] 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).

[20] Landewe R, Boers M, Verhoeven A, Westhovens R, van de Laar M, Markusse H, et al. COBRA combination therapy in patients with early rheumatoid arthritis: long-term structural benefits of a brief intervention. Arthritis & Rheumatism. 2002;46(2):347.
[21] National Institute for Health and Clinical Excellence. Final appraisal determination: Adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis. 27 November 2006. 2006 [cited 20 April 2007]; Available from: http://www.nice.org.uk/page.aspx?o=388554
[22] National Institute for Health and Clinical Excellence. Anakinra for rheumatoid arthritis. London: National Institute for Health and Clinical Excellence; 2003.
[23] 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.
[24] Australian Rheumatology Association. A practical guide for the use of biologic agents in the treatment of rheumatoid arthritis. APLAR Journal of Rheumatology. 2006;9(2):123-6.
[25] British Society of Rheumatology (BSR). Effect of a second course of anti-TNF therapy on HAQ following lack of response to the first course. 2006 27 November [cited; Available from: http://guidance.nice.org.uk/download.aspx?o=388219
[26] 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
[27] Westhovens R, Cole JC, Li T, Martin M, MacLean R, Lin P, et al. Improved health-related quality of life for rheumatoid arthritis patients treated with abatacept who have inadequate response to anti-TNF therapy in a double-blind, placebo-controlled, multicentre randomized clinical trial. Rheumatology. 2006 October 1, 2006;45(10):1238-46.
[28] Sibilia J, Schiff M, Genovese MC, Becker JP, Li T, McCann T, et al. Sustained improvements in disease activity score 28 (das28) and patient (pt)-reported outcomes (pro) with abatacept (aba) in rheumatoid arthritis (ra) pts with an inadequate response to anti-tnf therapy: the long-term extension (lte) of the attain trial. *EULAR* Amsterdam, The Netherlands 2006.
[29] Clinical study report. Attain LTE. A phase III study to evaluate the safety and clinical efficacy of intravenous infusions of BMS-188667 (abatcept) vs placebo in subjects with active rheumatoid arthritis on background DMARDs who have failed anti-TNF therapy. Open label period up to 6 June 2005. Final clinical study report. Protocol IM101029: Bristol Myers Squibb Co.
[30] Bristol Myers Squibb. 064: The safety of abatcept in patients with active rheumatoid arthritis and an inadequate response to anti-TNF therapy: Results from the ARRIVE trial (data on file).
[31] Food and Drug Administration. Guidance for industry clinical development programs for drugs, devices, and biological products for the treatment of rheumatoid arthritis. FDA 1999.
[32] EMEA, Committee for proprietary medicinal products (CPMP). Points to consider on clinical investigations of medicinal products other than NSAIDS for the treatment of rheumatoid arthritis. 2003 02 April [cited; Available from: http://www.tga.gov.au/docs/pdf/euguide/ewp/055695enrev1.pdf
[33] Clinical study report. IM101029 A phase III, Multi-centre, randomised, double-blind, placebo-controlled study to evaluate the efficacy and safety of abatcept in subjects with active rheumatoid arthritis on background DMARDs who have failed anti-TNF therapy: Bristol Myers Squibb Co. October 2004.
[34] Wang L, Smittten A, Simon T. Risk of site specific malignancies in adult patients with rheumatoid arthritis. A systemic literature review. Ann Rheum Dis. 2006;65(suppl II):488.
[35] British Society of Rheumatology. Effect of a second course of anti-TNF therapy on HAQ following lack of response to the first course. Additional analysis submitted to NICE for

the Technology Appraisal of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis. 2006.
[36] 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.
[37] Davies A, Cifaldi M, Weisman M. Cost-effectiveness of the 3 TNF protagonists vs. abatacept in the treatment of moderate to severe rheumatoid arthritis. EULAR Congress; 2006 June 21-24; Amsterdam; 2006.
[38] Brennan A, Bansback N, Nixon R. Modelling the cost-effectiveness of TNF inhibitors in the management of rheumatoid arthritis: results from the British Society for Rheumatology Register. An independent report: University of Sheffield School of Health and Related Research (ScHARR); 2006. Report No: 06/12.
[39] Thorne G, Beresniak A, Haraoui B, et al. Cost-effectiveness of abatacept versus other biologic agents in DMARD and anti-TNF inadequate responders for the management of moderate to severe rheumatoid arthritis in Canada. Canadian Rheumatology Association Conference and Annual Meeting; 2007 Feb 21-24; Lake Louise, Alberta; 2007.
[40] Kremer JM, Genant HK, Moreland LW, Russell AS, Emery P, Abud-Mendoza C, et al. Effects of Abatacept in Patients with Methotrexate-Resistant Active Rheumatoid Arthritis: A Randomized Trial. Annals of Internal Medicine. 2006 June 20, 2006;144(12):865-76.
[41] NICE. Guide to the methods of technology appraisal. London: National Institute for Health and Clinical Excellence; 2004.
[42] The General Practice Research Database. Data until October 2005.
[43] Wolfe F. National Data Bank for Rheumatic Diseases. US 2004.
[44] Hurst N, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). British Journal of Rheumatology. 1997(36):551-9.
[45] 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.
[46] Department of Health. NHS Reference Costs 2005-2006.
[47] Sokka T, Hakkinen A, Krishnan E, Hannonen P. Similar prediction of mortality by the health assessment questionnaire in patients with rheumatoid arthritis and the general population. Annals of the Rheumatic Diseases. 2004;63:494-7.
[48] British National Formulary. BNF 52. 2006 March [cited 2007; Available from: http://www.bnf.org/bnf/
[49] Department of Health. NHS Reference Costs. 2005/6 [cited 2007 March]; Available from: http://www.dh.gov.uk/prod_consum_dh/idcplg?IdcService=GET_FILE&dID=115839&Rendition=Web
[50] Deighton C, Gadsby K. The new British Society for Rhuematology disease modifying anti-rheumatic drugs guidelines Musculoskeletal Care. 2006;4(3):174-81.
[51] 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.
[52] Hurst N, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of Euroqol EQ-5D. British Journal of Rheumatology. 1997;36:551-9.
[53] Hawthorne G, Richardson J, Day N, Osborne R, McNeil H. Construction and utility scaling of the assessment of quality of life (AQOL) instrument. Melbourne, Australia: Centre for Health Program Evaluation; 2000.
[54] 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.

[55] Scott DL, Garrood T. Quality of life measures: use and abuse. Baillieres Best Pract Res Clin Rheumatol. 2000;14(4):663-87.
[56] Brennan A, Bansback N, Reynolds A, Conway P. Modelling the cost-effectiveness of etanercept in patients with rheumatoid arthritis in the UK. Rheumatology. 2004;43(1):62-72.
[57] Genovese M, Bathon J, Martin W, et al. Etanercept versus methotrexate in patients with early rheumatoid arthritis. Two-year radiographic and clinical outcomes. Arthritis & Rheumatism. 2002;46:1443-50.
[58] 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.
[59] Scott D, Pugner K, Kaarela K, et al. The links between joint damage and disability in rheumatoid arthritis. Rheumatology. 2000;39:122-32.
[60] 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.
[61] 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.
[62] Sherrer YS, Bloch DA, Mitchell DM, Young DY, Fries JF. The development of disability in rheumatoid arthritis. Arthritis & Rheumatism. 1986 Apr;29(4):494-500.
[63] Ward MM, Leigh JP, Fries JF. Progression of functional disability in patients with rheumatoid arthritis. Associations with rheumatology subspecialty care. Arch Intern Med. 1993;153(19):2229-37.
[64] 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. Br J Rheumatol. 1993;32(8):724-8.
[65] 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 Res. 1997;10(6):381-94.
[66] Leymarie F, Jolly D, Sanderman R, Briancon S, Marchand AC, Guillemin F, et al. Life events and disability in rheumatoid arthritis: a European cohort. Br J Rheumatol. 1997;36(10):1106-12.
[67] 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.
[68] 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.
[69] Kapral T, Stamm T, Machold K, Montag K, Smolen J, Aletaha D. Methotrexate in rheumatoid arthritis is frequently effective, even if re-employed after a previous failure. Arthritis Research and Therapy. 2006;8:R46.
[70] Russell AS, Wallenstein GV, Li T, Martin MC, Maclean R, Blaisdell B, et al. Abatacept improves both the physical and mental health of patients with rheumatoid arthritis who have inadequate response to methotrexate treatment. Annals of the Rheumatic Diseases. 2007 Feb;66(2):189-94.
[71] Weinblatt M, Combe B, Covucci A, Aranda R, Becker JC, Keystone E. Safety of the selective costimulation modulator abatacept in rheumatoid arthritis patients receiving background biologic and nonbiologic disease-modifying antirheumatic drugs: A one-year randomized, placebo-controlled study. Arthritis & Rheumatism. 2006;54(9):2807-16.
[72] Emery P, Kosinski M, Li T, Martin M, Williams GR, Becker JC, et al. Treatment of rheumatoid arthritis patients with abatacept and methotrexate significantly improved health-related quality of life. Journal of Rheumatology. 2006;33(4):681-9.
[73] Kremer JM, Dougados M, Emery P, Durez P, Sibilia J, Shergy W, et al. Treatment of rheumatoid arthritis with the selective costimulation modulator abatacept: twelve-month

results of a phase iib, double-blind, randomized, placebo-controlled trial. Arthritis Rheum. 2005 Aug;52(8):2263-71.
[74] Kremer JM, Westhovens R, Leon M, Di Giorgio E, Alten R, Steinfeld S, et al. Treatment of Rheumatoid Arthritis by Selective Inhibition of T-Cell Activation with Fusion Protein CTLA4Ig. New England Journal of Medicine. 2003 November 13, 2003;349(20):1907-15.
[75] Weinblatt M, Schiff M, Goldman A, Kremer J, Luggen M, Li T, et al. Selective costimulation modulation using abatacept in patients with active rheumatoid arthritis while receiving etanercept: a randomised clinical trial. Annals of the Rheumatic Diseases. 2007 Feb;66(2):228-34.
[76] Clinical study report. IM101043 A phase III, Multi-centre, randomised, double-blind, placebo-controlled comparative study of abatcept or infliximab in combination with methotrexate in controllling disease activity in subjects with rheumatoid arthritis having an inadequate clinical response to mehtotrexate Bristol Myers Squibb Co. September 2006.

Appendix 1

Table A-1: Characteristics of the five RCTS outside of the licensed indication

Study Name	Study intervention(s), comparator(s), drug, dose(s) and follow-up	Study design, location and enrolment	Study inclusion and exclusion criteria	Study outcomes
IM101102 (AIM) Kremer 2006[40] Russell 2006[70]	Abatacept 10mg per kg of body weight + MTX (N=433) Placebo + MTX (N=219) Medication was administered via a 30-minute intravenous infusion on days 1, 15, 29 and every 28 days thereafter up to and including day 337 Study duration 52 weeks	Double-blind phase III RCT from 102 sites between November 2002 and October 2004 Patients completing the double-blind phase of the study were allowed to enter a long-term, open-label extension phase (during which all patients received a fixed dose of abatacept of 10mg per kg of body weight)	Patients must have/be: • Aged ≥18 years who have RA (according to ACR criteria) for ≥1 year • Persistent and active RA • Taken MTX for ≥3 months (at stable dose for ≥28 days) • ≥10 SJC • ≥12 TJC • CRP levels of ≥1 mg/ • Use of oral corticosteroids (≤10mg of prednisone or its equivalent per day) was allowed if the dose had been stable for ≥25 days Patients were excluded if they had: • Active vasculitis of a major organ system, • Uncontrolled renal, hepatic, haematological, gastrointestinal, pulmonary, cardiac, neurological, or cerebral disease • History of cancer within the last 5 years • Serious bacterial infection in the previous 3 months • Active TB or herpes zoster history, Had surgery on more than 5 joints	Primary outcomes (at 1 year except where stated): • ACR20 response (at 6 months) • HAQ improvements of ≥0.3 • Radiographic progression of joint erosions (Genant-modified Sharp score) Secondary outcomes (at 1 year except where stated): • ACR20, ACR50 response, ACR70 response (at 6 months and 1 year) • DAS28 • HAQ mean improvement • SF-36 changes • Radiographic progression of joint space narrowing (Genant-modified Sharp score) • AEs • Immunogenicity testing
IM101031 (ASSURE) Weinblatt	Abatacept 10mg per kg of body weight + background DMARD (N=959)	Double-blind phase III RCT from 161 sites between December 2002 and June 2004	Patients must have/be: • Aged ≥18 years • Met ACR criteria for RA	Primary outcomes were safety assessments during the study period (of 1 year) described as: • AEs

Study Name	Study intervention(s), comparator(s), drug, dose(s) and follow-up	Study design, location and enrolment	Study inclusion and exclusion criteria	Study outcomes
2006[71]	Placebo + background DMARD (N=482) Medication was administered via a 30-minute intravenous infusion on days 1, 15, and 29, and every 4 weeks thereafter, for a total of 14 doses Study duration 26 weeks	Patients completing the double-blind phase of the study were allowed to enter a long-term, open-label extension phase (during which all patients received a fixed dose of abatacept of 10mg per kg of body weight)	• Had RA for ≥1 year • RA functional class I, II, III or IV • Active RA as defined by VAS ≥20mm • Taken any DMARD or for ≥3 months (at stable dose for ≥28 days) Patients were excluded if they had • Unstable or uncontrolled renal, endocrine, hepatic, haematologic, gastrointestinal, pulmonary, cardiac, or neurologic diseases • Any autoimmune disorder other than RA as the main diagnosis • Active or chronic recurrent bacterial infections unless treated and resolved • Active herpes zoster infection within the previous 2 months • Hepatitis B or hepatitis C virus infection • Active or latent tuberculosis	• SAEs • Discontinuations due to AEs Secondary outcomes were: • HAQ score • 100mmVAS for: o Patient global assessment of disease activity o Patient global assessment of pain o Physician global assessment of disease activity
IM101100 Emery 2006[72] Kremer 2005[73] Kremer 2003[74]	Abatacept 2mg per kg of body weight + MTX (N= 105) Abatacept 10mg per kg of body weight + MTX (N= 115) Placebo +MTX (N= 119) Medication was administered via a 30-minute intravenous infusion on days 1, 15, 29 and every 30 days thereafter Study duration 52 weeks	Double-blind phase IIb RCT from 66 sites Dates of trial enrolment:11/12/02 – 13/06/02 Patients completing the double-blind phase of the study were allowed to enter a long-term, open-label extension phase (during which all patients received a fixed dose of abatacept of 10mg per kg of body weight)	Patients must have/be: • 18-65 years old • Met ACR criteria for RA • Functional class I, II, III • ≥10 SJC • ≥12 TJC • C-RP≥1mg/dL • Treated with MTX for at least 6 months with a stable dose for ≥28	Primary outcome at 6 months: • ACR20 Secondary outcomes at 6 months: • ACR50, ACR70 • SF-36 • Safety assessment Secondary outcomes at 12 months: • ACR20, ACR50, ACR70 • MHAQ • DAS28 • SF-36 • Safety assessment • Immugenecity testing (serum samples)

Study Name	Study intervention(s), comparator(s), drug, dose(s) and follow-up	Study design, location and enrolment	Study inclusion and exclusion criteria	Study outcomes
IM101101 Weinblatt 2007[75]	Abatacept 2mg per kg of body weight + etanercept 25mg (N=85) Placebo + etanercept 25mg (N=36) Etanercept was administered twice weekly while abatacept was administered on days 1, 15, and 29, and every 4 weeks thereafter Study duration 52 weeks (patients who achieved at least a 50% reduction in both their SJC and TJC at 6 months discontinued abatacept)	Double-blind phase IIb RCT from 40 centres in the US Trial carried out between 26/02/01-18/09/02	Patients must have/be: • Met ACR criteria for RA • Functional class I, II, III • ≥8 SJC • ≥10 TJC • CRP ≥2mg/dL in original protocol but this was modified due to the effect of etanercept on normalising CRP levels Patients were excluded if they had • Latent infection • Recent opportunist infection • Tuberculosis requiring treatment within last 3 years • History of cancer within last 5 years • History of drug or alcohol abuse • Pregnant and nursing women	Primary outcome at 6 months: • Modified ACR 20 (exclude CRP levels) Secondary outcome at 6 months: • Modified ACR50, ACR70 (excludes CRP levels) • Standard ACR20 • ACR50, ACR70 • Improvements in individual ACR components • SF-36 changes The primary objective of the open-label long-term extension was to assess the safety and tolerability of abatacept in combination with etanercept during long-term administration in patients with active RA: • AEs • SAEs • Blood samples • Clinical laboratory tests
IM101043[76]	Abatacept: 500mg for patients <60kg, 750mg for patients 60–100kg and 1g for patients >100kg + MTX Infliximab: 3mg/kg IV (approved labelled dose) + MTX Placebo +MTX: After 6 months patients receiving placebo were allowed to receive abatacept Abatacept was administered on days 1, 15, 29 and every 28 days thereafter for a total of 14 doses Infliximab administered on days 1, 15, 43, 85 and every 56 days thereafter for a total of 8 doses Saline was administered where necessary to prevent unblinding	Double-blind phase IIIb RCT from 86 study sites Dates of trial enrolment:03/02/05-19/06/06	Patients must have/be: • Met ACR criteria for RA • Functional class I, II, III • ≥10 SJC • ≥12 TJC • CRP >1.0 mg/dL • Treated with MTX for at least 3 months with at least a weekly dose of 15 mg, and at a stable dose for 28 days prior to treatment. • At least 18 years of age. Patients were excluded if they had: • Active vasculitis of a major organ system, • Uncontrolled renal, hepatic, haematological, gastrointestinal, pulmonary, cardiac, neurological, or cerebral disease, • History of cancer within last 5 years,	• DAS28: high disease activity (>5.1), low disease activity (<3.2) and remission (<2.6). • Physical function measured by HAQ at day 365 • Physical function was evaluated using the HAQ, which is a validated assessment for functional status and degree of disability • HRQoL - SF36 • Symptomatic relief measured by ACR20 • Pharmacodynamic markers: RF, CRP, ESR, IL-6, and receptor activator of NF-κB ligand (RANKL)

Study Name	Study intervention(s), comparator(s), drug, dose(s) and follow-up	Study design, location and enrolment	Study inclusion and exclusion criteria	Study outcomes
	Study duration 52 weeks		• Serious bacterial infection in the previous 3 months • Active TB or herpes zoster history, • Had had surgery on more than 5 joints • Patients who have previously received treatment with an approved biologic RA therapy (infliximab, etanercept, anakinra, adalimumab). • Patients who have failed more than 4 DMARDs (including MTX) due to lack of efficacy.	

SJC=swollen joint count, TJC=tender joint count, MHAQ=modified HAQ, CRP=C-reactive protein, ESR= Erythrocyte sedimentation rate, AE = adverse event, SAE = serious adverse event, DAS28 = Disease activity score, SF-36 = Short form 36