Adalimumab for the treatment of psoriatic arthritis

Erratum for Premeeting briefing and ERG report

After issuing the premeeting briefing and ERG report to the Appraisal Committee the following error was identified.

The manufacturer submission reported that utilities where calculated using M02-570 and data from another study (Table 6.2.6.1, page 68 in the manufacturer submission). The evidence review group (ERG) considered that this additional study involved patients with more severe disease leading to a potential over estimate of the impact of psoriasis on quality of life estimates in the model. The manufacturer clarified that these data had been used in an earlier version of the model, but were not utilised in the submitted version of the model and that reference to the use of these data was an error within their submission report.

This information was not included in the clarification response forwarded to the ERG and therefore not incorporated into the premeeting briefing and the ERG report. Statements on the use and potential effects of the additional study data are therefore retracted and should be disregarded in these documents.

Premeeting briefing:

- Page 3, 2nd bullet, 3rd subbullet
- Page 12, 5th bullet

ERG report:

- Page 66; section 5.3.7 (Health-related quality of life); paragraph 2, 3rd sentence
- Page 89; section 5.5.6 (Health-related quality of life); paragraph 1, 2nd through to 9th sentance.

End of erratum

ADALIMUMAB FOR THE TREATMENT OF MODERATE TO SEVERE PSORIATIC ARTHRITIS

THE EVIDENCE REVIEW GROUP'S REPORT

The Evidence Review Group consists of:

Centre for Health Economics, University of York

Y Bravo Vergel, S Palmer

&

Regional Drug and Therapeutics Centre (Newcastle)

S Erhorn, V Young, S Brent, A Dyker, W Horsley, K Macfarlane, J Martin, B Reddy, S White, S Thomas

About the authors:

The ERG is a collaboration between two centres, the Regional Drug and Therapeutics Centre (RDTC) and the Centre for Health Economics (CHE).

The RDTC was established in 1991 to promote safe effective prescribing and economical drug usage, and to provide a source of independent authoritative advice on pharmaceutical and therapeutic issues throughout the former Northern and Yorkshire region. The RDTC co-ordinates prescribing activities, provides a poisons and medicine information service and is the teratology information service for the UK. The Centre is one of four NHS regional monitoring centres for the Medicines and Healthcare Regulatory Authority (MHRA).

CHE is a research unit of the University of York. The Centre's aim is to undertake high quality research that is capable of influencing health policy decisions. The largest programme of work at CHE is that on economic evaluation and health technology assessment which focuses on a range of methodological and applied work. This includes full technology assessment reviews and evidence review reports for the National Institute for Health and Clinical Excellence (NICE).

Contact the authors:

Regional Drug and Therapeutics Centre
Wolfson Unit, Claremont Place, Newcastle upon Tyne, NE2 4HH

Centre for Health Economics
Alcuin 'A' Block, University of York, Heslington, York, YO10 5DD

Acknowledgements:

The ERG received clinical advice from Dr Ian Bruce (Reader in Rheumatology, ARC Epidemiology Unit, University of Manchester) and Dr Robert Chalmers (Consultant Dermatologist and Honorary Clinical Lecturer, Manchester Royal Infirmary). The interpretation and views expressed in this document, and any errors, relate to the ERG authors alone.

Conflicts of interest:

The authors to this report have no conflicts of interest.

CONTENTS

TABLES

FIGURES

Chapter 1
Summary

1. Introduction

This document critically evaluates the evidence submission, from Abbott Laboratories Ltd, on the clinical and cost-effectiveness of adalimumab (Humira®) for the treatment of moderate to severe psoriatic arthritis (PsA).[1] This report identifies the submission's strengths and weaknesses, supplemented, where appropriate, with our own analysis. Clinical experts were asked to advise the Evidence Review Group (ERG) to help inform the review.

1.1 Scope of the submission

The stated aim of the submission was to evaluate the clinical and cost-effectiveness of adalimumab for the treatment of moderate to severe PsA in accordance with the licensed indication. Adalimumab has a marketing authorisation for the treatment of active and progressive PsA in adults when the response to previous disease-modifying antirheumatic drugs (DMARD) therapy has been inadequate.[2]

1.2 Summary of submitted clinical evidence

The manufacturer based the submission on four clinical trials. The first, a double-blind randomised controlled trial (RCT) in patients with active PsA and an inadequate response or intolerance to non-steroidal anti-inflammatory drug (NSAID) therapy, has been published in full. The second, a double-blind RCT in patients with active PsA and an inadequate response to DMARD therapy, has been published in abstract form only. Both RCTs were of adults (aged ≥18) with active PsA defined as ≥3 swollen joints and ≥3 tender or painful joints. The third study, an open-label, long term extension of the two RCTs in which all participants received adalimumab has also been published in abstract form. The fourth, an open-label study in patients with active PsA who had previously failed treatment with other ant- tumour necrosis factor (TNF) agents, is also published in abstract form only.

The submission also included a meta-analysis of the two aforementioned RCTs. Only the outcomes relating to the arthritis component were included in this meta-analysis as the smaller of the two studies (M02-570) did not consider outcomes relating to the psoriasis components of the disease. Given that there are no studies directly comparing adalimumab with other anti-TNF agents the manufacturer provided an indirect/mixed treatment comparison to strengthen inference concerning the relative efficacy of adalimumab. All

studies included in the clinical evidence section of the Abbott submission were subjected to a detailed critical appraisal (see Appendix 2).

The limited data available indicate that adalimumab is efficacious in the treatment of PsA with beneficial effects on both joint and psoriasis symptoms and on functional status, in patients who had an inadequate response to previous treatment with NSAIDs (24 week (M02-518) ADEPT trial, n=313) or DMARDs (12 week M02-570 trial, n=100). In both these studies the improvement in arthritis response was significantly greater with adalimumab than with placebo, as assessed by ACR20, ACR50, ACR70 and PsARC responses. Subgroup analyses in the ADEPT study indicate that ACR response rates with adalimumab in combination with methotrexate were similar to those achieved with adalimumab alone. In patients with ≥3% body surface area (BSA) affected by psoriasis, and an inadequate response to NSAIDs, adalimumab significantly improved the signs and symptoms of psoriasis, as measured by PASI. The disability and health related quality of life (HRQoL) of patients was significantly more improved with adalimumab than with placebo in both trials. This improvement in both joint and psoriasis symptoms appeared to be maintained for up to 88 weeks in an open-label long-term extension study

Because three out of the four studies included in the submission are not fully published and report only preliminary results in abstract form there are insufficient data presented to fully assess their validity. While the data presented in these abstracts is supplemented with additional data provided by the manufacturer in the submission, this supplemental data are not in the public domain and therefore cannot be externally validated. Until these studies are fully published and the complete data made available for evaluation, these results and any assumptions based thereon should be interpreted with due caution.

Adalimumab was generally well tolerated in the clinical trials of PsA, with similar incidences of adverse events as with placebo. In combined data from the two pivotal RCTs 3% of adalimumab-treated patients and 4.3% of placebo treated subjects experienced a serious adverse event. The most common treatment-emergent adverse events were upper respiratory tract infections, nasopharyngitis, injection site reactions, headache, psoriasis aggravated and diarrhoea.

Overall the adverse event profile appears to be similar to that associated with use of the drug in rheumatoid arthritis (RA). Additional data from a pooled analysis of the clinical trials of adalimumab in RA included in the submission and the Summary of Product Characteristics (SPC) are relevant, as long-term safety data for adalimumab in PsA are very limited. The most common adverse reaction to adalimumab in trials in patients with RA was

injection site reactions (20% vs. 14% with placebo), most of which were mild and did not require treatment discontinuation.[2] Other adverse events versus placebo included upper respiratory tract infections (17% vs. 13%), headache (12% vs. 8%), rash (12% vs. 6%), and sinusitis (11% vs. 9%).[2] Gastrointestinal symptoms, abnormal laboratory test results, back pain, urinary tract infections and hypertension occurred in ≤10% in any treatment group.[2] In common with other anti-TNF agents, the manufacturer's SPC carries a warning regarding the increased risk of infection (including tuberculosis and other opportunistic infections), and warnings pertaining to the onset or exacerbation of demyelinating disease, including multiple sclerosis.[2] There is concern that anti-TNF agents may increase the risk of cancer, particularly lymphoproliferative malignancies.

1.3 Summary of submitted cost-effectiveness evidence

In the absence of any previous studies assessing the cost-effectiveness of adalimumab for the treatment of PsA, the manufacturer submitted their own *de-novo* cost-effectiveness analysis. The model was based on a probabilistic, micro-simulation approach over a lifetime horizon. Alternative time horizons were also presented. The model evaluated the cost-effectiveness of adalimumab compared to alternative anti-TNF agents (etanercept and infliximab) and a "conventional DMARD" option. The model was designed to incorporate the impact of *both* the arthritis and skin components of PsA in estimating costs, health outcomes (expressed using Quality-Adjusted Life Years, QALYs) and cost-effectiveness. All patients were assumed to receive conventional DMARDs after the failure of initial therapy.

For the base-case model, initial response (and hence the decision to continue with the initial treatment) was defined as meeting the 12-week PsARC response criteria. Alternative decision rules were also explored in the sensitivity analysis based on different response criteria (PsARC and PASI) and alternative time periods for assessing whether to continue treatment (12-weeks and 24-weeks). The correlation between PsARC response and other response parameters (PASI and ACR) was also assessed. The degree of response to PASI and ACR (conditional upon PsARC response status) was then linked to overall HAQ and PASI scores, which in turn were used to predict costs and QALYs using regression approaches. In the absence of direct head-to-head RCTs for the different anti-TNF agents (adalimumab, etanercept and infliximab), the manufacturer applied indirect approaches to estimate the relative effectiveness of the treatments considered in the model.

In the base-case analysis, treatment with adalimumab dominated etanercept (i.e. less expensive and more effective). The incremental cost-effectiveness of adalimumab, compared to conventional DMARDs, was reported to be £25,991 per additional QALY.

Compared to adalimumab, infliximab was estimated to provide a QALY gain of 0.09, but at a significant additional cost (£81,614), with an associated ICER of £209,572 per QALY. At a threshold willingness to pay of £30,000 per QALY, the probability adalimumab is cost-effective was estimated to be 0.8. The manufacturer undertook a number of separate sensitivity analyses, reporting that the results remained robust to a wide-range of alternative assumptions. The results appeared most sensitive to the assumptions related to the annual progression rate of HAQ assumed for patients receiving DMARDs, the source of utility data (i.e. SF-6D or EQ-5D measures) and the exclusion of costs and quality of life associated with the psoriasis component of the disease. Alternative assumptions for all 3 of these components resulted in an ICER of adalimumab, compared to DMARDs, in excess of £30,000 per QALY.

A number of additional analyses were requested by the ERG to explore further the robustness of the base-case results to a series of alternative assumptions identified as part of a critical review of the manufacturer's submission. These analyses focused on the data used to estimate response parameters in the model (in particular the decision to exclude results from the 12-week trials from the base-case analysis) and the cost-effectiveness for particular subgroups (based on previous DMARD use and level of skin involvement). The manufacturer submitted a number of separate addenda to address the ERG's queries, including a more comprehensive synthesis of 12- and 24-week trials. Incorporating this additional evidence appeared to have an important impact on the base-case cost-effectiveness results. In contrast to the base-case analysis from their original submission, adalimumab was now extendedly dominated by etanercept in this revised analysis. However, further analyses presented by the manufacturer raised a number of significant concerns about the validity of the approaches used by the manufacturer to synthesise the response rate data which brought into question the overall validity of the cost-effectiveness estimates presented.

1.4 Commentary on the robustness of submitted evidence.

1.4.1 Strengths

The ERG felt that the majority of the data quoted within the submission was a fair and accurate representation of the original reference data.

The manufacturer's submission was considered to comprise the most relevant source of cost-effectiveness evidence related to the use of adalimumab for PsA. The ERG noted a number of strengths in the manufacturer's cost-effectiveness analysis. In particular, the

evaluation of the impact of the different treatments on both the arthritis and skin components of PsA addressed one of the major limitations of existing cost-effectiveness studies in this area. The overall model structure, approaches to estimating long term costs and outcomes (expressed using QALYs), time-horizon employed and the approach to handling parameter uncertainty were all consistent with the NICE Reference Case for cost-effectiveness analysis. A broad range of sensitivity analyses was also undertaken to explore alternative assumptions and the cost-effectiveness for different patient subgroups.

1.4.2 Weaknesses

The ERG felt that the Abbott submission was generally of poor quality. The document was inadequately referenced and the various sections within it were insufficiently cross-referenced. The sections reporting the results were difficult to interpret and unsatisfactorily discussed. Overall, the submission was not of the quality and clarity that the ERG had expected; consequently, the task of evaluating the submitted evidence was made significantly more problematic. The clinical efficacy data used in the submission were limited, being largely derived from just two RCTs in 415 patients, with only 204 patients having received adalimumab, and two uncontrolled long-term open-label studies. A significant proportion of the reference data presented in the submission was not fully published and only available in abstract form. Therefore the ERG felt that until these studies are fully published and the complete data made available for evaluation these results and any assumptions based thereon must be interpreted with due caution.

The participants in the pivotal RCTs were not entirely representative of the population for which adalimumab is currently licensed as neither population was made up exclusively of patients who had failed to respond to at least two DMARDs (43% & 40% in M02-518, and 61% & 51% in M02-570 for the placebo and adalimumab groups, respectively). However, independent expert clinical advice given to the ERG suggests that the participants in these trials nevertheless represented a population with relatively severe PsA similar to those currently being treated in UK clinical practice. Moreover, the patient population considered in these trials is in accordance with those included in the NICE technology appraisal of etanercept and infliximab for patients with PsA (TA104), by the Scottish Medicines Consortium (SMC) in their assessment of adalimumab and etanercept for PsA, and by the All Wales Medicines Strategy Group (AWMSG) in their assessment of adalimumab for PsA. Therefore, the ERG considered that the patient population considered is representative of current UK practice.[3]

The ERG noted a lack of transparency relating to both the description of the methods in the manufacturer's submission report and the implementation of the cost-effectiveness model. The model itself was considered complex and difficult to follow, particularly given the amount of hidden columns and data contained within the spreadsheet and associated macros. The ERG had a number of important concerns related to the evidence synthesis approaches used by the manufacturer in their base-case analysis. The ERG considered that the general approach employed a number of assumptions that were not adequately justified by the manufacturer. The ERG was concerned that these assumptions increased the possibility of introducing potential bias in the subsequent cost-effectiveness results. The ERG concluded that potentially relevant data were excluded by the manufacturer in their base-case analysis, which was likely to result in overly optimistic estimates of the cost-effectiveness of adalimumab. The ERG also felt that a number of subgroup analyses had not been adequately considered in the manufacturer's original submission. After reviewing the additional analyses submitted by the manufacturer in response to these queries, the ERG felt that the results lacked face validity, bringing into question the robustness of the overall evidence synthesis approach and/or assumptions used by the manufacturer.

1.4.3 Areas of uncertainty

Adalimumab is currently only licensed as a monotherapy and not in combination with DMARDs for the treatment of PsA. Therefore, it should be noted that a large proportion of those patients receiving adalimumab in the two RCTs were also receiving concomitant DMARD therapy (51% and 65%, in M02-518 and M02-570, respectively). It is possible that benefit perceived to be due to adalimumab treatment may in part be due to concomitant DMARD use.

The majority of participants in the two pivotal RCTs had PsA with a polyarticular pattern of disease. Although this form of PsA is the most frequent,[4] occurring in around 65% of patients,[3, 5] it is unclear whether these results can be extrapolated to those patients with other forms of PsA, such as spondylitis and arthritis mutilans. Nevertheless, by using mainly patients with polyarticular disease these trials provide appropriate data for evaluating the clinical and cost-effectiveness of adalimumab, as these patients represent the group with the worst prognosis.[6]

To date there are no trials directly comparing adalimumab with other anti-TNF agents for the treatment of PsA, therefore its relative efficacy and safety remains unclear. In the absence of direct head-to-head evidence against other anti-TNF agents, the relative effectiveness of adalimumab in relation to the full range of treatment alternatives relevant to the NHS can

only be assessed using indirect approaches. There remains considerable uncertainty as to the potential biases that this may introduce into the subsequent results. These biases could be minimised by considering all relevant trial evidence and maintaining the randomisation status of the comparisons used. To date there has been no attempt to utilise the full range of potentially relevant trial evidence without breaking the randomised comparisons from the individual trials. In addition, there remains considerable uncertainty as to the most appropriate approach to incorporate different trial follow-up periods and in relating these to the current guidelines on the management of PsA. The degree of heterogeneity between the different studies of the anti-TNF agents (particularly the proportion of patients with over 3% BSA affected by psoriasis) and the need to adjust for potential confounders, prior to any indirect comparison, remains an important source of uncertainty, as does the degree of correlation between different response types. Consequently, the current estimates of the relative effectiveness and cost-effectiveness of adalimumab appear highly uncertain and prone to a number of potential biases.

There is considerable uncertainty regarding the potential of adalimumab to trigger the development of autoimmune antibodies. The formation of antibodies to adalimumab may have significant clinical implications as immunogenicity may be associated with a shortened duration of clinical response. In three RCTs of adalimumab in RA (n=1062), approximately 6% of patients receiving the agent developed anti-adalimumab antibodies during treatment.[7] Furthermore, the incidence of antibodies was higher in patients receiving every other week dosing (6%) than those receiving weekly dosing (4%). In addition, concomitant MTX therapy was associated with a considerably lower rate of antibody development than patients on adalimumab monotherapy (1% versus 12%).[7]

In view of this the manufacturer was asked to provide any additional data relating to the appearance of antibodies to adalimumab in the PsA trials. The company stated in confidence that

‘[redacted]

Overall adalimumab appears to be less immunogenic than infliximab.[7] However, direct comparisons between the incidences of antibody formation to adalimumab with that of other TNF agents is difficult due to confounding factors such as lack of assay standardisation, concomitant medications, and underlying disease.

Chapter 2
Background

2.1 Adalimumab for the treatment of moderate to severe psoriatic arthritis

Psoriatic arthritis (PsA) is a progressive inflammatory arthritic condition occurring closely but not exclusively in co-existence with a history, or in the presence, of psoriasis. Sometimes the arthritis will precede psoriasis.[8] The association with psoriasis, as the name suggests, is strong – about 80% of new psoriatic arthritis patients will have a clinical history of psoriasis,[9] meaning that about 20% will first present with arthritis or with both aspects simultaneously.[10] The condition is characterised by latent periods with exacerbations of psoriasis and/or arthritis due to unknown trigger factors. Simultaneous exacerbations or 'flare-ups' of skin and joint features occur in about 35% of patients.[11] Various trigger factors have been postulated including stress, trauma, and infection.[12] Onset of arthritis is usually insidious but may be acute in about one third of patients. The peak age of onset is around 40 years, in common with RA.[13]

A number of distinct clinical features allow PsA to be distinguished from other forms of inflammatory arthritis.[14] The pattern of joint involvement is characteristic with distal interphalangeal joint involvement, asymmetry, dactylitis (inflammation of the fingers), flail-like or ankylotic deformed digits, enthesitis (inflammation of bone-tendon connections) and spinal involvement. Other features distinguishing the condition from RA are a less-intense violet/purple colouration of the skin of affected joints, a 'ray' distribution in which all of the joints of a single digit are affected, and less tender more fibrous joints.[15] It may present with a pattern of articular involvement similar to that seen with RA, but in these cases, rheumatoid factor and other systemic features of RA are usually absent.[16] Additionally, PsA affects both sexes equally, whereas RA is more common in females.[14] None of these features is by itself sufficiently diagnostically specific to the condition, however, when all the factors present are considered together, diagnosis can be made by experienced clinicians.[9]

2.1.1 Prevalence

The prevalence of psoriasis among patients in the general population is 2-3% but among patients with arthritis it is 7%. Inflammatory arthritis occurs in 2-3% of the general population but among patients with psoriasis this increases to between 6% and 42%.[14] Due to the lack of a precise definition or diagnostic marker for PsA the exact prevalence is unknown but estimates range from 0.3% to 1.0%.[14]

2.1.2 Diagnosis

There are no specific diagnostic serological or radiological tests for PsA, rather the whole clinical picture must be considered along with any objective clinical results. However, the following are often used to aid the diagnosis: acute phase reactants such as erythrocyte sedimentation rate (ESR) or C-reactive protein (CRP) may be elevated, and there may be anaemia of chronic disease. Rheumatoid factor is usually absent, but its presence does not exclude a diagnosis of PsA.[9, 12] Radiological examinations often reveal ankylosis, osteolysis of phalangeal bones, periostitis, absence of periarticular osteoporosis, and new bone formation within enthesis (e.g. bone spur).[9, 12] As spinal involvement is common in PsA, a radiological examination may also reveal random asymmetrical involvement, apparent new bone formation, cervical spine disease, and sacroilitis which may be asymmetrical.[12]

Despite some distinct differences between PsA and RA, in practice the condition can be difficult to distinguish from RA and some other conditions such as osteoarthritis, gout, and Reiter's syndrome.[12] It should be acknowledged that psoriasis patients without arthritis are still at risk of developing non-psoriatic arthritis such as osteo- or rheumatoid arthritis, and this is another confounding factor in diagnosing the condition. Diagnosis should be confirmed by a specialist.

Historically the Moll and Wright criteria have been used as a framework for both classification and diagnosis of PsA.[13, 17] Essentially they require the presence of an inflammatory arthritis (peripheral, sacroiliitis or spondylitis), the presence of psoriasis, and the (usually) absence of serological markers for RA.[17] A new classification system for PsA based on joint involvement has recently been developed by the Classification of Psoriatic Arthritis study group (CASPAR)[18] consisting of established inflammatory articular disease with at least three points from the following features: current psoriasis (assigned a score of 2; all other features were assigned a score of 1), a history of psoriasis (unless current psoriasis was present), a family history of psoriasis (unless current psoriasis was present or there was a history of psoriasis), dactylitis, juxta-articular new bone formation, rheumatoid factor negativity, and nail dystrophy.[18]

2.1.3 Prognosis

About 20% of patients with PsA develop a destructive disabling form of arthritis. After 10 years, 55% of patients have five or more deformed joints. Bone and joint erosion has a high rate of occurrence with 47% demonstrating ≥ 1 erosion within two years of onset.[14] About 5% of PsA patients will exhibit a highly deforming and destructive form of PsA called arthritis mutilans. Joint deformity can continue to progress despite reduced inflammation.[10, 12] PsA is

associated with an increased risk of death at a ratio of 1.62, although the cause of death is generally in line with that of the general population.[14] PsA patients have reduced HRQoL and functional capacity compared with psoriasis patients or healthy controls, with an overall impact similar to that seen with RA.[14] In general, PsA is considered to have a less severe course than that seen in RA. There is some evidence that radiological changes progress at a faster rate than clinical symptomatic expression.[10]

2.1.4 Treatment

PsA is a complex and multifaceted disease with prominent involvement of the skin and joints. Despite the lack of correlation in onset and severity of the two aspects of the disease, the effective treatment of PsA should target both the skin and joint manifestations. Patients with psoriasis and PsA should be managed by collaboration between dermatologists and rheumatologists.[19]

A number of different therapies have been adopted for the treatment of PsA providing differing levels of symptomatic relief and impact on the progression of permanent joint damage. NSAIDs are commonly used as initial therapy for patients with mild PsA.[19, 20] However, NSAIDs occasionally cause a worsening of psoriasis,[15] and long-term use may be associated with gastrointestinal, renal and hepatic toxicity.[21] Corticosteroids may be used when only one or two persistently actively inflamed joints are involved, either intra-articularly or systemically via the oral route.[20] This latter option is generally avoided due to potential for provoking a pustular flare in psoriasis upon withdrawal. Although both NSAIDs and corticosteroids may provide symptomatic relief in PsA, neither has been shown to affect the progression of the disease.[15, 19] For extensive or severe PsA that is unresponsive to anti-inflammatory therapy traditional DMARDs, commonly used in the treatment of RA have been used with varying efficacy.[22] As with RA there are no prognostic factors that will identify which patient will respond to a particular therapy.

The efficacy of DMARDs in the treatment of PsA was reviewed in the technology appraisal for etanercept and infliximab.[23] The available DMARD treatments for PsA, with the exception of sulphasalazine and possibly leflunomide, have not been investigated thoroughly. The available limited data indicate some degree of efficacy for all DMARDs. Currently, methotrexate and sulphasalazine are considered the DMARDs of choice, although the evidence for the use of methotrexate in PsA is still largely empirical, while the clinical benefit induced by sulphasalazine appears to be modest.[22] Other treatment options include gold salts, leflunomide, azathioprine, ciclosporin, hydroxychloroquine, penicillamine, colchicine,

and psoralen-UVA photochemotherapy. Amongst DMARDs, only leflunomide is specifically licensed for the treatment of PsA.

Drugs such as methotrexate and sulfasalazine may substantially help the arthritic component of the condition, but not always the skin manifestations and agents such as ciclosporin may benefit the skin but not adequately help the joints. However, none of the traditional DMARD therapies has a significant beneficial effect on the spine.[24] Newer strategies for the treatment of PsA have focused on reducing the various pro-inflammatory chemokines and cytokines that have been implicated in the pathogenesis of both psoriasis and psoriatic arthritis. Because of the central pro-inflammatory role played by TNF in autoimmune diseases such as RA several anti-TNF agents have been investigated for the treatment of various diseases, including PsA. Three TNF-alpha antagonists (etanercept, adalimumab and infliximab) have been licensed for the treatment of adults with active PsA who have had an inadequate response to DMARDs.

2.1.5 NICE Guidance[25]

The National Institute for Health and Clinical Excellence (NICE) published guidance on etanercept and infliximab for the treatment of PsA in July 2006. NICE Technology Appraisal Guidance No 104 made the following recommendations for the treatment of adults with severe active PsA:

1) Etanercept, within its licensed indications, is recommended only when the following criteria are met.
    - The person has peripheral arthritis with three or more tender joints and three or more swollen joints.
    - The psoriatic arthritis has not responded to adequate trials of at least two standard DMARDs, administered either individually or in combination.

2) Infliximab, within its licensed indications, is recommended if, under the circumstances outlined in 1) above, treatment with an anti-TNF agent is considered appropriate and the person has been shown to be intolerant of, or have contraindications to, treatment with etanercept or has major difficulties with self administered injections.

3) Etanercept and infliximab treatment should be discontinued in patients whose PsA has not shown an adequate response when assessed using the PsA Response Criteria (PsARC) at 12 weeks. An adequate response is defined as:

- An improvement in at least two of the four PsARC criteria, one of which has to be joint tenderness or swelling score, with no worsening in any of the four criteria.

2.1.6 Adalimumab

Adalimumab is a fully humanised monoclonal IgG1 antibody and TNF antagonist. Recent evidence has indicated the presence of raised concentrations of TNF in psoriatic skin and synovial tissue and fluid.[2] The mechanism whereby adalimumab exerts its anti-TNF action involves direct binding of TNF molecules, subsequently preventing TNF from binding to the p55 and p75 cell surface receptors. Adalimumab also modulates biological responses that are induced or regulated by TNF, including changes in the levels of adhesion molecules responsible for leukocyte migration (ELAM-1, VCAM-1, and ICAM-1). By inhibiting the release of proinflammatory cytokines, adalimumab and other TNF antagonists, aim to reduce the degree of joint damage exhibited in inflammatory arthritis conditions.[2]

Adalimumab solution is available commercially as Humira® (Abbott Laboratories Ltd) in a 40mg pre-filled syringe and pre-filled pen.[2] It is licensed for the treatment of active and progressive PsA in adults when the response to previous DMARDs has been inadequate.[2] Administration of adalimumab involves a single subcutaneous injection every two weeks which may be given by the patient or carer after proper training in injection technique. In patients who respond clinically to adalimumab, response is usually achieved within 12 weeks of treatment. The British Society for Rheumatology (BSR) guidelines for anti-TNF-alpha therapy in PsA state that response at 12 weeks should be assessed for all patients and that adalimumab should be withdrawn in those patients who are unresponsive to treatment.[26]

2.2 Critique of the manufacturer's description of the background

The Abbott submission provided a comprehensive and detailed background. The disease and current treatment options were discussed in detail. The rationale for the development of the technology and its proposed place in therapy were clearly defined. The description of the technology under assessment was detailed and appropriate and covered all the relevant aspects. Overall this section appears a fair and accurate summary, confirmed by an extensive literature review of the subject. Several factual statements are not referenced but most of these are not contentious. However, one statement made whilst describing the factors which have limited the development of a uniform approach to the treatment of PsA (page 11, paragraph 3 - 'a reluctance to use DMARDs') is not supported by the available literature or clinical experts.[15, 22, 27] The statement linking an increased mortality rate among PsA patients to disease progression and cardiovascular complications is incorrect. The

authors of the referenced article[28] conclude that the leading causes of death in their PsA patients were similar to those in both the general population and patients with RA but with a significantly greater than expected incidence of respiratory-associated mortality.[28] With respect to monoclonal antibody therapies, the annual number of injections for each drug is stated for adalimumab and etanercept, but not for infliximab, which is associated with the least frequent injection schedule.

Chapter 3
Defining the Decision Problem

3.1 Scope

The scope for this single technology appraisal (STA) was clearly defined in the Abbott Laboratories Limited submission.[1] The decision problem considered was the clinical and cost-effectiveness of adalimumab for the treatment of active and progressive PsA in adults who have responded inadequately to previous DMARD therapy.

3.2 Intervention

The intervention considered in the decision problem was adalimumab (Humira®) administered as a single dose of 40 mg via subcutaneous injection every other week.

Adalimumab is manufactured by Abbott Laboratories Ltd. The list price of (£357.50 for one 40mg prefilled syringe) is correct at the time of writing.[29]

3.3 Patient population

The manufacturer stated that the patient population considered in the decision problem should be in accordance with the licensed indication for adalimumab. Adalimumab has a marketing authorisation for the treatment of active and progressive psoriatic arthritis in adults when the response to previous DMARD therapy has been inadequate.[2] The participants in the pivotal RCTs, however, were not entirely representative of the population for which adalimumab is currently licensed, as neither population was made up exclusively of patients who had failed to respond to at least two DMARDs (43% & 40% in M02-518, and 61% & 51% in M02-570 for the placebo and adalimumab groups, respectively). However, independent expert clinical advice given to the ERG suggests that the participants in these trials nevertheless represented a population with relatively severe PsA similar to those currently being treated in UK clinical practice.[27]

3.4 Comparators

In accordance with the licensed indication,[2] the comparators chosen by the manufacturer were conventional management strategies for active and progressive PsA that has responded inadequately to previous DMARD therapy, excluding adalimumab, but including other available biologic therapies such as etanercept and infliximab. Because there are no head-to-head studies comparing the different anti-TNF agents, indirect treatment comparison methods were deemed appropriate.

3.5 Trial Outcomes

Assessment of the effectiveness of treatments for PsA relies on there being outcome measures that enable disease activity to be evaluated accurately and sensitively. The outcome measures used in PsA have generally been adapted from similar measures used in the assessment of RA and psoriasis.[24] A variety of domains are assessed including peripheral joint and skin symptoms, function, quality of life, fatigue and imaging of structural damage. Approaches to assessing enthesitis, dactylitis and spinal involvement are in development. Improved outcome measures are being developed and validated specifically for PsA by both rheumatologists and dermatologists through the Group for Research and Assessment of Psoriasis and Psoriatic arthritis (GRAPPA).[24, 30]

The evidence submitted considered a range of outcomes to assess the impact of adalimumab treatment on both the joint and skin components of PsA (Table 3.1.).

Table 3.1: Summary of common assessment tools used in studies of Psoriatic Arthritis

Name	Principal dimension	Description	Notes
Psoriatic arthritis response criteria **(PsARC)**	Peripheral joint disease activity	Positive treatment response defined as improvement in joint pain/tenderness score or joint swelling score plus improvement in at least one of patient or physician assessment, with no worsening in any domain. Cut-off point for effect with respect to joint scores is ≥ ±30%.	Developed specifically for use in a clinical study of sulphasalazine in PsA. Does not assess dermatological involvement. British Society for Rheumatology guidelines recommend the PsARC to assess the primary joint response to anti-TNF therapy in psoriatic arthritis.[26]
American College of Rheumatology joint count **(ACR)**	Peripheral joint disease activity	Different cut-off points are used, the most common being 20% known as the ACR20. Other common points include the ACR50 and ACR70 relating to 50% and 70% respectively. The criteria require, for example, ≥ 20% reduction separately in tender and swollen joint count, and a ≥ 20% reduction in 3 of the 5 following domains: patient assessment of pain, patient assessment of disease activity, physician assessment of disease activity, the health-assessment questionnaire disability index, and presence of acute-phase reactants.	Originally developed for RA. When used to assess PsA a modified version of the test, incorporating the distal interphalangeal joints must be used. This has been demonstrated to be a reliable measure of activity in PsA.[4] The ACR20 is generally accepted to be the minimal clinically important difference indicating a response. Does not assess dermatological involvement.
Psoriasis area and severity index **(PASI)**	Psoriatic skin disease	The total PASI score (range 0 to 72) is made up of the individual weighted scores for four body areas: head (10%), trunk (30%), and upper (20%) and lower (40%) extremities. The score for each area is made up of an extent score (range 0 to 6) and severity score (range 0-12) and the severity score is in turn made up of a symptom score (range 0 to 4) for three symptoms: erythema, induration, and desquamation. A score of ≥ 50% (PASI 50, i.e. ≥ 50% reduction in PASI score) is considered the minimum significant response and a score of ≥ 75% is a common target in psoriasis studies.	Requires a minimum involvement of 3% of the body surface area. Scores > 36 are uncommon, even in severe psoriasis. Likely to be insensitive to changes from mild to moderate psoriasis. Assessment of disease extent (i.e. percentage skin involvement) is prone to inaccuracy. Poorly defined parameters. The British Society for Rheumatology guidelines for anti-TNF therapy in PsA recommends using PASI 75 to assess the primary response of psoriasis.[26]

Physician Global Assessment **(PGA)**	Psoriatic skin disease	The PGA is simple and quick to use. It is described as either static (measured at that point in time) or dynamic (measured with reference to the patients baseline condition). The static measure is most commonly used. It is the standard measure of psoriatic disease severity in current clinical practice. Usually scored on a discrete seven-point scale with 0 = clear, and 1-6 = increasing severity. On the dynamic PGA 1-5 = increasing severity and 6 = worsened.	Highly subjective measure. Changes between scores may not be linear.
Health Assessment Questionnaire **(HAQ)**	Generic patient-assessed health outcome measure	Rheumatology specific generic health survey with five domains: death, disability, drug side-effects, discomfort/pain, and economic costs. Common sub-measures are the HAQ disability index (HAQ-DI) and the HAQ pain scale. The HAQ-DI has eight domains: dressing, rising, eating, walking, hygiene, reach, grip and usual activities. A short version is often used comprising of the HAQ-DI, a visual-analogue pain scale and visual-analogue global health scale, is often used.	The HAQ-DI is a composite component of the ACR joint count score. The HAQ-DI has a range from 0 to 3, with 25 possible values at 0.125 intervals. Higher scores represent greater disability. It has been validated in RA.
Short-form, 36-question health survey **(SF-36)**	Generic patient-assessed health outcome measure	Standard, widely used, generic health survey consisting of 36 questions from eight domains: physical function, social function, role limitations due to physical problems and emotional problems, mental health, energy/vitality, pain, and general health perception. The result is an 8-scale health profile as well as summary measures of HRQoL often divided into two domains: physical component score and mental component score.	The SF-36 has been extensively used in medical research. It has been tested in PsA and found to be reliable, valid and responsive to change.[4]
Total Sharp Score **(TSS)**	Radiological assessment of disease progression	This method of assessing radiological progression was developed for use in RA. All joints of the hands are graded separately for erosions (score 0-5) and joint space narrowing (score 0-4) with a maximum score of 149. A modified version of this score which incorporates the distal interphalangeal joints and metatarsophalangeal joints of the feet and interphalangeal joint of the first toe has been proposed for use in PsA.	Neither the TSS nor the modified TSS have been validated for use in PsA. Due to the difficult and unpredictable pattern of joint damage in PsA, it remains to be determined which joints should be scored to get a valid measurement of joint damage.[4]

Chapter 4
Clinical Effectiveness

4.1 Search strategy

A systematic literature search was undertaken by the ERG to verify the completeness of the methodology used by the manufacturer to retrieve relevant clinical studies presented in the submission. The inclusion and exclusion criteria and the search strategy used by the ERG are presented in Appendix 1.

Other than several abstracts derived from the two aforementioned RCTs included in the submission no other relevant clinical studies were found during the literature search. Data from these abstracts were included only if they provided additional evidence with respect to the decision problem.

4.2 Evidence of clinical efficacy

The manufacturer included two randomised controlled trials and a meta-analysis thereof, and two non-randomised controlled trials in the submission (summarised in Tables 4.1-4.6). An indirect/mixed treatment comparison was also considered as evidence of the clinical efficacy of adalimumab in PsA.

Table 4.1: Summary of trials included in the manufacturer's submission

Trial	Interventions		Key Issues
M02-518 (ADEPT)[31]	Adalimumab 40 mg SC eow	Placebo SC eow	• Fully published. • Patients had inadequate response to NSAIDs rather than DMARDs. • ACR 20/50/70, PsARC and PASI 50/75/90 responses at 12 and 24 weeks were significantly greater with adalimumab vs. placebo. • ACR 20/50 and 70 response rates did not differ between patients taking adalimumab alone or concomitantly with methotrexate. • Mean improvement in HAQ, FACIT and DLQI significantly better with adalimumab vs. placebo.
M02-570[32]	Adalimumab 40 mg SC eow	Placebo SC eow	• Abstract. • Patients had inadequate response to DMARDs. • ACR 20/50/70 [redacted] responses at 12 weeks were significantly greater with adalimumab vs. placebo. • PGA and mean improvement in TLS, HAQ and SF-36 scores significantly better with adalimumab vs. placebo.

M02-537[32-34] (open-label extension of M02-518 & M02-570)	Adalimumab 40 mg subcutaneously (SC) every other week (eow) (Adalimumab 40 mg SC weekly was permitted in subjects who failed to respond after at least 12 weeks).	• Abstract. • On-going, >120 week open label trial. • Patients had completed M02-518 or M02-570. • Improvements in ACR 20/50/70 and PASI 50/75/90 responses appeared to be maintained up to week 88. • ACR 20/50/70 and PASI 50/75/90 response rates did not differ between patients taking adalimumab alone or concomitantly with methotrexate.
M04-724 (STEREO)[35]	Adalimumab 40 mg SC eow	• Abstract. • On-going, prospective 12-week open label trial. • Adalimumab added to existing therapy in clinical practice. • ACR20 response achieved in 72% of patients by week 12. • Unclear how many of the enrolled patients are being treated concomitantly with DMARDs or NSAIDs.

Summary of clinical trials

Abbreviations key: ACR20/50/70 – American College of Rheumatology 20%/50%/70% Improvement Response*;* ADA - adalimumab; AE – adverse effects; BSA – body surface area; CI – confidence interval; DAS 28 – disease activity score; DB – double blind; DLQI – dermatology life quality index; DMARD – disease-modifying anti-rheumatic drug; eow – every other week; HAQ – health assessment questionnaire; HIV – human immunodeficiency virus; IBD – inflammatory bowel disease; MC – multicentre; mTSS – modified total sharp score; MTX – methotrexate; NSAID – non-steroidal anti-inflammatory drugs; NSD – no significant difference; PbO – placebo; PASI50/75/90 – the psoriasis area and severity scale, 50%, 75% and 90% improvement; PGA – physicians global assessment; PsA – psoriatic arthritis; PsARC – psoriatic arthritis response criteria; RCT – randomised controlled trial; SAE – serious adverse event; SC - subcutaneously; SJC – swollen joint count; TJC – tender joint count; URTI – upper respiratory tract infection; wk - week.

Table 4.2: Summary of trial: M02-518 (ADEPT)[31]

Reference	Design	Intervention	Inclusion criteria	Exclusion Criteria	Outcomes	Results	Adverse Effects
Mease P, Gladman D, Ritchlin C et al al. 2005 (ADEPT)[31]	24-Wk phase III, RCT, PbO, DB, stratified according to MTX use (yes/no) and extent of psoriasis (<3% and ≥3% BSA) in patients who had failed NSAID therapy	Patients were randomised to SC injections of either PbO or ADA 40mg eow	Patients were included if they were ≥ 18 years old, had moderate to severe PsA and either active psoriatic skin lesions or documented history of psoriasis. They needed to have a history of inadequate response or intolerance to NSAID therapy for PsA. MTX use was only allowed if it had been taken for at least 3 months previously with the dosage stable for at least 4 weeks prior to baseline visit.	Treatment within 4 weeks of baseline visit with ciclosporin, tacrolimus, DMARDs other than MTX, or oral retinoids; Topical treatments for psoriasis within 2 weeks of baseline, other than shampoos or low potency steroids; Concurrent treatment with MTX at dosages > 30mg/week and/or corticosteroids > 10mg/day; Anti-TNF therapy at any time.	Primary efficacy endpoints were ACR20 at week 12 and the change in mTSS of structural damage at week 24. Secondary endpoints included the ACR20 at week 24, and ACR50 and ACR70 response rates at weeks 12 and 24, modified PsARC response rates, the disability index of the HAQ and the SF-36 health survey at weeks 12 and 24. Dactylitis and enthesitis were assessed. For patients with psoriasis involving at least 3% BSA further endpoints at weeks 12 and 24 were PASI50 and PASI75, the PGA and the DLQI was completed by patients.	ACR20 response at week 12 was 58% for the ADA group and 14% for the PbO group (between-group difference 44%, 95% CI 33-54%, p<0.001). The mean change in mTSS at week 24 was -0.2 for ADA patients and 1.0 for PbO patients (p<0.001). ACR20 response rates at 24 weeks 57% and 15% in the ADA and PbO groups respectively (between-group difference 42%, 95% CI 31-52%, p<0.001). The PASI75 response at 24 weeks was 59% in the ADA group and 1% in the PbO group (n=69 per group, p<0.001). Mean change in HAQ -0.4 for ADA vs. -0.1 for PbO at week 24 (p<0.001)	Most AEs were similar in the PbO and ADA groups. The most common were URTI (14.8% vs. 12.6%), nasopharyngitis (9.3% vs. 9.9%), injection site reactions (3.1% vs. 6.6%) and headache (8.6% vs. 6.0%). Seven PbO and five ADA patients experienced serious AEs, four patients (three in the ADA group) prematurely discontinued treatment due to AEs. Two ADA patients additionally discontinued treatment due to abnormal laboratory results.

Table 4.3: Summary of trial: M02-570[32]

Reference	Design	Intervention	Inclusion criteria	Exclusion Criteria	Outcomes	Results	Adverse Effects
Genovese M, Mease P, Thomson G et al 2005. Proceedings from the Annual European Congress of Rheumatology (EULAR) 2005, Vienna, Austria.	12-Wk phase III, RCT, PbO, DB, MC, stratified according to DMARD use (yes/no).	ADA 40 mg SC eow (n = 51) or PbO SC eow (n = 49) for 12 wks.	≥ 18 years old, moderate to severely active PsA (≥ 3 swollen joints & ≥ 3 tender or painful joints), inadequate response to DMARD therapy, active cutaneous chronic plaque psoriasis lesions present or documented history of chronic plaque psoriasis.	Prior anti-TNF therapy, the following therapies before baseline: cyclosporin or tacrolimus within 4 wks; systemic psoriasis therapy within 4 wks; alefacept or siplizumab within 12 wks; other biological or investigational therapy within 6 wks; phototherapy or topicals within 2 wks.	Primary endpoint: ACR20 response at wk 12. Secondary endpoints: HAQ disability index, and target lesion evaluation and PGA for psoriasis in patients with a psoriasis target lesion (ADA n = 32; PbO n = 30). ACR50 and ACR70responses also reported.	**Results at wk 12** ACR20 response: ADA 39% of patients; PbO 16% (p = 0.05). Mean change in HAQ from baseline: ADA -0.3; PbO -0.1 (p = 0.01). PGA clear or almost clear: ADA 40.6%; PbO 6.7% (p = 0.01). Mean % change in Target Lesion Score: greater with ADA (-47.0%) than with PbO (-1.6%, p = 0.001). ACR50: ADA 25%; PbO 2% (p = 0.001). ACR70: ADA 14%; PbO 0% (p = 0.05).	Number of patients with any AE occurring in ≥ 5% of patients lower with ADA (27/52; 52.9%) than PbO (39/49; 79.6%; p ≤ 0.01). One (2.0%) and 2 (4.1%) serious AE with ADA and PbO, respectively. Psoriasis and aggravated PsA occurred in more patients on PbO than ADA, respectively (16.3% vs. 3.9% and 14.3% vs. 2.0%; p ≤ 0.05 for both). NSD between groups for incidences of other AEs (URTI, injection site pain, diarrhoea, back pain, headache).

Table 4.4: Summary of trial: M02-537 (Open-label extension)[32-34]

Reference	Design	Intervention	Inclusion criteria	Exclusion Criteria	Outcomes	Results	Adverse Effects
Mease P, Gladman D, Ritchlin C et al. 2005. (ACR, Annual Scientific Meeting, 2005).	24-Wk open-label extension of the ADEPT trial.	ADA 40 mg SC eow for 24 wks, after receiving ADA or PbO for 24 wks (n = 285).	Patients who completed the ADEPT trial.	As for the ADEPT trial.	Outcomes reported at 48 wks (i.e. after 24 wks of open-label therapy) were: ACR20/50/70 responses; improvement in HAQ score; PASI 50/75/90 responses (in patients with psoriasis ≥ 3% BSA); % change in PASI; PGA of clear or almost clear. ACR 20/50/70 and PASI 50/75/90 response rates in patients using and not using MTX who received ADA for 48 wks were compared.	Results at wk 48 272 (95%) completed open-label wks 24 – 48. At 36 wks, 30 patients increased ADA dose to 40 mg every week. ACR20/50/70 responses: ADA 61%/46%/31% of patients (c.f. 57%/39%/23% at wk 24). PASI 50/75/90: ADA 70%/58%/46% of patients (c.f. 75%/59%/42% at wk 24; n = 69). There was NSD between MTX users and non-users with respect to ACR and PASI response rates. HAQ mean change from baseline: ADA -0.4 (c.f. -0.4 at wk 24). Mean % change in PASI: ADA -67% (c.f. -66% at wk 24). PGA clear/almost clear: ADA 63% (c.f. 67% at wk 24; n = 70). PbO patients achieved similar responses in the open-label period after 24 wks of ADA	AE profile after ADA for 48 wks similar to that after 24 wks in the controlled phase of the ADEPT trial.
Mease PJ, Sharp JT, Ory P et al 2005. (EULAR, Annual Scientific Meeting, 2005).	As above but different outcome measure (radiological changes of hands and feet)	As above	As above	As above	Radiographs of hands and feet assessed by mTSS at 48 wks (i.e. after 24 wks of open-label therapy).	Evaluable films at wk 48: ADA n = 128; PbO/ADA n = 134. Mean change in mTSS at wk 48: ADA 0.1 (c.f. -0.1 at wk 24); PbO/ADA 1.0 (c.f. 0.9 at wk 24).	Not reported

Table 4.5: continued M02-537 (Open-label extension)[32-34]

Reference	Design	Intervention	Inclusion criteria	Exclusion Criteria	Outcomes	Results	Adverse Effects
Genovese MC, Mease PJ, Thomson GTD et al 2005 (presented at EULAR Annual Meeting, 2005).	12-Wk open-label extension of the M02-570 trial	As above (n = 97; 46/49 of original PbO group and all 51 of original ADA group).	Patients who completed 12 wks of therapy in MO2-570 trial.	As for MO2-570 trial.	ACR20/50/70, mean change in HAQ from baseline, PGA clear or almost clear, mean % change in target lesion score at wk 24 (i.e. after 12 wks of open-label therapy)	Results at wk 24 92/97 patients (97%) completed open-label wks 12 – 24. ACR20/50/70 responses: ADA 64%/43%/27% (c.f. 39%/25%/14% at wk 12). Mean change in HAQ from baseline: ADA -0.3 (c.f. -0.3 at wk 12) PGA clear or almost clear: ADA 56.3% (c.f. 40.6% at wk 12). Mean % change in Target Lesion Score: ADA -58.8% (c.f. -47.0% at wk 12). Responses of PbO/ADA group after 12 wks of open-label therapy increased to levels similar to those of the ADA group.	During open-label period, 2 additional AEs reported in ≥ 5% of all patients: cough and nasopharyngitis (both 5/97; 5.2%).

Table 4.6: Summary of trial: M04-724 (STEREO)[35]

Reference	Design	Intervention	Inclusion criteria	Exclusion Criteria	Outcomes	Results	Adverse Effects
Van den Bosch et al. 2006. Annual Scientific Meeting of the American College of Rheumatology (ACR). 2006, Washington, U.S.A (STEREO)[35]	Non-RCT, prospective 12 week open label study.	All patients received ADA 40mg SC eow	Patients were included if they were ≥ 18 years old, had active PsA with ≥3 tender joints, ≥3 swollen joints despite standard PsA therapy. They needed to have a history of inadequate response to at least one DMARD. Patients who had previously failed treatment with other anti-TNF agents were eligible.	History of cancer or lymphoproliferative disease, HIV, Hepatitis B or C. History of drug or alcohol abuse. History of co-morbidities (e.g. uncontrolled diabetes, unstable ischaemic heart disease, IBD, chronic leg ulcer)	Primary outcomes were ACR20, ACR50 and ACR70, and change in DAS28 at week 12. For skin symptoms PGA was measured at week 12. Other endpoints included responses to HAQ and DLQI. Safety was evaluated in terms of adverse events reported by the patients and routine safety evaluations at weeks 2, 6 and 12 were conducted. No distinction was made between primary and secondary outcomes.	72% of patients achieved an ACR20 response by week 12. (vs. 38% and 60% at week 2 and 6). 49% and 27% of patients achieved an ACR50 and ACR70 response by week 12 (vs. 14% and 2% at week 2 and 35% and 13% at week 6). At week 12 mean TJC was 7.3 compared to baseline of 17.6. At week 12 mean SJC was 2.3 compared to a baseline of 9.3. Mean DAS28 scores showed a decrease (2.6) compared to baseline (4.8) at week 12. (vs. 3.4 and 2.9 at weeks 2 and 6). Mean HAQ scores decreased over the 12 week period (1.20 to 0.86). The percentage of patients with a PGA of 'clear or almost clear' increased over the 12 week period, from 35% at baseline to 65% at week 12. (vs. 39% at week 2 and 53% at week 6). DLQI scores showed a decrease (6.4 at baseline down to 2.8.)	Preliminary safety data for all 441 patients is included to April 2006. ADA was well tolerated overall with only 19 patients (4%) experiencing SAE's Investigator defined SAE's included: abdominal pain, anaemia, dental abscess, urosepsis, fever with reduced general condition, allergic reaction, severe hip pain and hypersomnia. The spectrum of AE's was similar to that highlighted in earlier RCTs and withdrawal rates, so far, appear to below (8%).

4.3 Submission trial analysis

All studies included in the clinical evidence section of the Abbott submission were subjected to a detailed critical appraisal (see Appendix 2). Studies were appraised by one reviewer and independently checked by a second reviewer. Disagreements were resolved through consensus, consulting a third reviewer if necessary. The resultant appraisals were then compared to the data presented in the submission. Data from studies presented in multiple abstracts were extracted and reported as a single study with all other relevant publications listed.

4.3.1 Trial M02-518 (ADEPT)[31]

Trial summary

This study was designed to assess the safety and efficacy of adalimumab when compared to placebo in patients with active psoriatic arthritis (PsA). A structured critical appraisal of this trial is presented in Appendix 2.1.

This fully published trial showed that in the short term (24 weeks), adalimumab was statistically significantly superior to placebo in patients with active PsA with regards to ACR20 improvement and the change in score of structural damage on radiographs of hands and feet at 24 weeks.

Important trial points

Key trial points are outlined below:

- Eligible patients had moderately to severely active PsA ($\geq$ 3 swollen and $\geq$ 3 tender or painful joints) with either active psoriatic skin lesions or a documented history of psoriasis. Patients were also required to have a history of an inadequate response or intolerance to NSAID therapy (defined by the investigator).
- Only 42% of patients had previously received treatment with two or more DMARDs.
- The study was generally well conducted with adequate blinding of treatment allocation with all patients accounted for in the final results and no significant differences between treatment groups at baseline.
- Patients treated with adalimumab demonstrated significantly higher ACR 20 at week 12 than those treated with placebo. In the adalimumab group 58% achieved this primary endpoint compared with 14% of the placebo group ($p < 0.001$).
- ACR20, 50 and 70 response rates did not differ significantly between patients receiving adalimumab combined with methotrexate (response rates of 55%, 36% and 17% respectively) and those receiving adalimumab alone (61%, 36% and 23% respectively).

- The second primary end point measured modified total sharp score (mTSS) at 24 weeks. The mean change in the mTSS in patients with both baseline and week 24 radiographs was -0.2 for adalimumab patients compared with 1.0 for placebo patients ($p<0.001$). However, the study duration does not reach the recommended 12 to 24 months required to assess radiological deterioration in PsA.
- The PASI 75 response rate at 24 weeks was 59% in the adalimumab group and 1% in the placebo group ($n = 69$ per group, $p < 0.001$).
- Upon completion of 24 weeks therapy, patients were eligible to enter the open-label extension study (M02-537).

Critique of the Abbott submission

A significant proportion of the data included in the submission was not presented in the published study.

- Section 4.4 refers to the improvement in PASI score stating responses were 'significant at week 12 and were maintained through to week 48'. The published trial only reported data to 24 weeks.
- Section 4.5 bullet point four details the difference in response rate between patients given adalimumab alone, and in combination with methotrexate. The published clinical trial did not report data for this sub-analysis.
- Table 5.4.2 reports results of the PGA at 24 weeks. The figures quoted in the submission do not agree with those stated in the published clinical trial.
- In the published clinical trial the reported data on adverse events was minimal. Table 5.7.1 in the submission was far more detailed. However the data relating to 'any AE leading to discontinuation of study drug' was inconsistent (submission: placebo = 5 patients, adalimumab = 6 patients; published clinical trial: placebo = 1 patient, adalimumab = 3 patients)
- The submission reported additional information on the 12 and 24 week responses stratified by previous NSAID use (Table 5.9.2.1), improvements in the specific domains of the SF-36 score (Table 5.4.5 and 5.4.6), statistical analysis for the modified total sharp score, change in sharp score components at week 24 (Table 5.3.1.1), proportion of subjects with no change in modified sharp score (Table 5.3.4.1) and sensitivity analysis of the primary end points. The supplemental information provided by the manufacturer in the submission is not in the public domain and therefore cannot be validated externally.

Additional information requested from Abbott

The inclusion criteria defined in this study related to inadequate response to NSAID therapy. Patients were included if they had experienced an inadequate response to DMARDs although these data were only available in the submission not the published clinical trial. Data detailing the number of patients having had an inadequate response to two or more DMARDs were crucial in determining the proportion of patients who would fit the BSR criteria for treatment with anti-TNF therapy.

The following information (Table 4.7) was supplied by Abbott in response to this request.

Table 4.7: Number of different types of previous DMARDs – M02-518

Number of different types of previous DMARDs	Placebo (n=162) N (%)	Adalimumab (n=151) N (%)
0	[redacted]	[redacted]
1	[redacted]	[redacted]
2	[redacted]	[redacted]
3	[redacted]	[redacted]
4	[redacted]	[redacted]
5	[redacted]	[redacted]
6	[redacted]	[redacted]
Any Previous DMARD	[redacted]	[redacted]

(Academic in confidence)

In order to be enrolled in this trial, subjects were required to have active PsA and were allowed to continue prior treatment with methotrexate therapy (≤ 30 mg weekly). It was not clear from the information supplied whether or not this on-going treatment with methotrexate was classed as a failure to respond to methotrexate therapy. The ERG group asked Abbott to clarify this. The information received from Abbott stated 'The entry criteria for M02-518 stipulated that subjects must have active psoriatic arthritis. As such, it can be assumed that patients participating in study M02-518 taking concomitant methotrexate have failed this DMARD as they have active disease'.

In the original submission it was stated that after week 12, patients who failed to have at least a 20% decrease in both swollen and tender joint counts on two consecutive visits could receive 'rescue therapy' with corticosteroids or DMARDs. No further details were supplied regarding the number of patients who received 'rescue therapy'. On request by the ERG Abbott clarified that six placebo-treated patients and two adalimumab-treated patients received 'rescue therapy'. The ERG felt this was unlikely to introduce bias to the intention-to-treat analysis due to the small number of patients involved.

Summary

The data presented in this trial show that, in the short term (24 weeks), adalimumab was superior to placebo with regard to arthritis response measured by ACR 20/50/70. Adalimumab also showed greater improvements in the signs and symptoms of psoriasis compared with placebo, measured using PASI response. However, only 42% of adalimumab patients had previously received treatment with at least two or more DMARDs. Therefore, the patient population in this study may not be fully representative of the UK population for whom, according to current guidelines, anti-TNF therapy such as adalimumab would be considered.

This trial does not indicate how adalimumab performs in the longer term, or how it compares with other anti-TNF drugs in the same class, or even how it compares with DMARDs.

4.3.2 Trial M02-570[32]

Trial summary

This trial was designed to evaluate the efficacy and safety of adalimumab compared with placebo in patients with moderately to severely active PsA with an inadequate response to DMARD therapy. A structured critical appraisal of this trial is presented in Appendix 2.2.

This study, published in abstract form only, showed that in a small (n=100) 12-week trial adalimumab was more effective in terms of the ACR20 response than placebo. Several secondary endpoints, including effects on disability, the skin component of psoriasis and ACR50 and ACR70 responses, also suggested that adalimumab was more effective than placebo.

Since this study has not been fully published and the results are reported only in abstract form there are insufficient data presented to fully assess the quality and validity of this study. Although the abstract data are supplemented with additional information provided by the manufacturer in the submission, this supplemental information is not in the public domain and therefore cannot be validated externally.

Important trial points

Key trial points are outlined below:

- Published in abstract form only.
- Submission included additional unpublished data.

- Eligible patients had moderately to severely active PsA (defined as ≥ 3 swollen joints and ≥ 3 tender joints) and an inadequate response to DMARD therapy based on current or historic DMARD treatment.
- Only 56% of adalimumab patients had previously received treatment with two or more DMARDs
- ACR20 response at 12 weeks was significantly greater with adalimumab compared with placebo (39% vs. 16%, p=0.05). ACR50 and ACR70 response rates were also significantly greater with adalimumab (25% vs. 2%, (p=0.001), and 14% vs. 0% (p=0.05), respectively).
- [redacted] (Reported in the submission only).
- HRQoL, assessed via mean change in HAQ significantly improved compared with placebo (-0.3 vs. -0.1 on a 0-3 scale).
- In subjects with a psoriasis target lesion, adalimumab treatment had a significant effect on target lesion score (TLS) compared with placebo (mean change from baseline of -47% vs. -1.6%, p=0.001), and significantly more patients receiving adalimumab were assessed as clear or almost clear on the PGA of disease activity (PGA) (40.6% vs. 6.7%, p=0.01).
- Overall, significantly more adverse effects were reported in the placebo group than the adalimumab group (39 (79.6%) vs. 27 (52.9%); p ≤ 0.01, respectively).
- Adverse effects occurring in ≥ 5% of patients were similar in the placebo and adalimumab groups, with only aggravation of psoriasis and PsA occurring more frequently with placebo than adalimumab.
- Upon completion of 12 weeks therapy, patients were eligible to enter the open-label extension study (M02-537).

Critique of the Abbott submission

The submission was accurate according to the data in the published abstract, and appears to be a fair interpretation of the trial. Specific points are:

- The abstract data is supplemented with additional information regarding the study methodology. Outcome measures such as the PsARC and SF-36 are provided by manufacturer in the submission, but are not reported in the published abstract. This supplemental information, with particular respect to the additional outcome measures is not in the public domain and therefore cannot be validated externally.

- The submission states that subgroup analyses were undertaken for age, gender, race, site, PsA subtype and rheumatoid factor. However, no subgroup results are reported in the document.

Additional information requested from Abbott

The inclusion criteria defined in this study related to inadequate response to DMARD therapy. Data detailing the number of patients having had an inadequate response to two or more DMARDs was crucial to determining the proportion of patients who would fit the BSR criteria for treatment with anti-TNF therapy. The following information (Table 4.8) was supplied by Abbott in response to this request.

Table 4.8: Number of different types of previous DMARDs – M02-570

Number of different types of previous DMARDs	Placebo (n=[redacted]) N (%)	Adalimumab (n=[redacted]) N (%)
1	[redacted]	[redacted]
2	[redacted]	[redacted]
3	[redacted]	[redacted]
4	[redacted]	[redacted]
5	[redacted]	[redacted]
6	[redacted]	[redacted]
Any Previous DMARD	[redacted]	[redacted]

('Academic in confidence')

The submission also states that subjects were required to maintain baseline DMARD usage and dosage. However, adalimumab is currently only licensed as a monotherapy and not in combination with DMARDs for the treatment of PsA. Therefore, the ERG group asked Abbott to provide data regarding response parameters at 12 weeks based on DMARD use at baseline. The following information (Table 5.9) was supplied by Abbott in response to this request.

Table 4.9: Response rates at week 12, DMARD use at baseline – M02-570

DMARD use at Baseline	Yes		No	
Response	Placebo	Adalimumab	Placebo	Adalimumab
ACR20	[redacted]	[redacted]	[redacted]	[redacted]
ACR50	[redacted]	[redacted]	[redacted]	[redacted]
ACR70	[redacted]	[redacted]	[redacted]	[redacted]
PsARC	[redacted]	[redacted]	[redacted]	[redacted]
HAQ mean change from baseline ±SD	[redacted]	[redacted]	[redacted]	[redacted]

Summary

Results from this small study show that adalimumab recipients experienced a significantly greater improvement in arthritis response than those receiving placebo, as assessed by the ACR20/50/70 and PsARC response rates. The signs and symptoms of psoriasis in patients with PsA were also significantly improved with adalimumab compared with placebo, according to the TLS and PGA responses. However, only 56% of adalimumab patients had previously received treatment with two or more DMARDs. Therefore, the patient population in this study may not be fully representative of the UK population for whom, according to current guidelines, anti-TNF therapy such as adalimumab would be considered. Furthermore, until this study is fully published and the complete data are made available for evaluation these results should be interpreted with due caution.

4.3.3 Trial M02-537 (Open-label extension)[32-34]

Trial Summary

This trial was designed to evaluate the long-term efficacy and safety of adalimumab in patients with moderately to severely active PsA. Patients who completed 24 or 12 weeks of adalimumab or placebo in the ADEPT and M02-570 trials were eligible to enter this open-label extension study (M02-537), in which all patients received adalimumab 40 mg subcutaneously every other week. A structured critical appraisal of this trial is presented in Appendix 2.3.

This study published in abstract form only showed that ACR responses after adalimumab treatment for 24 and 12 weeks (ADEPT and M02-570 trials, respectively) appeared to be maintained after a further 24 and 12 weeks of treatment. Furthermore, ACR and PASI response rates did not differ significantly between patients taking adalimumab alone or concomitantly with methotrexate. The adverse effect profile after adalimumab treatment for 48 weeks was similar to that observed after 24 weeks.

Since this study has not been fully published and the results are reported only in abstract form there is insufficient data presented to fully assess the quality and validity of this study. Although the abstract data is supplemented with additional information provided by the manufacturer in the submission, this long-term supplemental information is not in the public domain and therefore cannot be validated externally. With the exception of responses in patients using and not using methotrexate at baseline, no statistical analyses of the open-label results are presented either in the manufacturer's submission or the three published abstracts. Therefore, the precision and robustness of the results cannot be assessed.

Important Trial Points

Key trial points are outlined below:

- Published in abstract form only.
- The submission included a large amount of additional unpublished data derived from a more recent analysis (full analysis set) of the data from the M02-537 trial.
- The open-label M02-537 trial included patients who completed the 24-week ADEPT and 12 –week M02-570 placebo-controlled trials.
- Eligible patients had moderately to severely active PsA and had an inadequate response to NSAIDs or DMARDs.
- In the published abstracts ACR responses after adalimumab treatment for 24 and 12 weeks (controlled phase) appeared to be maintained after a further 24 and 12 weeks of treatment (open-label phase). In the full analysis set, presented in the submission the percentage of patients achieving an ACR20/50/70 response over time showed a rapid increase and was maintained up to week 88.
- PASI responses are presented in the manufacturer's submission, but not in the abstracts. At week 12 PASI 50/75/90 responses were 73.8%, 53.3% and 33.6%, respectively and were maintained up to week 88 at 86.2%, 75.9% and 55.2%, respectively.
- ACR20 response rates by baseline DMARD use did not differ significantly.
- ACR and PASI response rates did not differ significantly between patients taking adalimumab alone or concomitantly with methotrexate. These finding are not presented in the submission.
- The mean change in mTSS from baseline to week 48 was -0.2 (outliers included)
- [redacted] (Reported as 'academic in confidence' in the submission only).
- The mean change in TLS was assessed in the M02-570 subjects only. At 24 weeks the mean percent change in the TLS was -58.8% compared with -47.0% at the end of the 24-week RCT.
- The percentage of subjects from both RCTs with a PGA of 'clear' or 'almost clear' was 51.6% after 12 weeks of adalimumab, and increased to 68.6 after 48 weeks. (Combined results presented in the submission only)
- In the full analysis set a mean change from baseline in HAQ of -0.3 was achieved at week 12, and increased to -0.4 at week 24. The mean change of -0.4 was maintained at each time point up to week 88.

- Results presented in the submission only indicate that adalimumab led to a sustained improvement in disability and physical function as measured by the mean change from baseline in FACIT-F and DLQI, and the SF-36 domain scores.
- Overall the safety profile during the open-label extension study was similar to that reported in the placebo-controlled trials.

Critique of the Abbott submission

The submission was reasonably accurate according to the data in the published abstracts, and overall appears to be a fair interpretation of the trial. Specific points include:

- The results of this study quoted in the submission are stated to be based on more recent analyses of the data from the M02-537 trial. This full analysis set included 395 patients, of which 382 received adalimumab.
- The results presented in the submission from the full analysis set are largely in accordance with those reported at differing time points in the published abstracts.
- The manufacturer's submission (pg 40-41) provides combined PGA results for patients in both trials at 12 and 48 weeks and states that this response was maintained up to 88 weeks; however, no response rate at 88 weeks is presented.
- The submission states (pg 41-2) that adalimumab demonstrated clinically and statistically significant improvements in mTSS, FACIT and DLQI. However, no statistical analyses of the open-label results are quoted in the manufacturer's submission.
- The disability and physical function patient reported outcome measures FACIT-F, SF-36 and DLQI are presented by the manufacturer in the submission, but are not reported in any of the published abstracts. This supplemental information presented in the submission is not in the public domain and therefore cannot be validated externally.

Summary

In patients with moderately to severely active PsA with a history of NSAID or DMARD failure, the efficacy of adalimumab in treating the joint and skin manifestation appeared to be maintained for up to 88 weeks. Adalimumab treatment also substantially improved disability and physical function. The adverse effect profile after long-term adalimumab treatment was similar to that observed after in the placebo-controlled trials (12 – to 24 weeks). However, the robustness of these findings cannot be fully assessed due to the lack of statistical analyses. The majority of the patients in this trial do not match those for whom anti-TNF therapy such as adalimumab are currently recommended, i.e. patients who have failed an adequate trial of two standard DMARDs.

4.3.4 Trial M04-724 (STEREO)[35]

Trial Summary

The STEREO trial is an on-going prospective 12 week open-label study to assess the efficacy and safety of adalimumab in patients with active PsA, in real-life clinical practice. Preliminary results show that adalimumab, when added to insufficient standard therapy, improves the ACR20 response in patients with acute PsA in the short term (12 weeks) in real life clinical practice. Results of several other endpoints, including effects on quality of life, psoriasis and ACR50 and 70 responses, also show improvement.

Since this study has not been fully published and the results are reported only in abstract form there are insufficient data presented to fully assess the quality and validity of this study. Although the abstract data are supplemented with additional information provided by the manufacturer in the submission, this supplemental information is not in the public domain and therefore cannot be validated externally. A structured critical appraisal of this trial is presented in Appendix 2.4.

Important Trial Points

Key trial points are outlined below:

- Published in abstract form only.
- Submission included additional unpublished data.
- Eligible patients had moderately to severely active PsA (defined as ≥ 3 swollen joints and ≥ 3 tender joints) and a history of an inadequate response or intolerance to at least one DMARD.
- The abstract reported the data from April 2006 when 253 patients out of a total of 441 had completed week 12.
- Although PsARC was identified as an outcome measure, so far no data regarding this have been reported.
- 72% of patients achieved an ACR20 response by week 12 with 49% and 27% of patients achieved an ACR50 and 70 response respectively by week 12.
- At week 12 the mean swollen and tender joint counts were reduced from baseline (reduction of 7.0 from 9.3 and reduction of 10.3 from 17.6, respectively).
- Mean DAS 28 scores showed a decrease (2.6) when compared to baseline (4.8) at week 12 as did the mean HAQ (baseline = 6.4, week 12 = 2.8).
- The percentage of patients with a PGA of ‘clear or almost clear’ increased over the 12 week period, from 35% at baseline to 65% at week 12.

- Some of the outcome measures are evaluated with respect to prior biologic use. However the numbers of patients with prior biologic exposure (n = 47; 17% of total numbers) is small and sub-group analyses need to be interpreted with caution.
- The spectrum of adverse effects was similar to those highlighted in earlier randomised controlled trials and withdrawal rates, so far, appear to be low (8%).

Critique of the Abbott submission

The submission was accurate according to the data in the published abstract, and appears to be a fair interpretation of the trial. Specific points include:

- The data from the sub-group analysis of prior biologic therapy compared with biologic naïve patients for ARC 20/50/70, DAS-28, HAQ and PGA measurements were not reported in the submission.
- PsARC and psoriasis target lesion assessment responses data at week 12 were not reported in either the submission or the abstract despite being included as outcome measures in the submission.
- The aim of the trial according to the abstract is to 'examine the efficacy and safety of adalimumab in a large number of patients with PsA, in real-life clinical practices including patients with various co-morbidities'. However the study design in the submission lists the presence of co-morbidities (e.g. uncontrolled diabetes, unstable ischaemic heart disease, IBD and chronic leg ulcer) as being a reason for exclusion from entry into the trial.

Summary

The preliminary data from this open-label trial show that in the short term (12 weeks) adalimumab appears to be effective and safe in patients with active PsA when added to insufficient standard therapy in real life clinical practice. However, it does not add significantly to the data obtained from the previous randomised controlled trials and the consistently high placebo response rates seen in PsA trials add to the difficulty in interpreting the data from uncontrolled trials. It is unclear how many of the enrolled patients are being treated concomitantly with DMARDS or NSAIDs. It is also unclear which proportion of the trial population would fit the BSR recommendations for anti-TNF therapy (inadequate response to $\geq$ 2 DMARDs)

Current results do not tell us how adalimumab performs in the longer term or how it compares to other anti-TNF drugs in the same class, or even how it compares to standard DMARDs.

4.3.5 Meta-analysis (M02-518 & M02-570)[1]

In the submission Abbott presented a meta-analysis of studies M02-518 and M02-570. The studies included in this meta-analysis were identified through a full literature search conducted by the authors of the submission and this has been commented upon elsewhere (Appendix 1). Only two RCTs were identified as suitable for analysis by the authors and subsequently included in the meta-analysis. The quality assessment of the two included studies is presented elsewhere in the evaluation report (Appendix 2). Although one study (M02-570) was available only in abstract form which limited the analysis, both studies were deemed to be of good quality.

The meta-analysis was conducted using RevMan 4.2.9. Heterogeneity was investigated for all outcome measures using the Chi-squared test and showed no significant statistical heterogeneity ($p<0.0001$ for all outcomes), suggesting it was appropriate to use a fixed effects model. A fixed effects analysis assumes that only within-study variation is taken to influence the uncertainty of results. However, there were some notable differences between the two study populations with regards to DMARD use. In the M02-570 study patients could use DMARDS other than methotrexate, which was the only DMARD allowed in M02-518, suggesting potential 'clinical heterogeneity', although this was not tested. The robustness of the results was, therefore, also examined using the more 'conservative' random effects model which includes both within-study and between-study variation in the assessment of uncertainty. Overall the results provided by these different modelling assumptions were very similar, with only marginally wider confidence intervals using the random effects model.

Only the outcomes relating to the arthritis component were included in the meta-analysis as M02-570 did not consider outcomes relating to the psoriasis components of the disease. The primary outcome variables were ACR20/50/70 and PsARC. Relative risk (RR) was calculated for the primary outcomes with 95% confidence intervals (CI). Across the two trials (n=413), at 12 weeks, around 53% of patients treated with adalimumab achieved an ACR20 (pooled RR 3.61 (2.55 to 5.12)) demonstrating a basic degree of efficacy in terms of the arthritis component of the disease. In addition, around 34% of patients receiving adalimumab achieved an ACR50 response (pooled RR 10.23 (4.81 to 21.75)) and around 18% an ACR70 response (pooled RR 26.04 (5.18 to 130.88)) demonstrating a good level of efficacy. Around 59% of patients treated with adalimumab achieved a PsARC (pooled RR [redacted] ([redacted])) which is the only outcome measure of joint disease specifically developed for people with PsA.

However, this analysis does not indicate how adalimumab performs in the longer term, or how it compares to other anti-TNF drugs in the same class, or even how it compares with DMARDs. It is limited both in the extra information it provides over the findings in the individual trials and the very short-term duration of therapy. Furthermore, patient numbers were significantly larger in the M02-518 trial in which patients had an inadequate response to NSAIDs and this has subsequently influenced the findings of the analysis.

4.3.6 Indirect/mixed treatment comparison

In order to complete the clinical evaluation of adalimumab it is necessary to compare the relative efficacy of the available treatments for patents with PsA who have had an inadequate response to DMARD therapy. Because there are no studies directly comparing adalimumab with other anti-TNF agents, the manufacturer provided an indirect treatment comparison to strengthen inference concerning the relative efficacy of adalimumab. This analysis was similar to that used in the previous assessment report for NICE on etanercept and infliximab.[23] Given that the licensed indications for adalimumab, infliximab or etanercept, indicate that they should only be offered following inadequate response to at least two DMARDs, it is reasonable not to compare the agents with DMARDs.

Two independent literature searches were conducted by the ERG to ensure all relevant clinical trials were assessed for inclusion in the indirect treatment comparison. Clinical trials assessing DMARDs in the treatment of PsA were conducted as part of the previous Health Technology Assessment (HTA) for infliximab and etanercept in PsA. As the search criteria used by Abbott were not detailed in the submission this specific search could not be reproduced. The search criteria used in the previous HTA were re-run including data from 01/01/04 to 01/01/07, details of which are given in Appendix 1. The clinical trials identified as part of this literature search were congruous with those included in the indirect treatment comparison table 5.6.2 (page 46) of the manufacturer's submission. A summary of the outcome measures from these clinical trials can be found in Table 2.6.1, Appendix 2. It should be noted that ACR response criteria and the PsARC unit of measurement were recently accepted as the basic measure of efficacy in RA. The BSR guideline recommends the use of these measures in clinical trials examining the efficacy of treatments for PsA. PsARC is recommended for assessment of the arthritis component and evaluation for continuation of treatment. However, a number of the DMARD trials were conducted prior to acceptance of these clinical measures and do not always use them as outcome measures.

The three assumptions on which the manufacturer's indirect treatment model was based are discussed in detail in section 5.3.2 of the ERG report. One of the main assumptions applied by the manufacturer's chosen model was that the results from the independent trials were the same as if they were from one study due to the similar placebo response in each of the trials identified in table 5.6.3 (pg 47 of Abbott's original submission). Other factors which should have been taken into account are; the degree of similarity in the trial methodology and outcome measures with an absence of clinically important differences in trial populations.

The ERG identified a number of differences in the inclusion criteria for the three phase III trials used in manufacturer's submission:

- The degree of active PsA – IMPACT II[36] recruited patients with $\geq$ 5 swollen or tender joints compared with ADEPT[31] and Mease (28)[37] (both $\geq$ 3)
- Only the IMPACT II[36] trial excluded patients who were rheumatoid factor (RF) positive. Mease (28♦)[37] and ADEPT trials included between 4 and 11% of patients who were RF positive.
- Mease (28)[37] and ADEPT[31] specified an inadequate response to previous NSAID treatment whereas IMPACTII[36] specified an inadequate response to DMARDs or NSAIDs
- Requirement for an active skin component to the disease was specified in two of the three trials. In the ADEPT[31] trial patients with a documented history of psoriasis were also included.
- Concomitant treatment with methotrexate was allowed in each of the anti-TNF agent trials ($\leq$25-30 mg/week and stable for preceding four weeks) however the proportion of patients in each study varied from 45-65%.

Synthesising the results to inform the cost-effectiveness model may also be affected by the differences in baseline characteristics:

- Patients who were RF-positive were included in the adalimumab trials (ADEPT[31] and M02-570[32]), (Mease[38] (28)) (etan), and (Mease[37] (27))(etan). The proportion of patients with PsA who are RF-positive varies from 4-15%. There is some suggestion that rheumatoid factor is a severity marker in patients with PsA (i.e. the presence of RF would suggest more severe disease). Therefore, the patients included in these trials may have more severe joint disease than those in other trials.
- PASI score – detailed in Appendix 2, table 2.6.2

♦ For clarity *(27)(28)* refer to the reference numbers cited in the Abbott submission.

The number of patients available for assessment of PASI varied widely in each group, as did the baseline PASI score. In the Etanercept trial conducted by Mease et al the difference between baseline PASI in the placebo and treatment group may have been large enough to introduce bias to this outcome measure.

Table 5.6.4 (pg 47 of the original submission) reports the results of a subgroup analysis of patients only on methotrexate. The data were taken from the subset of patients that were on methotrexate in the placebo arms of the studies (Mease et al Etanercept and ADEPT M02-518) versus anti-TNF agents. 'The manufacturer concluded that this provides evidence of the inferior efficacy of conventional DMARDs for both joint and psoriasis outcomes compared to anti-TNF agents. This comparison is the best evidence available of comparative conventional DMARD and anti-TNF agent efficacy'. The manufacturer confirmed (in the points for clarification response section A2) that patients enrolled in the ADEPT M02-518 study who continued to take methotrexate were classed as having an inadequate response to this DMARD. This would suggest the use of this data to evaluate the efficacy of methotrexate is inappropriate. It is not clear from the data available if this is also the case for the etanercept trial conducted by Mease et al but it seems a reasonable assumption that this holds true.

The assumptions made by the manufacturer in their indirect/mixed comparison are pivotal to the validity of the cost-effectiveness results. These assumptions and the ERG's concerns with this approach are discussed in detail in section 5.3.2 of the ERG report.

4.4 Review of Current Treatment Guidelines

The British Society for Rheumatology has produced guidelines on the use of anti-TNF drugs in PsA.[26] These guidelines state that if a patient has active disease (defined as three or more swollen or tender joints on two separate occasions one month apart) and has failed to respond to adequate trials of at least two standard DMARDs (leflunomide, sulfasalazine, methotrexate or ciclosporin, individually or in combination), and satisfy none of the exclusion criteria then they should be considered for licensed anti-TNF-α therapy. The details of these guidelines are attached in Appendix 3. In January 2006 the BSR released a statement on the use of Adalimumab for PsA.[39] This statement supports the use of adalimumab as a treatment for adult patients with active and progressive PsA, who have had an inadequate response to previous DMARDs and in accordance with the current BSR guidelines for the use of anti-TNF-α drugs in PsA.

In December 2005 the Scottish Medicines Consortium (SMC) recommended adalimumab for use within NHS Scotland for the treatment of adults with active and progressive PsA when the response to previous DMARD therapy has been inadequate;[40] and in June 2006 the All Wales Medicines Strategy Group (AWMSG) recommended that adalimumab should be available for use within NHS Wales for the treatment of PsA in accordance with the licensed indication subject to the following restriction: Adalimumab is used in accordance with the current BSR guidelines for anti-TNF-α therapy in adults with PsA.[41]

4.5 European Medicines Agency (EMEA) European Public Assessment Report (EPAR) Post authorisation opinion

In June 2005 the European Medicines Agency (EMEA) approved a license extension for adalimumab (Humira®) for the 'treatment of active and progressive psoriatic arthritis in adults when the response to previous DMARD therapy has been inadequate'. The data submitted to support this license extension application was derived from M02-518 (ADEPT), M02-570 and preliminary data from the open label extension study M02-537 (12 weeks).

The initial application proposed the use of adalimumab alone or in combination with DMARD products in the treatment of PsA. Approximately 50% of patients in the studies were on concomitant methotrexate at randomisation. However, the number of patients treated with other DMARDs was very low and deemed insufficient to ensure safe use concomitantly with adalimumab and was not therefore included in the indication.

The Committee for Medicinal Products for Human use (CHMP) noted the difference in the inclusion criteria relating to the previous inadequate response to NSAIDs (M02-518) or DMARDs (M02-570) and continuation of DMARD therapy during each study. Additional information on the subpopulation treated with methotrexate was obtained from the market authorisation holder (MAH). The heterogeneity of the population was not deemed to bias in favour of demonstrating response to adalimumab vs. placebo, which was shown in all subsets. It was not considered feasible to conduct clinical trials in a sufficient number of patients in all subpopulations and with possible concomitant treatment strategies due to the considerable heterogeneity in PsA. A tendency towards better efficacy was seen in patients with higher inflammatory activity, based on CRP levels, which is in line with active disease.

There were several areas of the application the CHMP felt that data were missing and declined to consider within the application. These were:

- Skin involvement – the number of patients with a PASI $\geq$ 10 (a score relevant for moderate to severe psoriasis) was insufficient to draw definite conclusions on the efficacy of adalimumab in the psoriatic domain of PsA.
- Radiological progression – a period of 24 weeks was deemed too short to allow an assessment of this parameter, and thus, these data were not considered further in the efficacy assessment.
- Maintenance of efficacy or increase in disease activity after stopping treatment – it was felt that the additional information supplied by the MAH was insufficient to allow conclusions to be drawn. In addition, the interim data from the M02-570 trial showed a tendency towards tapering of effect in the limit number of subjects (n = 40) at 48 weeks.
- Dose escalation to adalimumab 40mg every week.

No new safety concerns were identified as part of this evaluation process and the safety profile was recognised as being similar to that previously known from anti-TNF therapy.

4.6 Committee for Medicinal Products for Human use (CHMP) guidelines on clinical investigation of medicinal products for the treatment of PsA[4]

This guidance includes recommendations on the type of clinical trial that should be conducted and the efficacy outcomes that should be used.

Documentation required in studies assessing treatment efficacy in psoriatic arthritis includes:

- Demographic characteristics of patients.
- Duration of disease (psoriasis and arthritis).
- Previous and concomitant therapy.
- Concomitant disease.
- Severity and extend of disease (psoriasis and arthritis).
- Type of psoriasis.
- Disease activity (psoriasis and arthritis).
- Spinal and peripheral involvement.

Methodology for confirmatory studies:

1. Randomised, double-blind, parallel group design.

2. Efficacy of products claiming improvements in patients non-responsive to conventional DMARDs may be established by means of a placebo controlled add-on trial where all patients receive established standard therapy, e.g. methotrexate.
3. Comparison of the available treatment options for these patients, e.g. anti-TNF, may be necessary for an appropriate benefit/risk assessment, particularly if the product belongs to a new therapeutic class.

Main efficacy endpoints:

1. PsARC and ACR response for products intended to improve symptoms / physical function.
2. Any radiological scoring system as long as the choice is justified and the minimum relevant change should be established.
   Radiographs should be taken at fixed and predefined time points and be assessed by at least two assessors (blinded for treatment allocation, chronological sequence and initial assessment of the other assessors).

Secondary endpoints:

1. Axial involvement measured using ASAS response.
2. Individual components of the composite primary endpoints not assessed.
3. Individual assessments of the main domains of PsA (described in Appendix 4).
4. Dactylitis and enthesitis.

Skin lesions

1. The CHMP 'Guidance on clinical investigation of medicinal products indicated in the treatment of psoriasis' should be used to assess effects on skin lesions.
2. Demonstration of efficacy on skin disease will require separate specific trials, nevertheless the effect of any new therapy for PsA on skin lesions should be assessed (type of psoriasis, body surface area involved and presence of nail lesions).

4.7 Other relevant studies

No other relevant studies were identified by the ERG during a comprehensive literature search. See Appendix 2 for the search strategy employed.

4.8 Relevant ongoing studies

All relevant trials were included in the manufacturer's submission. Other than the ongoing MO2-537[32-34] and MO4-724 (STEREO)[35] studies there are no other relevant ongoing RCT studies examining the use of adalimumab in the treatment of PsA.

An open label study to further assess the safety and effectiveness of adalimumab 40mg when added to inadequate therapy for the treatment of PsA (ACCLAIM) has recently completed recruitment. This study sponsored by Abbott is assessing ACR, PsARC, PASI and HAQ responses in patients who have had an unsatisfactory response or intolerance to at least two prior or ongoing DMARDs (one of which has to be methotrexate).[42]

A small prospective study of adalimumab in patients with psoriasis and psoriatic arthritis is currently ongoing which aims to find the best predictive biomarker for response to treatment. This study which is investigating changes in cellular infiltrate and cytokine expression in biopsies of skin and synovium is expected to end in May 2007.[43]

A study evaluating the cost of the treatments of PsA refractory to conventional therapy has also completed recruitment. The purpose of this study is to conduct an economic analysis on the cost of conventional therapy as compared to biologic therapy and the direct/indirect costs of disease management in patients with refractory PsA. The primary outcomes are to qualify the economic burden of refractory PsA care. Secondary outcomes will assess the efficacy, safety, and cost effectiveness of the different therapies.[44]

Chapter 5
Economic Evaluation

5.1 Introduction

This section provides a structured critique of the original cost-effectiveness model submitted by Abbott Laboratories Ltd (the manufacturer). As part of the STA process, manufacturers are expected to perform a systematic review of existing cost-effectiveness evidence for the health care technology or process being assessed. Where there is no existing evidence or the existing evidence is insufficient, manufacturers may perform their own *de-novo* cost-effectiveness analysis.

The manufacturer's economic submission to NICE included (references in brackets refer to the manufacturer's submission):

(i) a description of the systematic search undertaken in an attempt to identify cost-effectiveness studies of therapies used in the treatment of psoriatic arthritis and a critical appraisal of relevant identified studies (p58-60; Appendix 3);

(ii) a report on the economic evaluation undertaken by the manufacturer (p61-94, in particular Figure 6.2.6.2, p67 the schematic of the model and Tables 6.2.6.1 - 6.2.9.8, p68-83 which provide information on the model parameters);

(iii) base-case cost-effectiveness results from the model (Table 6.3.1.1, p86-87 and Figures 6.3.1.1 – 6.3.1.2, p87-88);

(iv) stochastic sensitivity analysis results from the model (Table 6.3.3.1, p88-91);

(v) an Excel-based model comprising the manufacturer's economic model provided electronically; and

(vi) a copy of the psoriasis resource utilisation questionnaire used to inform the decision problem (Appendix 6).

Following a number of points of clarification raised by the ERG, a number of addenda were submitted by the manufacturer. These included:

(i) a table with response results at 12 and 24-weeks for all trials identified in the literature search, including mean change (SD) in HAQ scores (ERG clarification-NICE PsA Submission 120107.doc);

(ii) results of a full evidence synthesis of PsARC, ACR and PASI response at 12 and 24 weeks for all anti-TNF drugs using data from all identified studies in the literature search (ERG clarification-NICE PsA Submission 120107.doc);

(iii) the WinBUGS code used for the evidence synthesis of PsARC, ACR and PASI results (Appendix B, ERG clarification-NICE PsA Submission 120107.doc);

(iv) further information on the approach used to predict the correlation between different types of responses e.g. ACR/PASI/PsARC (ERG clarification-NICE PsA Submission 120107.doc);

(v) an Excel file containing the requested spreadsheets with Microsoft Solver calculations (PsA Solver Sheet.xls).

(vi) further analyses to test the robustness of the model under different assumptions, including a re-run of the base-case analysis based on a synthesis of 12-week and 24-week data (ERG clarification-NICE PsA Submission 120107.doc); and

(vii) further information about the methods used to estimate EQ-5D scores from SF-12 responses and summary table of baseline utilities, change in utilities based on PsARC response, for all patients and for patients with <3% and >=3% BSA affected by psoriasis and <2 and >=2 previous DMARD use (ERG clarification-NICE PsA Submission 120107.doc).

After consideration of the evidence submitted in response to the initial points for clarification and, with explicit authorisation from NICE, the ERG considered it necessary to request some additional analysis. The manufacturer's additional work included:

(viii) a revised evidence synthesis excluding the open label results from the adalimumab study M02-570 at 24 weeks (NICE PsA results removing MO2-580 open label results.doc);

(ix) revised cost-effectiveness results using the new evidence synthesis excluding the M02-570 study results at 24 weeks (NICE PsA results removing M02-580 open label results.doc); and

(x) an Excel file containing the new effectiveness parameters as implemented in the model (Updated NICE PsA model parameter sheet 260107.xls).

This section focuses on the economic evidence submitted by the manufacturer as part of their original submission. The submission is subject to a critical review on the basis of the manufacturer's report and by direct examination of the electronic version of the economic model. The critical appraisal is conducted with the aid of a checklist to assess the quality of economic evaluations and a narrative review in detail of the key limitations of the model. The critical appraisal and review are used to identify the key areas of uncertainty surrounding the original submission. These areas are then used to formulate the points for clarification raised by the ERG to the manufacturer. Section 6 presents a description of the additional work requested from the manufacturer and a critique of their re-submitted results.

5.2 Existing cost-effectiveness evidence

As part of the manufacturer's submission, a systematic search was undertaken with the aim of identifying published studies evaluating the cost-effectiveness of adalimumab for the treatment of moderate to severe psoriatic arthritis. The search strategy, key words, date range and sources searched to identify the economic studies were appropriate for this purpose. The manufacturer's search did not identify any studies which evaluated the cost-effectiveness of adalimumab for this indication. Two published studies assessing the cost-effectiveness of alternative anti-TNF agents were identified and subjected to a detailed critical appraisal by the manufacturer.

The searches undertaken by the manufacturer were replicated by the ERG in order to validate the evidence base considered. The ERG found that the search was reproducible, and the results were consistent with the original search. However, it was not clear from the company submission how many different results met their search criteria, as they did not show how many of them (4 in Medline, 3 in Medline IP, 6 in Embase, 3 in the NHS Economic Evaluation Database (NHS EED) and 3 in the Health Economic Evaluation Databse (HEED)) were duplicated among the databases.

A second search was conducted by the ERG using a much broader search strategy designed to capture all NHS EED records relating to PsA. In addition, the ERG ran searches of the NHS EED administrative system (CAIRS B) and of Medline, Cinahl,

Embase and EconLit to identify any recent studies not yet screened for NHS EED. After deduplication there were 16 potential studies that met the inclusion criteria. Of the 16 studies identified, only 2 of these were full cost-effectiveness analyses. Both studies assessed the cost-effectiveness of alternative anti-TNF agents (etanercept and infliximab). See Appendix 5 for details of the search strategies conducted by the ERG. The ERG concurs with the manufacturer that there are no existing published cost-effectiveness studies evaluating the use of adalimumab for the treatment of moderate to severe PsA. Our results were also consistent with the manufacturer's search in that only two published cost-effectiveness studies of other anti-TNF therapies used in the treatment of PsA were identified.[23 45]

Given that the cost-effectiveness of alternative anti-TNF agents was not considered directly relevant to the review of adalimumab the ERG has not undertaken a formal critique of these studies. However, we concur with the manufacturer's conclusion that existing models in this area are potentially limited since they have not included the impact of the alternative anti-TNF agents on the skin component of the disease in the analysis.

5.3 Overview of manufacturer's economic evaluation

The manufacturer's submission is based on a *de-novo* economic evaluation to estimate the cost-effectiveness of adalimumab for the treatment of moderate to severe PsA. A brief overview of the key assumptions used in the analysis, alongside a narrative description of the main approach used, is reported below. This is followed by a more detailed critique of the economic evaluation and its assumptions.

The key assumptions used in the model include:

(i) The relevant comparators to adalimumab considered were other alternative anti-TNF agents (infliximab and etanercept) and conventional DMARDs. Palliative care is not considered a relevant treatment alternative until all conventional DMARDs have been exhausted (up to 5 alternative DMARDs after initial treatment are considered). Patients who do not respond to an anti-TNF treatment are assumed to revert back to conventional DMARD therapy (again, assuming that palliative care is only an option after all DMARDs have been exhausted).

(ii) PASI, ACR and PsARC responses were chosen as the primary outcome measures and only trials that reported these three measures were included in the analysis. Trials of duration less than 6 months (24-weeks)

were excluded on the basis that shorter duration studies may underestimate the efficacy of anti-TNF agents.

(iii) The presence of skin disease was determined according to whether patients showed an affected BSA with psoriasis > 3%. The level of skin involvement was considered to be a potentially confounding factor for response to treatment. Response rates were adjusted for the skin involvement in the patient population of the trials, using the average percentage of patients with BSA > 3% across the treatment arms for adalimumab, etanercept and infliximab (66%) as representative of the population of interest.

(iv) An indirect synthesis of response data was undertaken under the assumption that the correlation between response types for patients with, and without, skin disease observed in the M02-518 adalimumab trial was identical for the rest of anti-TNF agents and representative of the population of interest.

(v) Patient level data from the M02-518 adalimumab trial was used to predict HAQ and PASI changes, under the assumption that the response measures (ACR, PASI, PsARC) discriminate all of the change in HAQ and PASI and that the baseline patient characteristics of the M02-518 adalimumab trial are representative of the population of interest.

(vi) Initial treatment failure (and the decision to continue with treatment) was analysed in terms of 12-week PsARC response for the base-case analysis.

(vii) The model uses 6-month cycles as it is assumed that this time period reflects consultation times at which treatment can be changed.

(viii) Anti-TNF agents are assumed to halt HAQ progression whilst a patient is responding to treatment. Patients receiving DMARDs are assumed to experience an annual progression in HAQ progression regardless of response status.

(ix) The extent of HAQ rebound after treatment failure is assumed to be of the same magnitude as the initial gain (i.e. rebound equal to gain); the PASI rebound is also to the starting level.

(x) Adverse events of treatments were not included in the analysis on the basis that these would have a negligible impact on cost-effectiveness.

The results for the economic evaluation are presented for the base-case, and thereafter, for several other scenarios using sensitivity analysis. A stochastic sensitivity analysis has also been undertaken.

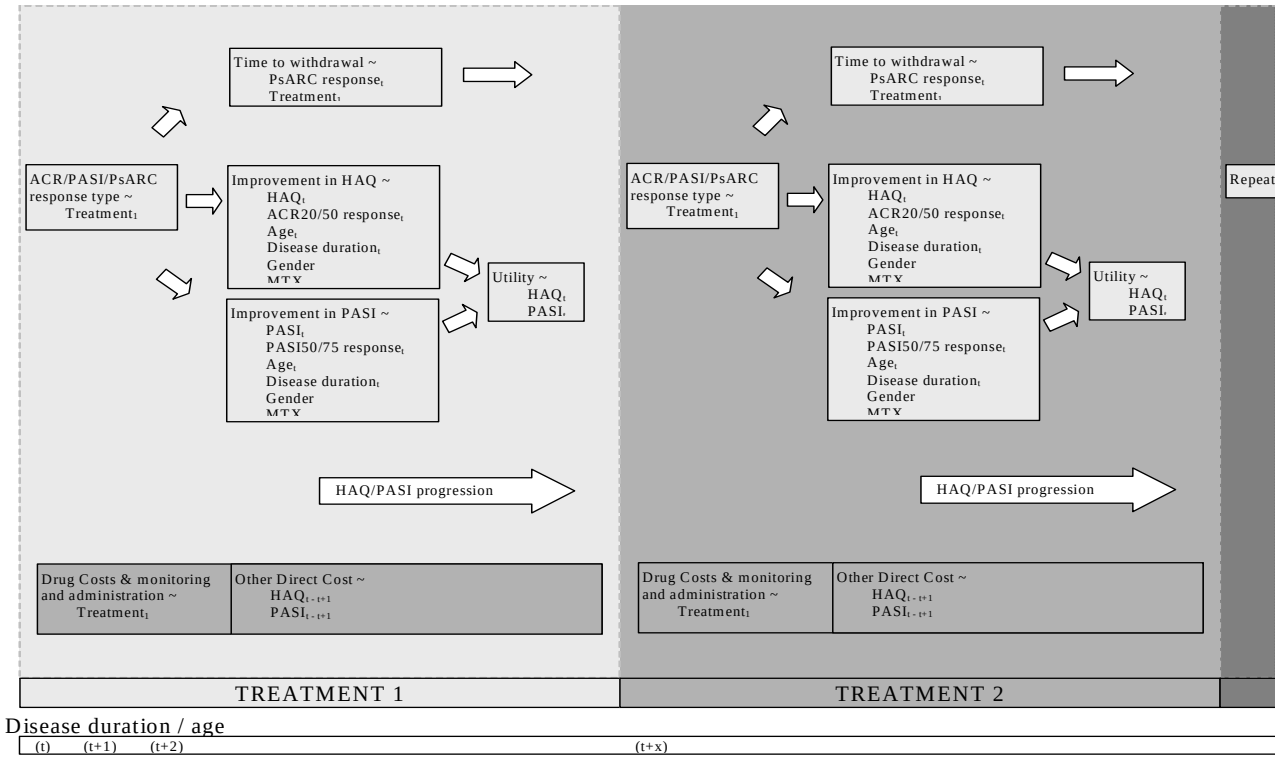
5.3.1 Model structure

The model uses a form of micro-simulation known as an individual sampling model to describe the natural history of psoriatic arthritis disease, modelling individual patient histories from time of entry into the model until death (i.e. lifetime horizon). The model uses Monte-Carlo simulation at the patient level. At each decision node a random number decides the route a patient takes based on calculated probability, so each hypothetical patient represents only one possible route that can be taken.

Patients with active disease having failed two previous DMARD therapies enter the model. They start on their first treatment (adalimumab, etanercept, infliximab or conventional DMARDs) and remain on it for 6-month cycles until they no longer respond, at which point they then switch onto the next drug in the treatment sequence. Alternative treatment sequences for the anti-TNF agents are not considered such that all patients are assumed to receive conventional DMARDs after failure of initial therapy. For the base-case analysis, initial response (and hence the decision to continue with the first treatment) is defined as meeting the 12-week PsARC response criteria, which is sampled from the joint distribution of their ACR, PsARC and PASI responses (details discussed in section 5.3.2). Based on the type of response and the baseline characteristics for each simulated patient, their improvement in HAQ and PASI is then predicted (details discussed in section 5.3.3).

Finally, based on the HAQ and PASI scores the health utility and direct costs for the cycle are determined. After the initial 6-month cycle, long term withdrawal is based on evidence from an observational study.[46] Patients remain or move onto the next drug following the same process as described above, until exhausting all treatment options (up to 5 alternative DMARD treatments are considered at which point patients progress to best supportive/palliative care). At any point in the model, patients may die and exit to the absorbing death state. A schematic of the model is presented in Figure 5.1.

Figure 5.1: Schematic of manufacturer's model



5.3.2 Treatment effectiveness

The methods of evidence synthesis were a central component of the manufacturer's submission. In the absence of direct head-to-head RCTs for the different anti-TNF agents (adalimumab, etanercept and infliximab), the manufacturer applied indirect approaches to estimate the relative effectiveness and cost-effectiveness of these agents. The approach used by the manufacturer is outlined in Section 5.6 (p45) of their submission.

The manufacturer cited a number of reasons why they felt that conventional approaches to indirect/mixed treatment comparisons may not be appropriate for this analysis. These included:

- Evidence from a recent published meta-analysis (using a mixed treatment comparison model) of biologics in RA identified significant heterogeneity across studies, based on differences in study-level prognostic factors such as disease duration.[47] Meta-regression approaches were thus required in order to obtain a more reliable estimate of the relative effectiveness of the different treatments. The manufacturer notes that these approaches require a large number of trials in order to obtain robust estimates.

- Previous approaches using mixed treatment comparisons for the assessment of etanercept and infliximab for PsA[23] were based on a single outcome of interest related to the extent of joint disease and did not simultaneously consider the impact on psoriasis outcomes. The manufacturer argues that methods of mixed treatment comparisons for multiple, related outcomes are not available.

- The difference in the number of patients across the trials who had active psoriasis amenable to severity scoring (e.g. with a BSA > 3%) across the trials. The manufacturer claims that the response to treatment has been found to be different in patients with BSA <3% and >3%. The manufacturer notes that access to patient level data would be required in order to adjust for this potential imbalance across existing studies.

Faced with these concerns the manufacturer outlines an alternative approach which requires three key assumptions:

1. Trials were only considered if they reported data on all primary outcomes considered (including PASI, PsARC and ACR responses). The rationale outlined by the manufacturer is that each response parameter plays a key role within the cost-effectiveness model (with treatment continuation being determined by PsARC response and ACR and PASI response data being used to estimate utility and costs).

2. As previously stated the manufacturer claims that response to treatment differs in patients with BSA <3% and >3%. In order to make a more reliable comparison between the different anti-TNF agents, the manufacturer argues that it is necessary to adjust for the number of patients with psoriasis at baseline across the studies. In the absence of patient-level data from each of the trials, the manufacturer assumes that the correlation between response types based on patient level data for adalimumab would be the same for the other anti-TNF agents.

3. After adjusting for skin involvement, the approach used by the manufacturer assumes that the results from the different trials can be treated as if they come from a single study. This approach assumes that both the relative and absolute effects for each of the different agents are exchangeable across the different studies. The manufacturer cites similar placebo response results from the three

main anti-TNF studies included in their synthesis as justification for this assumption.

Given the lack of head-to-head RCT data for the anti-TNF agents, and the need to consider conventional DMARDs as a comparator in the model, the manufacturer undertook a systematic search of Medline to identify relevant studies of existing anti-rheumatic therapies (including both conventional DMARDs and anti-TNF agents). We have previously discussed and critiqued this element in section 4.3.6 of this report.

In total the manufacturer identified 8 trials meeting the initial inclusion and exclusion criteria. The manufacturer then applied a secondary filter to ensure that the studies could be used as part of the cost-effectiveness model. Follow-up across these studies ranged from 12 to 54-weeks. The manufacturer excluded studies of less than 6-months duration on the basis that studies of shorter duration may potentially underestimate the true efficacy of anti-TNF agents. This is a key assumption which we address in detail in the critical review section. In addition, the manufacturer noted that previous use of DMARDs and disease duration has been shown to have prognostic value in studies of patients with RA. In order to reduce this potential source of heterogeneity in the evidence synthesis, trials were only included at the secondary filter stage if they recruited patients with disease duration of >8 years and DMARD therapy has been used at least once. Three out of the eight trials met these additional selection criteria.

The 3 trials included in the manufacturer's analysis were all studies of anti-TNF agents evaluated against placebo, with a single study for adalimumab (M02-518 ADEPT study), etanercept (Mease) and infliximab (IMPACT II). All 3 trials were of 24-week duration. Response rate data reported in the manufacturer's submission for each of these studies is shown in Table 5.1 below. Additional information was also presented on potentially important prognostic factors such as the % of subjects with BSA >3% and the % of subjects receiving methotrexate at baseline.

Table 5.1: Response Results at 24 Weeks for Trials Used in Economic Evaluation

Trial name	Treatment	N	% with BSA > 3%	% MTX*	ACR20	ACR50	ACR70	PASI50	PASI75	PASI90	PsARC
Mease	Etanercept	104	60%	49%	50%	37%	9%	47%	23%	6%	70%
	Placebo	101	65%	45%	13%	4%	1%	18%	3%	3%	23%
ADEPT	Adalimumab	151	43%	50%	57%	39%	23%	74%	59%	42%	60%
	Placebo	162	46%	51%	15%	6%	1%	11%	1%	0%	23%

IMPACT II	Infliximab	100	87%	45%	54%	41%	27%	75%	60%	39%	70%
	Placebo	100	83%	47%	16%	4%	2%	8%	1%	1%	32%

* MTX = methotrexate

The manufacturer noted that a major difference across the trials, in terms of baseline clinical demographics, was in the proportion of patients with active skin disease based on the percentage of subjects with BSA >3%. The manufacturer reported that an analysis of patient-level data from the ADEPT (M02-518) study of adalimumab demonstrated high correlation between response variables and, in particular, that response parameters for the arthritic component of the disease (ACR and PsARC) differed according to the level of skin involvement affected by psoriasis at the time of study entry. Given the reported differences between studies based on the numbers of patients with skin involvement at baseline (BSA>3%), ranging from 43% in ADEPT to 87% in IMPACT II, the manufacturer deemed that it was necessary to adjust the results to a common percentage of patients with psoriasis (BSA>3%) in order to provide a fairer comparison between all the trials. While the response data for patients with and without skin disease at the time of entry into the study could be estimated from the M02-518 ADEPT study, similar data was not available for the Mease and IMPACT II studies. In the case of these 2 trials the percentage of responders for each individual category (e.g. ACR20, PsARC, PASI75 etc) and not the percentage of responders for all response types were reported. An optional add-on to Excel (Microsoft Solver) was used to estimate these proportions assuming that the correlation between response types observed in the ADEPT trial would be the same for the other anti-TNF agents.

Appendix 5 of the manufacturer's submission provides further details on the observed data from the ADEPT study and the predicted response for the other treatments. The observed response rates and correlations between responses (PsARC, ACR and PASI) from the ADEPT study were used to estimate the most likely response rates for the other treatments using Microsoft Solver. Microsoft Solver is an optimisation package within Excel that can be used to find the values of certain cells in a spreadsheet that optimize (maximize or minimize) a certain objective subject to particular constraints. The manufacturer used the ADEPT study to estimate the correlation between PsARC response, ACR response and PASI response whilst adjusting for the proportion of patients with BSA >3% (studies were adjusted to a common percentage of patients with BSA >3%, applying a figure of 66% based on a pooled estimate across the treatment arms for adalimumab, etanercept and infliximab). This data was then applied to the overall response rates

from Mease and IMPACT II. Estimates of the distribution of responses for patients with and without skin involvement were then estimated for these additional studies using Microsoft Solver. This was achieved by minimising five separate components subject to a series on constraints outlined in Appendix 5 of the manufacturer's submission. The optimisation software was used to generate correlated response tables for each treatment. Table 5.2 summarises the data generated for infliximab to illustrate the response data used in the model.

Table 5.2: Estimated responses with infliximab (IMPACT II)

PsARC	ACR response type	PASI				TOTAL
		0%-50%	50%-75%	75%+	No PASI	
Non-responders	0%-20%	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	20%-50%	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	50%+	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Responders	0%-20%	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	20%-50%	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	50%+	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
TOTAL		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Uncertainty surrounding these was incorporated into the model using a Dirichlet distribution, maintaining the original sample size of each trial. Based on the response tables each hypothetical patient could then be sampled from the joint distribution of their ACR, PsARC and PASI responses based on alternative response classifications (e.g. PsARC, PASI 50+, PsARC and PASI50+ etc). This enabled the manufacturers to examine a number of alternative decision rules relating to treatment continuation.

By using response rate data from trials of 6-months duration there is a potential inconsistency in comparison to current guidelines on the management of PsA, which recommend that the decision to continue treatment should be based on response at 12-weeks. In particular, previous NICE guidance for etanercept and infliximab state that treatment should be discontinued in patients for whom their PsA has not shown an adequate response at 12-weeks. In order to adjust for this in the model the manufacturers estimated 12-week response data by adjusting the 6-month response estimates. The adjustments were made by estimating the proportion of patients responding at 6 months who were also responders at 12-weeks using data from the ADEPT study (see parameters table 6.2.6.1. in original submission, p68). These adjustments are reported in Table 5.3 below. In the absence of similar estimates for

the other agents, the manufacturer applied the same adjustment to all treatments included in the model.

Table 5.3: Adjustments applied to ACR, PASI and PsARC responders at 6 months

Parameter	Value	Distribution	Source
Response			
PsARC/PASI/ACR Responses	See Appendix 5	Dirichlet	Various trials
% PsARC Responders at 6 Months who were PsARC responders at 3 months	80	Beta	ADEPT
% ACR20 Responders at 6 Months who were ACR20 responders at 3 months	78	Beta	ADEPT
% ACR50 Responders at 6 Months who were ACR50 responders at 3 months	71	Beta	ADEPT
% PASI75 Responders at 6 Months who were PASI75 responders at 3 months	70	Beta	ADEPT

The base-case analysis from the manufacturer assumed that treatment would only be continued based on PsARC response at 12-weeks. A number of alternative decision rules based on PsARC and/or PASI75 response at 12-weeks and 6-months were also evaluated. In each case the evidence synthesis was used to estimate the response rates for PsARC, PASI and ACR. PASI and ACR response rates were then used to estimate health outcomes and costs in the model. Further details on how PASI and ACR response data were used are discussed in the following sections.

5.3.3 HAQ / PASI prediction

Based on the type of response (ACR and PASI) and the baseline characteristics for each simulated patient, the improvement in HAQ and PASI scores is predicted using regression analysis. The predicted HAQ and PASI scores were then used as the basis for estimating costs and health utilities.

Patient level data from the ADEPT trial was used as the basis for predicting HAQ and PASI scores at 24-weeks. A forward stepwise regression was used to select significant covariates in predicting HAQ and PASI scores. Covariates were based on ACR and PASI response parameters (at 24-weeks) and baseline patient characteristics (i.e. age, gender, PsA duration, baseline HAQ, baseline PASI and use of concomitant methotrexate). The results of the regression models are shown in Table 5.4 and 5.5 below, alongside the regression equations for predicting (i) HAQ and (ii) PASI scores at 24-weeks. Two response parameters were combined (ACR50-70 and ACR70+ were combined to form ACR50+, and PASI75 and PASI90

were combined to form PASI75+) since their coefficients were similar and as separate covariates added very little to the explanatory power of the statistical model.

(i) **HAQ24** = α+ β1HAQ0 + β2Age0 + β3Gender + β4Duration0 + β5MTX + β6ACR20-50$_{24}$ + β7ACR50+$_{24}$

(ii) **PASI24** = α+ β1PASI0 + β2Age0 + β3Gender + β4Duration0 + β5MTX + β6PASI50-75$_{24}$ + β7PASI75+$_{2}$

Table 5.4: HAQ at 24-weeks regression coefficients

Description	Covariate	Parameter estimate	Standard error	t Value	Pr > |t|
Intercept	α	0.0856	0.1116	0.77	0.4441
Baseline HAQ	β_1	0.6856	0.0351	19.55	<0.0001
Baseline Age	β_2	0.0044	0.0020	2.2	0.0284
Gender (1= Male)	β_3	-0.0575	0.0449	-1.28	0.2009
Baseline PsA Duration	β_4	-0.0011	0.0026	-0.4	0.6890
Whether on MTX (1=yes)	β_5	0.0095	0.0433	0.22	0.8269
Whether a ACR20-50 responder	β_6	-0.3747	0.0636	-5.89	<0.0001
Whether a ACR50+ responder	β_7	-0.6149	0.0532	-11.55	<0.0001

Table 5.5: PASI at 24-weeks regression coefficients

Description	Covariate	Parameter estimate	Standard error	t Value	Pr > |t|
Intercept	α	0.3759	0.2473	1.52	0.1312
Baseline PASI	β_1	0.8096	0.0675	11.99	<.0001
Baseline Age	β_2	0.0006	0.0039	0.15	0.8835
Gender (1= Male)	β_3	0.0606	0.0955	0.63	0.5271
Baseline PsA Duration	β_4	0.0014	0.0051	0.28	0.7819
Whether on MTX (1=yes)	β_5	0.0190	0.0946	0.2	0.8416
Whether a PASI50-75 responder	β_6	-0.9128	0.1396	-6.54	<.0001
Whether a PASI75+ responder	β_7	-2.1843	0.1014	-21.55	<.0001

This allowed more accurate matching of the different patient demographic variables from each individual trial, under the assumption that the selected response measures discriminate all of the change in HAQ and PASI and that the baseline patient characteristics of the ADEPT trial are representative of the population of interest.

Given that the adjusted response data was estimated at 24-weeks, the magnitude of response in the initial HAQ and PASI change is based on the estimated 6-month gain

but weighted dependent on results at 12-weeks observed in the ADEPT trial (see parameters table 6.2.6.1. in original submission, p68; Table 5.3. above).

5.3.4 HAQ / PASI progression

The model assumes there is no progression in HAQ whilst a patient is responding to anti-TNF agents, based on open label results (88 weeks follow-up) for adalimumab[31-33] and similar results from an open label study of etanercept (no reference provided). It is assumed that patients continue to experience a lack of progression as long as they continue to respond and remain on treatment with any anti-TNF agent. However, in order to reflect the fact that conventional DMARDs are not considered as efficacious, the manufacturer assumes a mean annual HAQ progression of 0.07 (SD 0.03) whilst patients are responding to DMARDs, based on data taken from a sample of a cohort of patients based at the Academic Unit of Musculoskeletal Disease, University of Leeds.[45] An alternative study was also used in a sensitivity analysis.[48] It is assumed that psoriasis is only symptomatic and not progressive so PASI progression is not modelled.

5.3.5 Rebound effect

The extent of HAQ rebound after treatment failure is assumed to be of the same magnitude as the initial gain (i.e. rebound equal to gain) and the HAQ worsening occurs immediately at the point of withdrawal. For PASI progression it is assumed that psoriasis is only symptomatic and not progressive, hence the PASI rebound is always to the starting level.

5.3.6 Withdrawal rates

Withdrawal is analysed in terms of initial response to treatment and long term withdrawal. Following BSR recommendations on treatment continuation,[26, 39] treatment failure is analysed in terms of 12-week PsARC response for the base-case analysis, although a number of sensitivity analyses are explored based on whether patients are PsARC and/or PASI75 responders at 12- and 24-weeks.

Given the limited follow-up from open label studies of the anti-TNF agents, estimates for long term withdrawal rates were based on evidence from two observational studies. The main analysis was based on the BIOBADASER study, a Spanish Biologics registry[46] which has evaluated the safety of all three anti-TNF agents over a four-year period. The majority of patients in the BIOBADASER registry have RA, although 10% of patients are reported to have PsA. The registry was used to

estimate an average withdrawal rate for all three anti-TNF agents, using a Weibull survival model. The impact of applying separate withdrawal rates for each anti-TNF agent was assessed in a sensitivity analysis. A second study by Flendrie et al.[48] reporting on the withdrawal rates for all three anti-TNF agents was used as an additional sensitivity analysis.

5.3.7 Health-related quality of life

The impact of psoriasis on health-related quality of life (HRQL) was considered in conjunction with the impact of the arthritis component of the disease, using the PASI and HAQ scores, respectively.

The ADEPT trial included a generic measure of HRQL, the SF-36. Using the patient-level data from the adalimumab trial, the SF-36 responses were used to derive utility values via the SF-6D using the Brazier algorithm.[49] In order to be able to discriminate between the more severe PsA patients and given concerns relating to the "floor effect" associated with the SF-6D, an alternative utility measure, the EQ-5D was also estimated from SF-12 responses following the methods described by Gray et al.[50] Although not explicitly described in the manufacturers submission (only appearing as a data source in Table 6.2.6.1. of the submission), an additional study reporting SF-36 data on adalimumab in patients with psoriasis was combined with the ADEPT study data prior to the application of these algorithms.[51]

Patient level data from these two sources were used to assess the relationship between utility (using both SF-6D and EQ-5D results), PASI and HAQ scores using regression analysis. Separate regressions were undertaken to estimate utility for patients with and without active skin involvement (BSA <3% and >3%).Health utilities for each treatment and for patients (i) with and (ii) without skin disease were calculated using the following functions:

(i) $\text{Utility(skin)}t = \alpha + \beta 1\text{HAQ}t + \beta 2\text{PASI}t$

(ii) $\text{Utility(no skin)}t = \alpha + \beta 1\text{HAQ}t$

Generalized estimating equations (GEE) methods were used in the regression analyses due to the correlated nature of the data. Results for the regressions using SF-6D and EQ-5D are reported in Tables 5.6 - 5.9 below. The regression results

based on EQ-5D were used in the base-case analysis and sensitivity analyses were undertaken using the SF-6D regression results.

Table 5.6: SF-6D utility regression incorporating HAQ and PASI variables

Variable	Parameter estimate	Standard error	Z	Pr > |z|
Intercept	0.7803	0.0088	88.85	<.0001
HAQ	-0.1005	0.0078	-12.91	<.0001
PASI_t	-0.0174	0.0037	-4.71	<.0001

PASI_t = Transformed PASI log(PASI+0.5)

Table 5.7: SF-6D utility regression incorporating only HAQ

Variable	Parameter estimate	Standard error	Z	Pr > |z|
Intercept	0.7603	0.0046	163.6	<.0001
HAQ	-0.1036	0.0049	-20.94	<.0001

Table 5.8: EQ-5D utility regression incorporating HAQ and PASI variables

Variable	Parameter estimate	Standard error	Z	Pr > |z|
Intercept	0.9144	0.0186	49.09	<.0001
HAQ	-0.2512	0.0189	-13.30	<.0001
PASI_t	-0.0355	0.0096	-3.70	0.0002

Table 5.9: EQ-5D utility regression incorporating only HAQ

Variable	Parameter estimate	Standard error	Z	Pr > |z|
Intercept	0.8624	0.0146	59.01	<.0001
HAQ	-0.2406	0.0207	-11.64	<.0001

5.3.8 Resource utilisation and costs

Resource utilisation and costs for the following three components were considered in the base-case analysis: (i) drug acquisition and associated monitoring/administration costs; (ii) acute care and hospitalisation costs for the arthritis component of the disease and (iii) costs due to psoriasis symptoms. An additional analysis of the direct non-medical and indirect costs was also undertaken but did not form part of the base-case analysis (further details reported on p81 of the manufacturer's submission). Resource utilisation was based on published literature, national databases and expert opinion. Unit costs were obtained from routine NHS sources and published literature.

Drug costs included the cost of drug acquisition, administration and monitoring based on the recommended dosages and vial prices by the Monthly Index of Medical Specialties[52] and BSR recommendations on specific monitoring requirements for all treatments.[26, 39] Resource use related with the administration of the different treatments seem to have been based on the corresponding SPCs and expert opinion. The unit cost of physician visits (Rheumatology consultant and GP) were estimated based on the PSSRU database.[53] The cost of a half day inpatient visit per infliximab infusion was estimated based on NHS Reference Costs[54] and the cost of laboratory tests were based on a recent published study.[55] The annual drug acquisition and monitoring costs for the first and subsequent years are shown in Table 5.10 below.

Table 5.10: Annual drug and monitoring costs

Drug	Drug cost (First year/Subsequent Years)	Monitoring cost (First year/Subsequent Years)
Etancercept	£9295/£9295	£423/£360
Adalimumab	£9295/£9295	£423/£360
Infliximab	£13847/£10910	£2526/£2248
Conventional DMARD	£238/£236	£765/£712
Methotrexate	£179/£179	£800/£737

The direct acute care and hospitalisation costs associated with the arthritis component of the disease were estimated as a function of patients' HAQ scores based on a published study[56] on the relationship in rheumatoid arthritis between HAQ and the costs of treatment. An alternative study was used in a sensitivity analysis,[57] although no further details were provided on the approach employed in this study.

The costs due to the psoriasis symptoms were estimated from a physician survey conducted to estimate typical levels of resource utilisation for a 6-month period for patients with different levels of psoriasis severity given by the PASI score. Details of the psoriasis resource utilisation survey are provided in Appendix 6 of the original submission and results presented in Table 5.11 below.

Table 5.11: Results of psoriasis resource utilisation survey

PASI score range	Median Cost (2004 £GBP)
PASI state 1: score=1.5 (1.5, 2.7)	£168.92
PASI state 2: score=9 (7, 11.2)	£931.74
PASI state 3: score=15 (12.6, 16.8)	£817.06
PASI state 4: score=40 (32.4, 43.2)	£1035.27

Only NHS direct costs were considered for the base-case analysis and all costs were adjusted for inflation and stated in 2004 GBP prices. Costs were discounted using an annual rate of 3.5%.

5.4 Critique of the manufacturer's economic evaluation

The ERG has considered the methods applied in the manufacturer's economic evaluation in the context of the critical appraisal questions listed in Table 5.12 which are drawn from common checklists for economic evaluation methods.[58]

Table 5.12: Critical appraisal checklist

Item	*Critical Appraisal*	*Reviewer Comment*
Is there a well defined question?	Yes	The manufacturer assessed the cost-effectiveness of adalimumab for those patients with active and progressive PsA who have responded inadequately to previous DMARDs.
Is there a clear description of alternatives?	Yes ?	The relevant comparators to adalimumab considered other alternative anti-TNF agents (etanercept and infliximab) and a "traditional DMARD" option; after treatment failure these initial treatments were followed by a sequence of up to 5 alternative DMARDs. Palliative care is only an option after all DMARDs have been exhausted. The definition of what constitutes "palliative care" and how it differs from the option "traditional DMARD" in terms of efficacy and costs estimates is not provided.
Has the correct patient group/ population of interest been clearly stated?	Yes ?	The population of interest is those patients with active and progressive PsA who have responded inadequately to previous DMARDs. However, the type of previous DMARDs failed before in the model is not clearly specified.

Is the correct comparator used?	Yes ?	The comparison with other anti-TNF agents is appropriate. However, the interpretation of the "traditional DMARD" treatment option in the way that the manufacturer has estimated its efficacy (i.e. the result of a meta-analysis of the placebo plus methotrexate arms response rates from the anti-TNF trials was assumed to be equivalent to treatment efficacy for the traditional DMARD option) is not clear.
Is the study type reasonable?	Yes	Cost-effectiveness analysis.
Is the perspective of the analysis clearly stated?	Yes	The model estimates costs from the perspective of the UK NHS, and health outcomes in terms of Quality-Adjusted Life Years (QALYs).
Is the perspective employed appropriate?	Yes	The manufacturer's submission adopts a UK NHS perspective for costs, although they fail to take account of costs to PSS, so it is only partially consistent with the NICE reference case. Perspective on outcomes is that of the patient with treatment health effects to the individuals being captured by QALYs.
Is effectiveness of the intervention established?	?	Indirect comparison methods are used to obtain the relative efficacy of adalimumab, etanercept, infliximab and the "traditional DMARD" option and to incorporate the impact of the alternative treatments on both skin and joint disease. There are concerns about the accuracy of the efficacy estimates included in the analysis, given the lack of transparency and complexity of the methods used and the strong assumptions required.
Has a lifetime horizon been used for analysis (has a shorter horizon been justified)?	Yes	A lifetime horizon has been used in the model.
Are the costs and consequences consistent with the perspective employed?	Yes	Costs are consistent with a NHS perspective, although there are some concerns over the way the resource utilisation for the treatment of psoriasis were estimated. Consequences are measured in QALYs.
Is differential timing considered?	Yes	Future costs and health outcomes were discounted at an appropriate rate.
Is incremental analysis performed?	Yes	
Is sensitivity analysis undertaken and presented	Yes	The uncertainty in model parameters was characterised using probabilistic

clearly?		sensitivity analysis. A number of sensitivity analyses of key parameters, including utilities and HAQ progression whilst responding to DMARDs (see Table 6.2.11.1 p.84-85 of the submission) were undertaken and the results clearly presented, (Section 6.3.3, p88 – 94; table 6.3.3.1)

Table 5.13 below compares the manufacturer's submission to that of the NICE reference case.

Table 5.13: NICE reference case checklist

Attribute	*Reference Case*	*Included in submission*	*Comment on whether de-novo evaluation meets requirements of NICE reference case*
Comparator(s)	Alternative therapies including those routinely used in NHS	Yes ?	The comparison with other anti-TNF agents is appropriate. However, the meaning of the "traditional DMARD" treatment option in the way that the manufacturer has estimated its efficacy is not clear.
Perspective - costs	NHS and PSS	?	NHS costs have been taken into account but no consideration of PSS costs was undertaken.
Perspective - benefits	All health effects on individuals	Yes	QALY benefits to treated individuals were considered.
Time horizon	Sufficient to capture differences in costs and outcomes	Yes	The economic model has a lifetime time horizon. Alternative time horizons are also explored.
Synthesis of evidence	Systematic review	No	Only 3 out of the 8 identified RCTs were included in the analysis, on the basis that studies of less than 6-months duration may potentially underestimate the true efficacy of anti-TNF agents. There are concerns that the manufacturer might have excluded relevant trial evidence.
Outcome measure	QALYs	Yes	Two different utility measures were considered: EQ-5D for the base-case analysis, SF-6D for the sensitivity analysis.
Health states for QALY measurement	Described using a standardised and validated instrument	Yes	Using the patient level data from the adalimumab ADEPT trial, the SF-36 responses were used to derive utility via the SF-6D using the Brazier algorithm. In order to be able to

			discriminate between the more severe PsA an alternative utility measure, the EQ-5D was also estimated from SF-12 responses.
Benefit valuation	Time Trade Off or Standard Gamble	?	N.A.
Source of preference data	Sample of public	?	N.A.
Discount rate	Health benefits and costs	Yes	Benefits and costs have both been discounted at 3.5%.
Equity	No special weighting	Yes	No special weighting was undertaken.
Sensitivity analysis	Probabilistic sensitivity analysis	Yes	Probabilistic sensitivity analysis was undertaken. Results presented graphically using cost-effectiveness acceptability curves (CEACs).

5.5 Detailed critique of evaluation methods

As a general comment, the ERG noted a lack of transparency relating to both the description of the methods in the manufacturer's report and in the implementation of the cost-effectiveness model. The description of the economic model in the manufacturer's submission lacked clarity (e.g. the indirect treatment comparison methods section) and the use of some relevant trial evidence was not mentioned in the text (i.e. use of additional trial evidence to estimate EQ-5D for patients with skin disease, only stated in the parameters table 6.2.11.1). The model itself was considered complex and difficult to follow, particularly given the amount of hidden columns and data contained within the spreadsheet and associated macros. A critical review of the methods used in the manufacturer's economic evaluation has been undertaken, using the previous checklists as a basis for the review.

5.5.1 Evidence synthesis methods

We have previously described the general approach used by the manufacturer to synthesise the effectiveness data for adalimumab, etanercept, infliximab and DMARD therapies required to inform the cost-effectiveness analysis. Given the lack of head-to-head RCT data on the relevant treatments, an indirect comparison was undertaken by the manufacturer. The methods and assumptions employed in the indirect comparison are now critiqued in order to outline the key areas of uncertainty identified by the ERG.

In general the ERG found that the methods employed by the manufacturer lacked transparency which made it difficult to assess the validity of the findings. Many of the approaches were insufficiently explained in the manufacturer's submission and, as such, the critical review required detailed interrogation of the electronic model itself. The ERG felt that the general approach was overly complex employing a number of assumptions which increase the possibility of significant bias in the subsequent results. The key issues identified by the ERG include:

- The inclusion/exclusion criteria applied by the manufacturer in selecting studies for the indirect synthesis;
- the assumption of exchangeability of response rates after adjustment for the number of patients with psoriasis at baseline;
- the approach used to estimate correlation between response parameters;
- the adjustment used by the manufacturer to estimate 12-week response parameters from 24-week trial results.

Each of the areas is considered in more detail below, outlining the key assumptions and the potential uncertainties surrounding them.

The inclusion/exclusion criteria applied by the manufacturer in selecting studies for the indirect synthesis

The ERG was concerned with the secondary filter (Section 5.6, p47 of the manufacturer's submission) employed by the manufacturer to select studies for the indirect evidence synthesis. This filter applied a number of additional exclusion criteria to the initial set of studies identified as part of a wider systematic search of anti-rheumatic therapies in PsA. From the initial 8 trials identified (comprising 9 studies in total with 1 trial reporting 2 separate publications for different follow-up periods) only 3 were subsequently included as part of the manufacturer's evidence synthesis. Trials were only included if they met the following criteria:

1. Response data for PsARC, ACR and PASI was reported;
2. Trial duration of at least 6 months;
3. Disease duration > 8 years;
4. DMARD therapy had been attempted at least once.

Of the 5 trials (6 studies) excluded, 3 were of alternative DMARD therapies (leflunomide, cyclosporine and sulfasazine) compared to placebo[59-61] and two were of trials of adalimumab and infliximab again compared to placebo. [32, 62] In addition, the 12-week results for the Mease et al study[38] comparing etanercept and placebo were excluded (although the 24-week trial results from this trial were included). The 3 DMARD trials were excluded based either on disease duration <8 years or on the basis that DMARD therapy had not been attempted at least once before. The 3 studies of alternative anti-TNF agents compared to placebo were all excluded on the basis that the study results were reported at a follow-up of less than 6-months (all 3 trials reported results at 12-weeks). While the ERG considers the exclusion of trials based on previous DMARD therapy and disease duration to be defensible (although the choice of 8 years duration appears potentially arbitrary and no supporting references are provided to support this particular assumption), the exclusion of trials with a shorter duration than 6-months is a particular concern, resulting in the apparent exclusion of relevant evidence. The manufacturer justified the decision to exclude 12-week studies on the basis that these data may potentially underestimate the true efficacy of anti-TNF agents. However, no supporting references were provided by the manufacturer to justify this particular assumption. Instead there was simply a comment (p47 of the manufacturer's submission) stating that evidence of anti-TNF agents in polyarthritic conditions suggests that initial response to therapy plateaus at 6-months and, therefore, using 12-week data may underestimate the true efficacy of anti-TNF agents.

The decision to exclude trials of 12-week follow up appears to contradict current BSR guidelines[26, 39] and the recent NICE appraisal of etanercept and infliximab[23] which recommend a 12-week period as sufficient to assess response to treatment and to inform the decision to continue a particular treatment. Discussions with our own clinical advisors strongly supported the view that patients who had not responded by 12-weeks were unlikely to respond over a longer duration, such that 12-weeks was considered an adequate period with which to assess treatment response.

In the manufacturer's response to the ERG initial points for clarification, further supplementary evidence was provided examining ACR20 response rates over time for patients in the adalimumab arms of studies M02-518 and M02-570 (see Figure 5.2 below). The manufacturers noted a delayed response in the rate of improvement in arthritis during the first 12-weeks of treatment with adalimumab in study M02-570 compared to the response rates at 24-weeks (M02-518 ADEPT trial). The

manufacturer stated that a similar delay in response was also observed in other response parameters including ACR50, ACR70, PsARC and HAQ-DI (although no supporting data were provided on these parameters). The manufacturer cites these data as additional justification for the decision to exclude 12-week trial results from their main analysis. However, as the manufacturer notes, the results from study M02-570 beyond 12-weeks are subject to a number of potential confounding factors, since results from this study beyond 12-weeks were part of the subsequent open-label phase. Despite some exploratory subgroup analysis, which proved inconclusive, the manufacturer fails to provide an adequate explanation for the apparent delayed effect observed in study M02-570. In addition, it should be noted that the same findings were not observed in the M02-518 study during which randomisation status was maintained at both the 12 and 24-week follow up. Closer examination of the ACR20 response rates at 12 and 24-weeks reveal that response rates at these two time periods were virtually identical.

Figure 5.2 : ACR20 response rates (observed) over time for patients in the adalimumab arms of studies M02-518 and M02-570. (Treatments were open label in M02-570 after Week 12, and in M02-518 after Week 24)

Figure 5.2: Removed

The ERG does not believe that the manufacturer has provided a justifiable rationale to support their decision to exclude the 12-week trial results, thus excluding potentially relevant evidence. Furthermore, the exclusion of the 12-week trial data appears to contradict current guidance on the management of PsA in which response at that time point is considered the main factor in deciding whether a patient should be continued on a particular treatment or switched to an alternative agent. Indeed, decision rule for treatment continuation employed by the manufacturer in the base-case cost-effectiveness analysis is actually based on a 12-week period (PsARC

response at 12-weeks). By excluding the actual 12-week trial data the manufacturer is then forced into making an unnecessary assumption in order to adjust 24-week trial data to inform the 12-week decision point (this particular issue is discussed in more detail in the following sections). It is clear that the exclusion of these data may introduce possible bias into the manufacturer's results such that the associated cost-effectiveness results should be treated with particular caution.

The assumption of exchangeability after adjustment for the number of patients with psoriasis at baseline

We have previously described the manufacturer's concerns relating to the applicability of conventional approaches to evidence synthesis using indirect and mixed treatment comparisons in this area. These were based largely on the concern that these approaches would not be able to adequately address the heterogeneity in response across the trials, caused by the different number of patients with active psoriasis BSA <3% and >3% (meta-regression was not considered possible by the manufacturer due to the small number of trials) and the need to consider multiple outcomes related to both arthritis and psoriasis outcomes. However, it should be noted that the methods subsequently employed by the manufacturer rely on a series of strong assumptions which the ERG does not feel have been adequately justified. In particular, the assumption of exchangeability, after adjustment for the proportion of patients with active psoriasis, is central to the validity of the manufacturer's approach. In essence, this approach assumes that the absolute response rates for individual treatments (as opposed to the relative treatment effects between individual treatments) can be assumed to hold across all studies and that the correlation between response parameters from a single study of adalimumab is applicable to all other treatments considered. By using the absolute response rates in the evidence synthesis, as opposed to the relative effects, the approach breaks randomisation and hence is prone to potential bias.

There are two key aims of the manufacturer's evidence synthesis: (i) to adjust for potential confounding between trials based on the different number of patients with active BSA <3% and >3% and (ii) to correlate response parameters for both arthritis and psoriasis measures. These appear to be largely based on the manufacturer's claim (p47 of the manufacturer's submission) that patient-level analysis from the ADEPT (M02-518) study showed a high correlation between response variables and, in particular, that there appeared to be a difference in efficacy of the arthritis component of the disease for the different subgroups with <3% and >3% BSA

affected by psoriasis. However, no data were reported in the original submission to justify this claim. Further subgroup analysis submitted as part of the manufacturer's response to initial points for clarification raised by the ERG provided a comparison of the response rates for ACR and PsARC based on these different subgroups. These are reported in Table 5.14 below.

Table 5.14: M02-518 - Response Rates at Weeks 12 and 24

12-week response rates from M02-518 (ADEPT) trial

Response	Placebo		Adalimumab	
	Psoriasis <3% BSA	Psoriasis >=3% BSA	Psoriasis <3% BSA	Psoriasis >=3% BSA
ACR20	15%	13%	56%	60%
ACR50	4%	3%	37%	36%
ACR70	1%	0%	21%	19%
PsARC	26%	26%	62%	61%

24-week response rates from M02-518 (ADEPT) trial

Response	Placebo		Adalimumab	
	Psoriasis <3% BSA	Psoriasis >=3% BSA	Psoriasis <3% BSA	Psoriasis >=3% BSA
ACR20	16%	14%	60%	53%
ACR50	5%	6%	42%	36%
ACR70	2%	0%	22%	23%
PsARC	26%	20%	64%	56%

No formal statistical analysis was undertaken by the manufacturer to establish the statistical significance of the difference in response rates between the subgroups. Closer examination of the response rates appears to show that the differences do not appear to be particularly large and, more importantly, no clear trend appears in either the magnitude or the direction of the results between the 12 and 24-week periods. Furthermore, in the absence of a formal statistical analysis, it is not possible to determine whether the observed differences are simply due to chance or not. The need to adjust for the proportion of patients with different levels of skin involvement to provide a fairer comparison between the trials does not appear to have been adequately justified.

Given that there does not appear to be a marked difference in the response rates between the different subgroups, the ERG did not consider that the adjustment made by the company for the different levels of skin involvement across the trials (ranging from 43% to 87% across the trials) would fundamentally alter the response rates from

the trials themselves. To explore this issue in more detail the ERG compared the original trial results to the adjusted response rates based on the optimisation algorithm used by the manufacturer. The adjustment was made based on a pooled estimate of 66% of patients with BSA >3% from the anti-TNF treatment arms of the three trials (ranging from 43% to 87% across the individual trials). The results of this comparison are presented in Table 5.15. As expected, the response rates were remarkably similar with only the results for adalimumab appearing to change marginally from the unadjusted estimates from the original trials. However, it should be noted that the response rate for PsARC was marginally higher (62% vs 60%) using the optimisation software for adalimumab (etanercept and infliximab appeared to remain unaffected). So, although the adjustment appears to have limited effect overall, it does appear to marginally improve the PsARC response estimates for adalimumab, which will improve the cost-effectiveness estimates for this treatment compared to the unadjusted estimates.

Table 5.15: ERG comparison of response rates using optimisation software with trial estimates

Trial name	Treatment	Source	ACR20	ACR50	ACR70	PASI50	PASI75	PASI90	PsARC
Mease	Etanercept	Trial	50%	37%	9%	47%	23%	6%	70%
	Etanercept	Solver	50%	37%	-	47%	23%	-	70%
ADEPT	Adalimumab	Trial	57%	39%	23%	74%	59%	42%	60%
	Adalimumab	Solver	57%	38%	-	75%	59%	-	62%
IMPACT II	Infliximab	Trial	54%	41%	27%	75%	60%	39%	70%
	Infliximab	Solver	54%	41%	-	75%	60%	-	70%

The approach used to estimate correlation between response parameters

The second key area addressed by the manufacturer's synthesis was the need to correlate the different response parameters. The issue of correlation was particularly important since each response parameter plays a distinct role within the manufacturer's cost-effectiveness model, with treatment continuation being determined by PsARC response and ACR and PASI response data being used to estimate utility and costs. Given the structural link relating the three main response parameters in the model, it was clearly important to be able to consider the correlation between them (i.e. in order to obtain the ACR and PASI responses conditional upon PsARC response status). In the absence of patient level data from the IMPACT II and Mease studies, the manufacturer used the observed data from the M02-518 study to predict the correlations for the other treatments.

The manufacturer used optimisation techniques (using Microsoft Solver) to predict the correlations for the other treatments based on the M02-518 study. The objective was to move from aggregate data for each of the response parameters (i.e. ACR20, PASI50 etc for all patients) to a much more detailed matrix based on a series of related response parameters (i.e. ACR20 and PASI 50 response assuming a PsARC response etc), allowing the correlation between response parameters to be factored into the economic model. Microsoft Solver was used to predict the response rates for the non-adalimumab strategies by employing a series of logical constraints and attempting to minimise the difference across five separate components (including the correlation observed in the M02-518 study, the overall PASI, PsARC and ACR responses from each of the non-adalimumab treatments).

In the absence of patient level data for the non-adalimumab strategies it is difficult for the ERG to validate the subsequent estimates predicted by Microsoft Solver. As we have previously shown a comparison of the aggregate responses from the trials and those the ERG calculated (by aggregating the correlated response parameters) shown in Table 5.15 demonstrated that there were only minor differences between the predicted results and those from the trials themselves. However, it is not possible to establish whether the correlated results themselves are valid or not. Given that these were informed by the correlations observed in the M02-518 study it is not clear whether these are generalisable to other non-adalimumab treatments.

While the ERG acknowledges the importance of correlating response parameters based on the structural relationship imposed by the model on the alternative response parameters, the ERG has significant concerns over the major assumption required by the manufacturer to implement their approach. After adjusting for skin involvement, the manufacturer assumes that the results from the different trials can be treated as if they come from a single study. The approach to correlating response parameters for the different treatments uses the response data for the treatment arms from the different trials without any adjustment for the control group response rates (an approach akin to pooling across single arms of trials – i.e. breaking randomisation). This approach assumes that the absolute response rates for each of the different agents are fully exchangeable across the different studies. While it is common to assume that trials are sufficiently homogeneous that the *relative* difference between treatments can be considered exchangeable across studies, the assumption that the *absolute* event rates can be treated in this manner requires a far stronger assumption. The manufacturer cites similar placebo response results from

the three main studies included in their synthesis as justification for this assumption (p45 of the manufacturer's submission). However, the ERG does not consider that this is adequate justification particularly since differences between the placebo PsARC response rates range from 23-32% across the studies. Without adequately controlling for the differences in the underlying placebo response rate across the studies it means that subsequent response rates employed in the model are prone to potential bias.

***The adjustment used by the manufacturer to estimate 12-week response parameters from 24-week trial results*.**

The final area of concern to the ERG is the adjustment made by the manufacturer to estimate 12-week response data. We have previously highlighted that the manufacturer excluded 12-week trial data from their analysis. However, since the decision to continue treatment in the base-case analysis is based on PsARC response at 12-weeks, the manufacturer is then forced into making an assumption in order to adjust the 24-week trial data to inform the 12-week decision to continue treatment. The adjustments made by the manufacturer for the main response parameters are summarised in Table 5.16 below (see Table 6.2.6.1 p68 of the manufacturer's submission).

Table 5.16: Adjustments used to estimate 12-week response parameters

Response parameter	Value
% PsARC responders at 6 months who were PsARC responders at 3 months	80
% ACR20 responders at 6 months who were ACR20 responders at 3 months	78
% ACR50 responders at 6 months who were ACR50 responders at 3 months	71
% PASI75 responders at 6 months who were PASI75 responders at 3 months	70

The adjustments themselves were based on results from the ADEPT M02-518 study for adalimumab. In the absence of similar estimates for the other treatments, the same adjustments were applied to all treatments. The ERG has a number of significant concerns regarding this approach:

- it ignores the response data from the 12-week trials themselves;
- it assumes that the relationship between response rates at 12 and 24-weeks observed for adalimumab applies to all other treatments;
- it forces the 12-week response rate to be lower than 24-week response rates.

In addition to the concerns noted above it is also unclear as to how the adjustments have actually been estimated. While the manufacturer states that these are based on the ADEPT M02-518, we have previously discussed the relationship between response parameters at different follow-up periods, highlighting that a comparison of the ACR20 response rates at 12 and 24-weeks revealed that response rates at these two time periods were virtually identical based on the ADEPT M02-518 study. It is therefore unclear what data have actually been used as the basis for these adjustments. It is likely that these data are based on the proportion of responders at 24-weeks who were also responders at 12-weeks. However, this does not appear to constitute an appropriate basis for estimating 12-week responses. The decision to continue with treatment would actually be made at 12-weeks and consequently the patients responding at this particular time point would be those continued on treatment. It is clear that the exclusion of 12-week trial data has significant implications not only for the inclusion of *all* relevant evidence but it would also appear necessary in order to accurately estimate the actual response at 12-weeks required for the base-case analysis undertaken by the manufacturer.

5.5.2 Choice of comparators

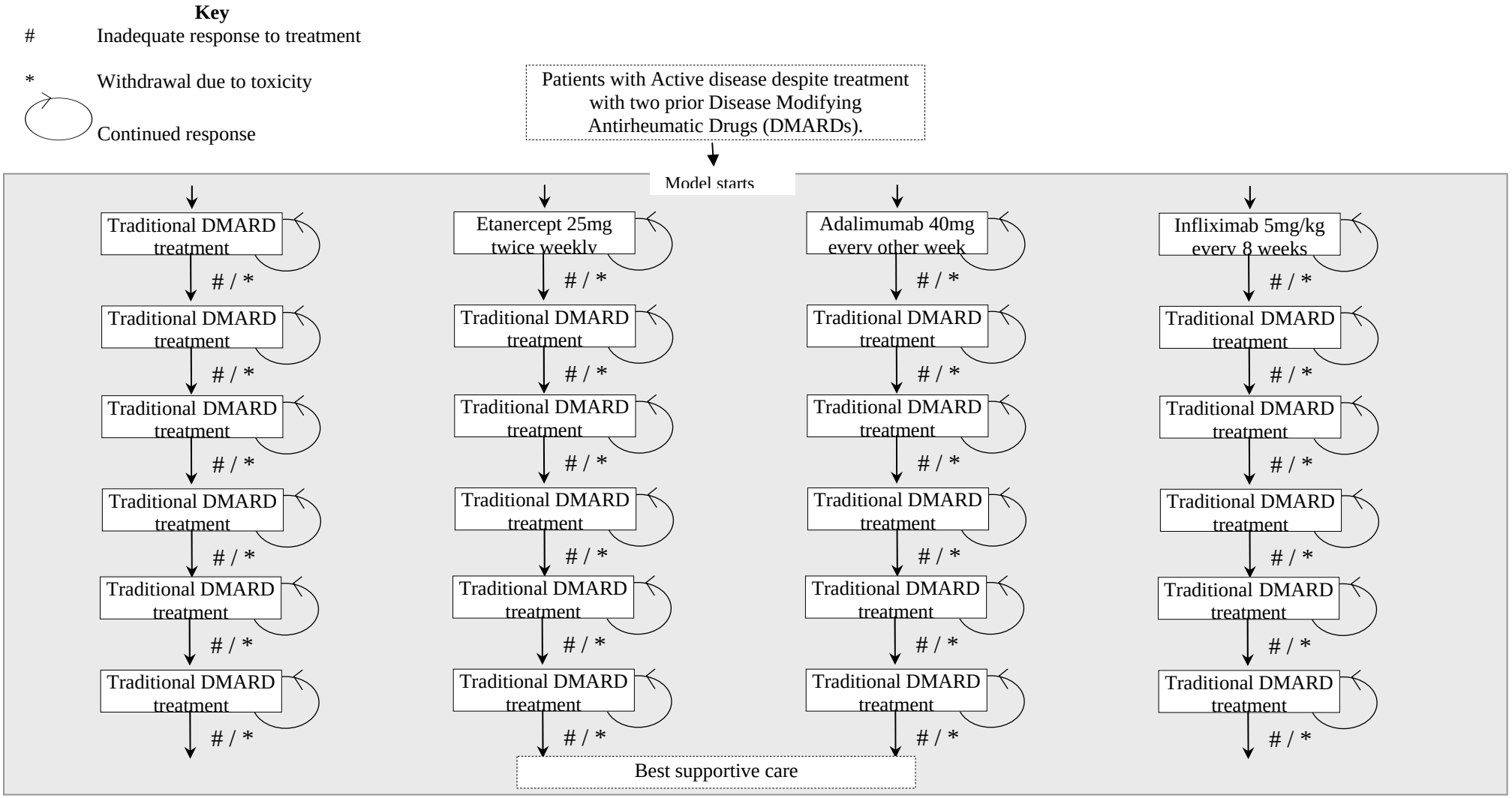
The manufacturer submission justifies the choice of "traditional DMARDs" as the comparator treatment to adalimumab, etanercept and infliximab on the basis that they are the most common alternative to anti-TNF therapies in the UK. The use of "palliative care" is only considered as a treatment option after all alternative DMARDs in the treatment sequence (up to 5 after failure of initial treatment) have been tried and failed, employing the assumption that in UK clinical practice rheumatologists are more likely to exhaust all conventional DMARD treatment options prior to administration of palliative care only. See Figure 5.3 for further details.

There are a number of concerns with the choice of comparators presented in the manufacturer's submission. Firstly, according to the BSR guidelines for the use of anti-TNF drugs for psoriatic arthritis[26, 39] at least two DMARDs individually or in combination should have been tried prior to the administration of biologics (i.e.

typically leflunomide, and either sulphasalazine or methotrexate) and, given the limited number of DMARDs routinely used in PsA compared to RA, the choice of alternatives with proven efficacy after the failure of anti-TNFs is in fact very restricted (i.e. the available limited data indicate some degree of efficacy for all DMARDs, but in some cases the evidence is particularly weak and may not be reliable, such as the case of IM gold and azathroprine[23]; the only option left may be cyclosporine). Secondly, a strict interpretation of the licenses of anti-TNF therapies would suggest that these should be used "end of line" once DMARD therapies have been tried and failed, and the use of a treatment sequence of up to 5 alternative DMARDs does not meet this interpretation.

While the ERG notes the points raised by the manufacturer, in which they highlight that a high proportion of patients in the control arm of the anti-TNF trials were receiving DMARD therapy, the ERG also considers that the manufacturer could have undertaken additional sensitivity analyses to explore the robustness of the base-case assumptions by assuming different proportions of patients (from 0 to 100%) would continue to receive DMARD therapy.

Figure 5.3: Treatment sequences



There are also a number of concerns regarding the precise interpretation of the "traditional DMARD" and "palliative care" options when it comes to their treatment efficacy and cost estimates applied in the model. Firstly, the manufacturer's submission is not explicit about which are the two DMARDs assumed to have failed before the hypothetical patient enters the model. Consequently, a common response estimate was assumed for the "traditional DMARDs" option based on the results of a meta-analysis of the placebo plus methotrexate arms response rates from the anti-TNF trials. A weighted average cost for all "traditional DMARDs" was then estimated based on the proportion of PsA patients on eight different DMARD therapies used at the University of Toronto database. However, even if the proportions and types of therapies used in this Canadian database were representative of NHS clinical practice, the final result is that the manufacturer's submission may potentially penalise unduly the DMARD options, adding the cost of drug acquisition, monitoring and administration (see table 5.17 and table 5.18 below) without considering their actual treatment efficacy.

5.17: Drug costs of each treatment

Treatment	Dose regimen	Unit cost	First Year	Subsequent years
Adalimumab	40mg; every other week	357.50	£9295	£9295
Etanercept	25mg; twice weekly	89.38	£9295	£9295
Infliximab (4 vials per infusion)	5mg/kg;0,2,6, then every 8	419.62	£13847	£10910
Infliximab (3 vials per infusion)	weeks 100mg vial		£10386	£8183
Methotrexate †	10-15mg weekly (2.5mg)	0.10 (oral), 10.08 (15ml injection)	£179	£179
Conventional DMARD	Weighted mix of DMARDs	-	£238	£236

†31% injection methotrexate (MTX), 69% oral MTX use taken from the University of Toronto dataset.

5.18: Monitoring and administration costs for each treatment

Treatment	First 12 weeks	Months 3-6	Subsequent 3 months	Months 0-12	Subsequent years
Etanercept	153.06	90.06	90.06	423.24	360.24
Adalimumab	153.06	90.06	90.06	423.24	360.24
Infliximab	839.56	562.06	562.06	2525.74	2248.24
Methotrexate	247.18	184.18	184.18	799.72	736.72
traditional DMARD	231.07	178.12	178.12	765.43	712.48

The ERG notes that the breakdown of the monitoring costs for the different DMARD therapies provided by the manufacturer indicates that a high proportion of resources are based on the use of leflunomide (5%), sulphasalazine (23%) and methotrexate (44%). However, it is clear that these three DMARDs will typically be used as first-

line or second-line before considering the use of biologics. In the model, these three DMARDs contribute a major proportion (72%) of the total weighted cost.for the average cost estimate applied to the conventional DMARD option.

Finally, the definition of what constitutes "palliative care" and how it differs from the "traditional DMARD" alternative in terms of efficacy and costs estimates is not provided. The ERG concludes that there are a number of potential uncertainties concerning the manufacturer's approach to modelling alternative strategies to the use of anti-TNF agents, including the assumptions made regarding the ongoing use of DMARD therapies and the effectiveness and cost-calculations applied therein.

5.5.3 HAQ progression

The manufacturer's submission assumes that there is no progression in HAQ whilst a patient is responding to adalimumab, based on evidence provided by an open label study (M02-537 trial). The same assumption is made for anti-TNF agents in general, as this evidence is assumed to be in line with open label data reported for etanercept. In contrast, the manufacturer assumes that HAQ will continue to progress whilst patients are responding to DMARDs, stating that: *"(…) conventional DMARDs are not as efficacious. The base-case analysis uses the estimate of HAQ progression whilst on DMARDs as per the York Model"* (p71). The authors refer here to the NICE review on the use of etanercept and infliximab for the treatment of active PsA, recently published as a HTA report.[23]

The York model used estimates of HAQ natural history progression for patients on palliative care taken from a sample of a cohort of patients based at the Academic Unit of Musculoskeletal Disease, University of Leeds.[45] According to the cohort patient characteristics, a mean annual progression rate of 0.07 (SD 0.03) was assumed for severe patients who were receiving palliative care in the York model (i.e. not DMARDs), after treatment failure with etanercept and at least 2 previous DMARDs (including leflunomide or ciclosporin), as used elsewhere.[63, 64] This figure is comparable with patients with severe RA[64] (HAQ = 0.066).

The ERG therefore considers that the use of this data to inform the annual HAQ progression in patients receiving DMARD therapy in the model may not be appropriate. The manufacturer's submission uses a mean HAQ progression whilst responding to DMARDs (0.07) which is 2.5 times that used in a recent published study on the cost-effectiveness of etanercept in patients with PsA.[45] Based on

available evidence (6), Bansback et al. used a HAQ progression rate of 0.028 for patients responding to DMARDs (leflunomide or ciclosporine). The manufacturer did, however, undertake a sensitivity analysis on the progression rate using the Toronto dataset, comprising a less severe population (mean annual HAQ progression 0.0085) than the Leeds cohort. The results from the sensitivity analysis demonstrated higher mean QALYs for all treatments and marginally lower costs, with adalimumab showing a higher ICER (£47,404 per QALY) with an associated probability of being cost-effective of 0.20 at a cost-effectiveness threshold of £40,000 per QALY. Consequently, the assumption on the annual progression in HAQ appears to be an important driver in the cost-effectiveness analysis.

In conclusion, the HAQ progression rate whilst responding to DMARDs used by the manufacturer's submission appears more appropriate for severe patients receiving palliative care or not responding to treatment. Consequently, the cost and QALY calculations (determined in part by HAQ scores) may potentially underestimate the cost-effectiveness of the DMARD strategy in the model; hence the relative cost-effectiveness of the anti-TNF agents may be overly optimistic.

A related issue for modelling HAQ progression concerns the logical constraint that exists based on the HAQ scoring system itself. The HAQ score itself focuses on two dimensions of health status: physical disability and pain, generating a score of between 0 (least disability) and 3 (most severe disability). In other words, the HAQ score is bounded between these two anchor points with the upper one being the most severe condition. After examination of the electronic model submitted by the manufacturer, the ERG noticed that the logical constraint did not appear to have been built into the manufacturer's calculations. As such, it was possible for patients to progress to HAQ scores beyond the upper bound of 3. Table 5.19 reports the percentages of HAQ scores > 3 and > 4 at different time points of the manufacturer's model calculated by the ERG.

5.19: Proportion of patients with HAQ >3.0 and HAQ > 4.0 (all treatments)

	Time horizon (years)				
	<30	**35**	**40**	**45**	**50**
HAQ > 3.0	0.00%	66.43%	78.83%	85.28%	86.63%
HAQ > 4.0	0.00%	0.00%	0.00%	0.00%	64.58%

The table shows that by 35 years more than half of the patients show a level of disability measured with HAQ over 3.0 (66%). This proportion of patients increases with longer time horizons. Furthermore, by 50 years a high percentage of patients have HAQ scores over 4.0 (65%). The ERG has also examined whether these proportion were evenly distributed or there were marked differences between treatments (see Table 5.20 below).

5.20: Breakdown of proportion of patients with HAQ > 3.0 by initial therapy

	Time horizon (years)				
	<30	**35**	**40**	**45**	**50**
Adalimumab	0.00%	58.70%	72.60%	81.00%	83.50%
Etanercept	0.00%	58.30%	72.80%	81.20%	83.20%
Infliximab	0.00%	55.50%	78.90%	78.90%	81.80%
Traditional DMARDs	0.00%	93.20%	100.00%	100.00%	98.00%

The table shows that there is an important difference between the HAQ scores of patients initially treated with traditional DMARDs (almost all of them exceed a HAQ score of 3 at 35 years) and results for the rest of anti-TNF agents, with infliximab showing the smaller proportion of patients with a HAQ > 3.0 at all time points. Given that HAQ scores are used subsequently to estimate costs and health outcomes in the model, the lack of bounding of HAQ at a score of 3 is a potential issue. Given that a higher proportion of patients in the DMARD option will experience this logical inconsistency, the subsequent estimates of the relative cost-effectiveness of the anti-TNF agents will be overly optimistic. However, the ERG also recognises that the lack of bounding does not become an issue in the model until after 30 years. Given that future years are discounted it is not envisaged that this inconsistency will have an important effect on the subsequent ICERs.

5.5.4 HAQ Rebound effect

It is commonly accepted that at withdrawal patients will experience some deterioration in HAQ (rebound). There are a number of possible rebound effects when a patient discontinues anti-TNF therapy but available trial data are too short-term to be able to characterise this accurately and there is no consensus among rheumatologists about the magnitude of this rebound.[65] A recent published study on the cost-effectiveness of etanercept and infliximab, identified and critically appraised in the manufacturer's submission,[23] considered two alternative rebound scenarios presented as limits, under the assumption that reality regarding rebound would be somewhere between them. In the rebound to natural history progression, the HAQ

after treatment failure returns back to the level and subsequent trajectory it would have been if the patient had not initiated treatment (i.e. conservative scenario, where anti-TNF agents only provide symptomatic relief). In the rebound equal to gain scenario, the patient's disability in terms of HAQ deteriorates by the same amount it improved when first responded to treatment (i.e. optimistic scenario, where anti-TNF agents can re-set the curve and delay disease progression). Cost-effectiveness results for the base-case analysis were shown to be sensitive to this structural assumption, with a difference in the ICER for etanercept between both rebound scenarios of about 40% (rebound back to natural history £16,801, rebound equal to gain £27,681, both for a lifetime horizon).

The manufacturer's submission presents base-case results under the assumption that the rebound effect *"… would be of the same magnitude as the initial (HAQ) gain"* (p72) and that the HAQ worsening occurs immediately at the point of withdrawal. The submission does not present a sensitivity analysis exploring the implications of a more conservative assumption of the rebound effect. Regarding the skin component of the disease, since it is assumed there is no PASI progression, the rebound is always to the starting level.

5.5.5 Long-term withdrawal rates

Estimates for long-term withdrawal rates were based on evidence from two observational studies. The base-case analysis was based on evidence provided by a Spanish Biologics registry,[46] whilst an alternative source, the Flendrie et al study[48] was used for the sensitivity analysis. The characteristics and differences of these studies were not discussed, and no justification for the choice of the Spanish registry as the source for the base-case analysis was provided. In addition, the reasons why the manufacturer's submission used an average withdrawal rate for all three anti-TNF agents instead of treatment specific rates were not made explicit either. The long-term withdrawal rates were subjected to sensitivity analysis by the manufacturer (see table 5.24 in the Results section below).The use of treatment specific withdrawal rates altered the mean QALY estimates, and although adalimumab remains the most cost-effective option (assuming a threshold of £30,000 per QALY), etanercept was no longer dominated, showing that this is a potentially important assumption in the ICER calculations.

5.5.6 Health-related quality of life

In the base-case anaysis, EQ-5D estimates, using UK population weights, were obtained via the SF-12 responses for both patients with and without active skin, based on analysis of individual patient data from the ADEPT trial. An additional study of adalimumab in psoriasis was used to estimate the utility scores for patients with psoriasis.[51] This is not made explicit in the text and is only reported in the table of parameters 6.2.6.1 (p68). The implications of the inclusion of this additional trial or its characteristics are not discussed by the manufacturer. There are some concerns relating to the severity of the patient population included in the Menter et al. study.[51] At baseline, patients with PsA showed a mean duration of psoriasis of 21 years, a % BSA affected of 28.7% and a PASI score of 17.6 (mean values). These results are indicative of a PsA patient with more severe chronic plaque psoriasis than the patient population considered in the model. As a point of reference, the percentage of patients with BSA ≥ 3% was only 46% in the adalimumab arm of the ADEPT trial, with a mean PASI score of 7.4 (± 6.1). Hence, the inclusion of this additional study may overstate the independent importance of psoriasis in the model.

Sensitivity analyses were also undertaken using similar regressions estimated using results from the SF-6D. As previously described the results of the SF-6D and EQ-5D regression (both for patients with and without skin disease) have been shown to produce quite different coefficient estimates (see table 5.6 and 5.9 above) in the regressions linking HAQ and PASI to utility scores. Using results from the SF-6D in the sensitivity analysis increases the ICER for adalimumab to £62,000 per QALY, compared to DMARDs, although etanercept remains dominated. The choice of whether to use EQ-5D or SF-6D utility estimates therefore appears to be an important factor in establishing the cost-effectiveness of adalimumab. The ERG did not feel that the differences between the EQ-5D and SF-6D estimates were sufficiently discussed by the manufacturers. While the manufacturer referred to the floor effect of SF-6D, such that utility values below 0.3 are not possible, the ERG was unsure as to whether this was the major factor determining the differences between the utility scores for the two algorithms.

In order to understand the reasons for these differences, the ERG requested some additional data from the manufacturer, such as the mean (SE) utility values for the ADEPT trial at week 12 and week 24. Table 5.21 below shows how the mean utility scores are well above the “floor point” for SF-6D (0.3) and even at 12 weeks, the difference between EQ-5D and SF-6D mean change scores for PsARC responders

and non responders was quite different (0.17 using EQ-5D measure, 0.06 using SF-6D, see Table 5.22 below). After consideration of this additional data, the ERG does not believe that differences between EQ-5D and SF-6D estimates can be explained wholly by the "floor effect".

Table 5.21: Utility Values for M02-518

Characteristic	Adalimumab N = 151	Placebo N = 162
Baseline : mean (SE)		
EQ5D	0.61(0.02)	0.59(0.02)
SF6D	0.67(0.01)	0.65(0.01)
Mean change (SE) at Week 12		
EQ5D	0.14(0.02)	0.03(0.02)
SF6D	0.04(0.01)	0.01(0.01)
Mean change (SE) at Week 24		
EQ5D	0.14(0.02)	0.05(0.02)
SF6D	0.06(0.01)	0.01(0.01)

Table 5.22: Change in utilities based on PsARC response at Week 12

Study	Characteristic	PsARC responders	PsARC non-responders
M02-518	Mean change (SE) at Week 12		
	EQ5D	0.17 (0.02)	0.01 (0.02)
	SF6D	0.06 (0.01)	0.00 (0.01)
M02-570	Mean change (SE) at Week 12		
	EQ5D	0.14 (0.03)	0.01 (0.04)
	SF6D	0.04 (0.01)	0.01 (0.01)

5.5.7 Psoriasis resource utilisation

In addition to assessing the potential impact of psoriasis on health utility, the manufacturer attempted to estimate the potential resource use and costs that could be attributed to this aspect as well. We have previously outlined that expert opinion was used to estimate the potential resource use associated with the management of psoriasis. In the critique of this component the ERG identified a number of potential areas of uncertainty. Firstly, the classification of psoriasis severity in the form of four representative patients (Appendix 6; see table 5.23 below) was not justified, and the relevance and appropriateness of this type of classification taking into account the baseline characteristics of the three anti-TNF trials included and the population of interest of this study are not discussed. Each of the four representative patients had BSA >5% (ranging from 5%-60%). Given that the classification applied in the model relates to BSA >3%, it is unclear how generalisable these resource estimates are to this particular population.

Table 5.23: Description of the four representative patients used in the psoriasis resource utilisation questionnaire

#1	Severe symptoms of psoriasis on approximately 60% of the body, covering part of the head and most of the trunk and lower extremities
#2	Moderate symptoms of psoriasis on approximately 10% of the body, covering some of the upper extremities and partially covering the lower extremities.
#3	Severe symptoms of psoriasis on approximately 20% of the body, partially covering the upper extremities and covering some of the trunk
#4	Moderate symptoms of psoriasis on approximately 5% of the body, partially covering the head

In addition, details of the physician survey used to estimate resource use associated with the treatment of psoriasis are not adequately described in the manufacturer's submission (e.g. size and representativeness of the sample of physicians, non-response rate etc.). Resource use and unit costs were not reported separately and results of the survey were not validated. See Table 5.24 below for further details.

Table 5.24: Results of psoriasis resource utilisation survey

PASI score range	Median Cost* (2004 £GBP)
PASI state 1: score=1.5 (1.5, 2.7)	£168.92
PASI state 2: score=9 (7, 11.2)	£931.74
PASI state 3: score=15 (12.6, 16.8)	£817.06
PASI state 4: score=40 (32.4, 43.2)	£1035.27

*All costs based on 6-month period.

Finally, details of the logarithmic regression used to fit the above median estimates to give an estimate of costs for all points on the continuous PASI scale (p81) are not provided and its implementation in the model is not described.

5.6 Results

5.6.1 Summary

The results of the model are presented in the manufacturer's submission from p86 to 94. In particular, it is worth noting that the submission includes: (i) the base-case results for a lifetime scenario showing an ICER of £25,991 per additional QALY for adalimumab compared to DMARDs, whilst etanercept is dominated and infliximab shows an ICER of £209,572 per QALY; (ii) results from a wide range of sensitivity analyses (Table 6.3.3.1, p.88-91).

5.6.2 Base-case analysis

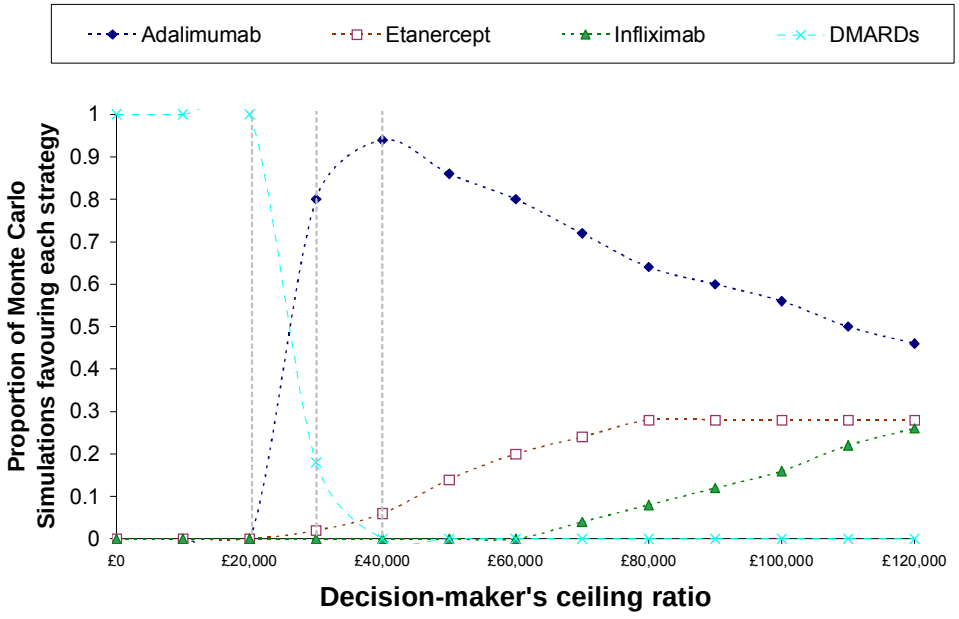
The mean costs and QALYs for the base-case lifetime scenario are presented in Tables 5.25 below.

Table 5.25 Results for base-case scenario from a lifetime perspective

	Mean Costs (£)	Mean QALYs	ICER (£)	Probability Cost Effective at (£): 20,000	30,000	40,000
Adalimumab	61,308	8.18	25,991	0%	80%	94%
Etanercept	65,627	8.15	Dominated	0%	2%	6%
Infliximab	81,614	8.27	209,572	0%	0%	0%
DMARD	28,518	6.91	-	100%	18%	0%

In the base-case analysis, the mean total costs of adalimumab were slightly less than etanercept (£61,308 compared to £65,627) but provided a modest QALY gain (0.03). Etanecept was therefore dominated by adalimumab. The ICER of adalimumab, compared to DMARDs, was £25,991 per QALY. Infliximab was estimated to provide a QALY gain of 0.09 compared to adalimumab, but at a significant additional cost (£81,614), providing an ICER of £209,572 per additional QALY. Figure 5.4 presents the cost-effectiveness acceptability curves for the base-case analysis considering a lifetime scenario. At a threshold willingness to pay of £30,000 per QALY, the probability adalimumab is cost-effective was 0.8.

Figure 5.4 Cost-effectiveness acceptability curves



The manufacturer's submission presented a number of alternative time horizons (1 year, 5 years, 10 years and lifetime). At each point in time adalimumab presents a lower ICER than infliximab, with etanercept consistently dominated by adalimumab. Both infliximab and adalimumab show almost the similar estimates of mean QALYs but infliximab has markedly higher costs (due to higher administration costs), resulting in a high ICER. The lifetime horizon is considered by the ERG to be the most appropriate to capture all future differences in costs and outcomes given the chronic nature of PsA.

5.6.3 Sensitivity analyses

Table 5.26 below presents a selection of the results of the sensitivity analysis based on those assumptions which appear to have greatest impact on the base-case results. These include:

- Variation of the baseline HAQ.
- Use of the Toronto dataset as an alternative source to estimate HAQ progression whilst responding to DMARD treatment.
- Use of SF-6D regression to estimate utility gains.
- Exclusion of PASI utility coefficients and PASI costs (separately and in combination) in order to analyse the impact of skin disease.
- Use of anti-TNF treatment specific withdrawal rates.

Table 5.26 Selected sensitivity analysis results

	Mean Costs(£)	Mean QALYs	ICER (£)	Probability Cost Effective at: £20,000	£30,000	£40,000
Baseline HAQ=0.5						
Adalimumab	60,598	9.74	26,067	0%	82%	94%
Etanercept	64,918	9.72	Dominated	0%	2%	6%
Infliximab	80,902	9.84	212,174	0%	0%	0%
DMARD	27,808	8.49	-	100%	16%	0%
Baseline HAQ=2						
Adalimumab	62,472	5.59	25,849	0%	82%	94%
Etanercept	66,794	5.56	Dominated	0%	2%	6%
Infliximab	82,777	5.69	204,890	0%	0%	0%
DMARD	29,687	4.32	-	100%	16%	0%
Toronto HAQ progression						

	Mean Costs(£)	Mean QALYs	ICER (£)	Probability Cost Effective at: £20,000	£30,000	£40,000
Adalimumab	60,536	9.82	47,404	0%	0%	20%
Etanercept	64,881	9.72	Dominated	0%	0%	0%
Infliximab	80,871	9.85	655,402	0%	0%	0%
DMARD	27,479	9.12	-	100%	100%	80%
SF-6D utility						
Adalimumab	61,307	9.72	62,360	0%	0%	0%
Etanercept	65,627	9.71	Dominated	0%	0%	0%
Infliximab	81,612	9.76	510,995	0%	0%	0%
DMARD	28,518	9.20	-	100%	100%	100%
No PASI utility coefficient						
Adalimumab	61,308	9.00	29,756	0%	48%	94%
Etanercept	65,627	8.95	Dominated	0%	0%	3%
Infliximab	81,614	9.08	22,510	0%	5%	3%
DMARD	28,518	7.89	-	100%	47%	0%
No PASI costs						
Adalimumab	54,537	8.18	27,099	0%	82%	94%
Etanercept	58,290	8.15	Dominated	0%	0%	6%
Infliximab	75,037	8.27	211,537	0%	0%	0%
DMARD	20,347	6.91	-	100%	18%	0%
No PASI costs or utility coefficient						
Adalimumab	56,220	9.00	31,606	0%	40%	98%
Etanercept	56,827	8.95	Dominated	0%	1%	2%
Infliximab	77,262	9.08	256,705	0%	6%	0%
DMARD	21,120	7.89	-	100%	53%	0%
Withdrawal rates for individual treatments						
Adalimumab	63,653	8.25	26,258	0%	80%	84%
Etanercept	72,430	8.35	86,237	0%	2%	16%
Infliximab	73,099	8.06	Dominated	0%	0%	0%
DMARD	28,518	6.91	-	100%	18%	0%

5.7 Summary of uncertainties and issues

As a general concern, the ERG noted a lack of transparency relating to the description of the methods in the manufacturer's submission report. The ERG identified a number of uncertainties and issues that may bias the model results, and consequently additional work was requested to address them and examine the potential robustness of the base-case results. The major concerns raised by the ERG include:

- The approach of the methods of evidence synthesis was overly complex, employing a number of assumptions which increase the possibility of significant bias in the subsequent results.
- The exclusion of relevant 12-week trial evidence from the analysis.
- The use of unnecessary assumptions in order to adjust 24-week trial results to estimate the response rates at 12-weeks (i.e. assuming that the relationship between response rates at 12 and 24-weeks observed for adalimumab applies to all other treatments, and forcing the 12-week response rate to be lower than 24-week response rates).
- The robustness of the cost-effectiveness results to different relevant subgroups (e.g. previous use of DMARDs, patients with and without skin involvement).

Chapter 6
Additional analyses requested by the ERG

6.1 Overview

The ERG requested additional analyses from the manufacturer to address several of the issues and uncertainties identified during the structured critique of their submission. Appendix 6 lists the complete list of issues raised by the ERG to the manufacturer as points for clarification. Following the manufacturer's response to these initial queries, a further analysis was requested to address a specific issue identified during the review of the manufacturer's initial response to the ERG. The additional analyses were requested to examine the potential robustness of the base-case results to several of the assumptions made in the manufacturer's model, and also to identify possible sources of bias. This section focuses primarily on the revised evidence synthesis undertaken by the manufacturer in response to the issues raised by the ERG in Chapter 5.

The ERG requested further clarification from the manufacturer on the following aspects:

- The assumptions and methods used for the indirect treatment comparison methods.
- Further analyses to test the robustness of the model results using additional trial response data, and especially its sensitivity to the PsARC response measure at 12-weeks. A re-run of the base-case analysis with the efficacy results of a Bayesian evidence synthesis of 12-week PsARC data for all comparators was requested.
- Further subgroup analyses to explore the robustness of results to alternative costing assumptions and patient characteristics, including:
    - Patients with and without skin involvement (defined according to the threshold of BSA > or ≤ 3%)
    - Patient use of previous DMARDs at baseline
    - Exclusion of the weighted average drug cost of the traditional DMARDs strategy.
    - A more conservative scenario for the HAQ rebound effect.

6.2 Critique of the re-submission

Critical review of revised evidence synthesis

In Chapter 5 the ERG identified a number of key areas of uncertainty relating to the approaches the manufacturer used to synthesise response data. In particular, the ERG was concerned with the decision to exclude trials reporting 12-week follow-up data from the manufacturer's analysis. The ERG felt that this was a key issue for a number of reasons:

- Potentially relevant evidence was excluded from consideration;
- the 12-week responses are central to the economic model since they determine initial response to treatment and the decision to continue treatment or not (based on PsARC response at 12-weeks in the base-case analysis and in accordance with current guidelines);
- the exclusion of 12-week trial data resulted in the manufacturer making an adjustment to 24-week trial results in order to estimate the response rates at 12-weeks.

The ERG felt that the exclusion of the 12-week trials represented a significant omission from the manufacturer's submission and considered that the approach taken by the manufacturer could lead to possible bias in the subsequent cost-effectiveness estimates. In order to fully assess the robustness of the cost-effectiveness results to this issue, the ERG requested additional data from the manufacturer. These included details of the response rates from the excluded 12-week trials and a full synthesis of PsARC response and other outcome parameters for etanercept, infliximab and adalimumab using data from all available studies (including both 12 and 24 week studies). In addition, the ERG requested that the manufacturer re-run their base-case analysis using the results from the revised evidence synthesis. An overview and critique of the manufacturer's response now follows.

Table 6.1 reports the results from the 6 anti-TNF trials, including the 3 studies excluded from the original synthesis, at 12 and 24-weeks. It is worth noting that the 12-week response rates for etanercept and infliximab appear higher than the 24-week results, consequently the adjustment made by the manufacturer to estimate 12-week response rates may not be appropriate (constraining the 12-week responses to be lower than 24-weeks). Furthermore, it is evident that the PsARC

response rates at 12-weeks for adalimumab based on M02-570 (excluded from the initial analysis) appear markedly lower than those from the ADEPT M02-518 study that was included. So, the adjustment made by the manufacturer to estimate 12-week response rates based on 24-week trial evidence may not only be inappropriate but it also appears to work in favour of adalimumab, at least regarding to PsARC response outcomes, potentially underestimating the cost-effectiveness of etanercept and infliximab in the model.

Table 6.1: Response Results at Week 12 for Adalimumab, Etanercept and Infliximab

Trial name	Treatments	N	% With BSA ≥ 3%	% MTX*	ACR20	ACR50	ACR70	PASI50	PASI75	PASI90	% HAQ change from baseline (SD) or mean change from baseline ± SD	PsARC	Length of blinded part of study
Mease (2000)	Placebo	30	63%	47%	13%	3%	0%	21%	0%	-	-0.1⁂	23%	12 Weeks
	Etanercept	30	63%	47%	73%	50%	13%	42%	26%	-	-1.2⁂	87%	
Mease (2004)	Placebo	104	60%	49%	15%	4%	0%	-	-	-	6%	31%	24 Weeks
	Etanercept	101	65%	45%	59%	38%	11%	-	-	-	54%	72%	
M02-570	Placebo	49	- *	47%	16%	2%	0%	-	-	-	-0.1 ± 0.3	[redacted]	12 Weeks
	Adalimumab	51	- *	47%	39%	25%	14%	-	-	-	-0.3 ± 0.5	[redacted]	
M02-518	Placebo	162	43%	50%	14%	4%	1%	15%	4%	0%	-0.1 ± 0.5	26%	24 Weeks
	Adalimumab	151	46%	51%	58%	36%	20%	72%	49%	30%	-0.4 ± 0.5	62%	
IMPACT I †	Placebo	52	42%**	65%	10%	0%	0%	0%	0%	0%	0.0	21%	16 Weeks
	Infliximab	52	33**	46%	65%	46%	29%	100%	68%	36%	-0.6	75%	
IMPACT II	Placebo	100	87%	45%	11%	3%	1%	9%	2%	0%	-18% (91)	27%	24 Weeks
	Infliximab	100	83%	47%	58%	36%	15%	82%	64%	41%	48% (43)	77%	

NA = Not Available

† The IMPACT I study was double-blinded for 16-weeks and therefore outcomes were evaluated at 16 weeks and not 12 weeks like the Mease 2000 study and M02-570. In addition, IMPACT II measures outcomes at Weeks 14 and 24, and not Weeks 12 and 24 like M02-518 and the Mease 2004 study and as such 14-week response rates are reported here.

* In study M02-570, patients with ≥3% psoriasis covering their body surface area (BSA) were not assessed, and as such no PASI scores can be calculated. This trial utilised the Target Lesion assessment as an indication of psoriasis severity, which evaluates target lesions for erythema, induration and scaling, each on a scale of 0 (best) to 5 (worst), with a total plaque score of 0 -15. Psoriasis-related assessments were conducted only for patients with a lesion that, at baseline, was ≥2 cm in diameter and had a plaque score ≥6.

** In the IMPACT I study, data on patients with ≥3% psoriasis covering their BSA were not presented, instead patients with a baseline PASI score of ≥2.5 were included in the efficacy evaluation of the skin.

⁂HAQ change from baseline for the Mease study is presented as a median change from baseline and not as a mean change from baseline (± SD) like M02-570 and IMPACT I.

Table 6.1 contd: Response Results at Week 24 for Adalimumab, Etanercept and Infliximab

Trial name	Treatments	N	% With BSA ≥ 3%	% MTX*	ACR20	ACR50	ACR70	PASI50	PASI75	PASI90	% HAQ change from baseline (SD) or mean change from baseline ± SD	PsARC	Length of blinded part of study
Mease (2000)	Placebo	30	63%	47%	-	-	-	-	-	-	-	-	12 Weeks
	Etanercept	30	63%	47%	-	-	-	-	-	-	-	-	
Mease (2004)	Placebo	104	60%	49%	13%	4%	1%	18%	3%	3%	6%	23%	24 Weeks
	Etanercept	101	65%	45%	50%	37%	9%	47%	23%	6%	54%	70%	
M02-570	Placebo	49	- *	47%	-	-	-	-	-	-	-	-	12 Weeks
	Adalimumab	51	- *	47%	65%	43%	27%	-	-	-	-0.3 ± 0.5	[redacted]	
M02-518	Placebo	162	43%	50%	15%	6%	1%	12%	1%	0%	-0.1 ± 0.4	23%	24 Weeks
	Adalimumab	151	46%	51%	57%	39%	23%	75%	59%	42%	-0.4 ± 0.5	60%	
IMPACT I †	Placebo	52	42%**	65%	-	-	-	-	-	-	-	-	16 Weeks
	Infliximab	52	33**	46%	-	-	-	-	-	-	-	-	
IMPACT II	Placebo	100	87%	45%	16%	4%	2%	8%	1%	0%	19% (103)	32%	24 Weeks
	Infliximab	100	83%	47%	54%	41%	27%	75%	60%	39%	46% (43)	70%	

As requested the manufacturer undertook a full Bayesian synthesis of the 6 anti-TNF trials, including both the 12-week and 24-week trials. Table 6.2 below reports the results from this synthesis.

Table 6.2 Response rates from Evidence Synthesis

Week 12

	Supportive Care	**Adalimumab**	**Etanercept**	**Infliximab**
N	**497**	**202**	**131**	**152**
%BSA>3	**57%**	**46%**	**65%**	**69%**
ACR20	13 % (10% - 17%)	53 % (24% - 80%)	61 % (31% - 87%)	64 % (34% - 89%)
ACR50	4 % (3% - 6%)	31 % (9% - 60%)	39 % (14% - 71%)	42 % (16% - 73%)
ACR70	1 % (0% - 2%)	14 % (2% - 35%)	19 % (4% - 47%)	21 % (5% - 49%)
PsARC	26 % (21% - 31%)	57 % (24% - 85%)	76 % (46% - 96%)	75 % (45% - 95%)
PASI50	12 % (3% - 25%)	65 % (11% - 92%)	39 % (3% - 81%)	82 % (42% - 97%)
PASI75	4 % (1% - 9%)	43 % (3% - 78%)	20 % (1% - 59%)	64 % (20% - 88%)
PASI90	1 % (0% - 3%)	23 % (1% - 56%)	9 % (0% - 35%)	42 % (7% - 72%)

Week 24

	Supportive Care	**Adalimumab**	**Etanercept**	**Infliximab**
N	**366**	**202**	**101**	**100**
%BSA>3	**60%**	**46%**	**65%**	**83%**
ACR20	14 % (8% - 24%)	58 % (25% - 83%)	48 % (10% - 82%)	56 % (12% - 86%)
ACR50	7 % (3% - 12%)	41 % (13% - 69%)	32 % (4% - 69%)	39 % (5% - 74%)
ACR70	2 % (1% - 4%)	23 % (5% - 48%)	17 % (1% - 47%)	22 % (2% - 53%)
PsARC	27 % (12% - 50%)	65 % (11% - 98%)	67 % (3% - 99%)	62 % (3% - 99%)
PASI50	14 % (2% - 40%)	75 % (8% - 98%)	43 % (2% - 88%)	84 % (20% - 99%)
PASI75	4 % (0% - 16%)	53 % (2% - 89%)	21 % (0% - 66%)	65 % (5% - 93%)
PASI90	1 % (0% - 5%)	32 % (0% - 73%)	9 % (0% - 42%)	45 % (1% - 81%)

NB: includes all phase III trials, plus the open label results from MO2-570

Comparing the results from the revised evidence synthesis with the original response rates applied in the model raises a number of important issues. While these issues focus on PsARC data, since this establishes the decision on whether to continue treatment, it should be noted that similar issues arise for other response parameters. The main issues identified were as follows:

- PsARC response rates at 12-weeks for both etanercept and infliximab appear higher than those at 24-weeks;
- PsARC response rates at 24-weeks for both etanercept (67%) and infliximab (62%) from the revised synthesis appear lower than the estimates previously assumed (70% in both), whereas the response rate for adalimumab at 24-weeks is higher than those previously assumed (65% compared to 62%).

The manufacturer commented on the drop in response rates for etanercept and infliximab in the discussion suggesting that this was due to (i) a slight deterioration in outcomes seen in the 6-month phase III studies and (ii) the fact that phase II studies were incorporated into the 12-week assessments, and not into the 24-week assessments. No explanation was provided for the increase in the response rate for adalimumab based on PsARC. The ERG was unable fully to validate the results of the Bayesian evidence synthesis performed (as part of the manufacturer's response to the initial points for clarification they submitted a copy of the WinBUGS code used in a previous assessment report for psoriasis[66] but did not supply the actual code that they used as the basis for these calculations). It is, therefore, not possible to determine whether (or how) the relationship between the two time periods (12-weeks and 24-weeks) have been considered, and hence it is unclear why incorporating additional 12-week trial data would have any influence on the 24-week results.

The ERG considers the most likely explanation for the disparity between the 24-week results from the revised and original analyses to be attributed to the inclusion of M02-570 in the 24-week response rate calculations for the new synthesis. The 24-week results from this study were based on data from the open-label follow-up phase with all patients (including the control group) receiving adalimumab after 12-weeks. In the absence of any control group data at 24-weeks for this particular study, the ERG has significant concerns about the validity and impact that including this trial has on the subsequent estimates of the response rate for both adalimumab (and also potentially for etanercept and infliximab since these may also be influenced by this data in the Bayesian synthesis).

It should also be recognised that the results, their correlation, and the associated uncertainty estimated from the revised synthesis are not directly used in the subsequent model:

- An adjustment is first made for the different levels of skin involvement to the mean response rates from the revised synthesis. The adjusted mean response rates are then used as part of the optimisation calculations to derive the associated Dirichlet distributions for the various groups of correlated response parameters. Consequently, the uncertainty surrounding the response rates are based on the Dirichlet simulations themselves rather than the estimates from the Bayesian synthesis. A comparison of the Bayesian credible intervals and those applied in the model suggests that the uncertainty surrounding these estimates has been significantly underestimated in the economic model.

- Secondly, it is important to be aware that the 12-week results themselves are not used directly. These results are simply used as the basis for estimating the adjustment subsequently applied by the manufacturer to the 24-week trial results to estimate 12-week response data.

The ERG has serious reservations about the way the 12-week trial data have been used to inform the cost-effectiveness estimates. Of particular concern is that these are not used directly in the model but are simply used as the basis for adjusting 24-week trial data. As we have previously noted (see Section 5.5.1), these adjustments constrain the 12-week response rates subsequently applied in the model to be equal or lower than those reported at 24-weeks. While in the original analyses all treatments were adjusted using the proportions observed in the ADEPT trial, in the revised synthesis different proportions are now assumed for each individual treatment based on the difference between 12 and 24-week results. The constraint imposed by the manufacturer is an important issue in the revised cost-effectiveness analysis since 12-week PsARC response rates from the Bayesian synthesis are higher than the 24-week response rates for both etanercept and infliximab. Table 6.3 below summarises the adjustments made by the manufacturer and clearly highlights the ERG's concerns about the impact this has on the 12-week response rates assumed for etanercept and infliximab. The table clearly shows that where the 12-week results were greater than the 24-week results, a ceiling of 100% was used. In other words, 12-week responses were assumed to be identical to those at 24-weeks when the ceiling estimate is applied. By imposing this constraint the manufacturer underestimates the actual response rates for infliximab and etanercept at 12-weeks.

Table 6.3: Adjustments applied in model for determining 12-week results from 24-week results[λ]

	Supportive Care	Adalimumab	Etanercept	Infliximab
ACR20	93%	91%	100%	100%
ACR50	57%	76%	100%	100%
ACR70	50%	61%	100%	95%
PsARC	96%	88%	100%	100%
PASI50	86%	87%	91%	98%
PASI75	100%	81%	95%	98%
PASI90	100%	72%	100%	93%
HAQ†	67%	76%	100%	98%
PASI‡	95%	80%	95%	96%

λ where 12 week results were greater than 24 week results, a ceiling of 100% was used.
† the average ACR20/50/75 between 12 and 24 weeks
± the average of PASI50/75/90 between 12 and 24 weeks

Table 6.4 shows the revised cost-effectiveness results presented by the manufacturer for a lifetime horizon. In contrast to the base-case analysis from their original submission, etanercept is no longer dominated by adalimumab. Furthermore, the ICER for adalimumab is now higher than the ICER for etanercept (£31,458 vs £19,856). Consequently, adalimumab is now actually subject to extended dominance[67] by etanercept in the revised cost-effectiveness analysis.

Table 6.4: Revised cost-effectiveness analysis results (based on new synthesis and incorporating the M02-570 study)

Intervention	Total cost	Total QALY	ICER	Probability Cost-effective at: £20,000	£30,000	£40,000
Adalimumab	69,677	8.14	31,458	0%	29%	35%
Etanercept	72,729	8.29	19,856	0%	39%	65%
Infliximab	87,675	8.41	114,234	0%	0%	0%
DMARD	28,518	6.83	-	100%	32%	0%

Based on the additional concerns noted in this section regarding the revised synthesis approach, the ERG believes these results should still be considered with some caution and are only indicative of the potential impact of considering all relevant evidence from both 12 and 24-week trial data. These ERG concerns are largely based on (i) the inclusion of the open-label results from M02-570 and (ii) the constraint applied to the PsARC response rates for etanercept and infliximab. Both of these issues were considered by the ERG to provide an overly optimistic estimate of the potential cost-effectiveness of analimumab relative to the other anti-TNF agents. It could be argued that, since adalimumab is already subject to extended dominance by etanercept with the current (optimistic) assumptions, that any further analysis addressing these issues would simply reinforce these conclusions. However, due to the complexity of the analysis and model, the ERG was keen to establish the logical consistency of the results to these issues.

The ERG therefore requested a further analysis from the manufacturer to examine the robustness of the response rate data and the cost-effectiveness results to the inclusion/exclusion of the 24-week (open-label) results from study M02-570. The ERG requested to re-analyse the response rate data and cost-effectiveness results by excluding the 24-week results from this study. The revised response rates, the new adjustments applied and the resulting cost-effectiveness estimates are summarised in Table 6.5 to 6.7.

Table 6.5: Response rates from Evidence Synthesis (excluding 24-week results from M02-570)

12-weeks

	Supportive Care	Adalimumab	Etanercept	Infliximab
N	**497**	**202**	**131**	**152**
%BSA>3	**57%**	**46%**	**65%**	**69%**
ACR20	13 % (10% - 17%)	53 % (24% - 80%)	61 % (31% - 87%)	64 % (34% - 89%)
ACR50	4 % (3% - 6%)	31 % (9% - 60%)	39 % (14% - 71%)	42 % (16% - 73%)
ACR70	1 % (0% - 2%)	14 % (2% - 35%)	19 % (4% - 47%)	21 % (5% - 49%)
PsARC	26 % (21% - 31%)	57 % (24% - 85%)	76 % (46% - 96%)	75 % (45% - 95%)
PASI50	12 % (3% - 25%)	65 % (11% - 92%)	39 % (3% - 81%)	82 % (42% - 97%)
PASI75	4 % (1% - 9%)	43 % (3% - 78%)	20 % (1% - 59%)	64 % (20% - 88%)
PASI90	1 % (0% - 3%)	23 % (1% - 56%)	9 % (0% - 35%)	42 % (7% - 72%)

24-weeks

	Supportive Care	Adalimumab	Etanercept	Infliximab
N	**366**	**151**	**101**	**100**
%BSA>3	**60%**	**46%**	**65%**	**83%**
ACR20	16 % (3% - 44%)	53 % (0% - 100%)	48 % (0% - 100%)	52 % (0% - 100%)
ACR50	8 % (1% - 28%)	41 % (0% - 99%)	36 % (0% - 99%)	40 % (0% - 99%)
ACR70	3 % (0% - 13%)	27 % (0% - 96%)	23 % (0% - 95%)	27 % (0% - 96%)
PsARC	27 % (7% - 57%)	57 % (0% - 100%)	60 % (0% - 100%)	58 % (0% - 100%)
PASI50	14 % (2% - 40%)	75 % (8% - 98%)	43 % (2% - 88%)	84 % (20% - 99%)
PASI75	4 % (0% - 16%)	53 % (2% - 89%)	21 % (0% - 66%)	65 % (5% - 93%)
PASI90	1 % (0% - 5%)	32 % (0% - 73%)	9 % (0% - 42%)	45 % (1% - 81%)

NB: includes all phase III trials, and excludes open label results from MO2-570. *WinBUGS convergence was difficult to achieve with only 3 trials entered, hence the large confidence intervals*

Table 6.6: Adjustments applied in model for determining 12 week results from 24 week results[λ]

	Supportive Care	Adalimumab	Etanercept	Infliximab
ACR20	81%	100%	100%	100%
ACR50	50%	76%	100%	100%
ACR70	33%	52%	83%	78%
PsARC	96%	100%	100%	100%
PASI50	86%	87%	91%	98%
PASI75	100%	81%	95%	98%
PASI90	100%	72%	100%	93%
HAQ†	55%	76%	94%	93%
PASI‡	95%	80%	95%	96%

λ where 12 week results were greater than 24 week results, a ceiling of 100% was used.
† the average ACR20/50/75 between 12 and 24 weeks
± the average of PASI50/75/90 between 12 and 24 weeks

Table 6.7: Revised cost-effectiveness analysis results (based on new synthesis and excluding the 24-week data from the M02-570 study)

	Total	Total	ICER	Probability Cost-effective at:		
				£20,000	£30,000	£40,000

	cost	QALY				
Adalimumab	67,457	8.35	25,893	0%	84%	90%
Etanercept	67,670	8.20	Dominated	0%	2%	10%
Infliximab	84,542	8.51	122,532	0%	0%	0%
DMARD	28,271	6.83	-	100%	14%	0%

Previously the ERG hypothesised that the exclusion of the 24-week data from M02-570 would result in a more conservative estimate of the cost-effectiveness of adalimumab. However, it is clear that this is not the case, with adalimiumab now dominating etanercept based on a lifetime time-horizon. The ICER of adalimumab relative to DMARD therapy is now estimated to be £25,893 per QALY (remarkably close to the original base-case ICER presented in the original submission of £25,991 per QALY). After reviewing the requested re-analysis the ERG felt that these results lacked face validity bringing into question the robustness of the evidence synthesis approach and/or assumptions used by the manufacturer. We now look at the different component parts to identify possible areas of remaining uncertainty.

The PsARC response from M02-570 at 24-weeks in this study (74% for adalimumab) was markedly higher than the corresponding 24-week estimate from ADEPT M02-518 study (60%). The ERG had anticipated that the exclusion of the 24-week results from study M02-570 would result in a reduction in the PsARC response rate from the revised synthesis for adalimumab. While this appears to have had the anticipated impact on the PsARC response rates at 24-weeks for adalimumab (reducing this from 65% to 57%), the PsARC data for both infliximab and etanercept also appear to have been significantly reduced in the process (from 67% to 60% for etanercept and from 62% to 58% for infliximab). The ERG is unclear why the exclusion of the 24-week results M02-570 would alter the PsARC response data for infliximab and etanercept in this manner. Indeed the ERG had envisaged that these might actually increase. The ERG noted the manufacturer's comment that convergence in the Bayesian synthesis was difficult to achieve with only 3 trials entered, hence the large credible intervals. Indeed the Bayesian credible intervals for PsARC response at 24-weeks across the 3 anti-TNF agents varies from 0-100% (see Table 6.5). The ERG therefore concludes that the revised synthesis requested does not appear to be sufficiently robust and that the subsequent cost-effectiveness estimates may not be valid.

Subgroup analysis – use of DMARDs at baseline

The ERG observed a potential trend in terms of different arthritic and psoriatic response rates in the ADEPT trial at both 12 and 24-weeks between patients having failed less than two or at least two previous DMARDs (e.g. at 12 weeks, patients on adalimumab who had

failed < 2 DMARDs showed a ACR70 response of 22% vs. 16% for those who had failed ≥ 2 DMARDs; 32% on adalimumab showed a PASI90 response vs. 26%, respectively. See Table 5.9.2.1 p.57 original submission for further details). In the case of the adalimumab arm of the ADEPT trial, 60% of the patients had failed < 2 and 40% ≥ 2 DMARDs. On this basis, the ERG requested a subgroup analysis of patients having failed < 2 or ≥ 2 previous DMARDs to test the robustness of the cost-effectiveness results to the use of sub-group specific response rates.

However, the manufacturer did not perform the subgroup analysis requested, commenting that "(…) *the number of previous DMARDs used proved not to be a significant predictor of HAQ or PASI in the model* (…)", presumably given that (…) "*the variation in number of baseline DMARDs was small."* (ERG clarification-NICE PsA Submission 120107.doc). The ERG did not feel that this was an entirely satisfactory justification to reject the requested analysis. While the ERG accepts that the use of previous DMARDs may not have been a significant predictor of HAQ or PASI (after controlling for the independent influence of ACR and PASI responses), the ERG did feel that the differences in the response parameters themselves (ACR and PASI) might be partially explained by the number of previous DMARDs. Instead, the manufacturer provided results from an additional subgroup analysis on disease duration, intended as a proxy of previous number of DMARDs failed. The ERG did not consider this additional analysis directly addressed their clarification query.

Subgroup analysis – skin involvement

Given the differences in terms of PASI and PsARC response observed between the anti-TNF agents (i.e. higher response rates for etanercept and infliximab compared to adalimumab in terms of PsARC response, and higher PASI responses for infliximab and adalimumab compared to etanercept; see Table 5.6.3, p47 of the original submission) and the focus of the model on quantifying the impact of both the arthritis *and* skin components, the ERG considered it appropriate to request a subgroup analysis of patients with and without skin disease (i.e. < 3% or ≥3% BSA). This was requested in order to explore the importance of the skin component of the disease in the overall cost-effectiveness results. The ERG also considered that the level of BSA could be used as a potential basis for making separate treatment decisions, conditional upon on the level of skin involvement. Table 6.8 below presents the sensitivity analysis results provided by the manufacturer for the base-case scenario.

Table 6.8: Impact of changing the percentage of psoriasis involvement of BSA >3% on cost-effectiveness estimates

				Probabilty cost-effective at:		
	Total costs	Total QALYs	ICER	£20,000	£30,000	£40,000
All skin						
Adalimumab	59,487	7.83	23,039	14%	98%	98%
Etanercept	66,881	7.88	Dominated (extended)	0%	0%	2%
Infliximab	82,625	8.04	107,177	0%	0%	0%
DMARD	32,381	6.65	-	86%	2%	0%
No skin						
Adalimumab	56,204	8.58	32,461	0%	16%	44%
Etanercept	60,516	8.69	39,646	0%	8%	52%
Infliximab	73,487	8.58	Dominated	0%	0%	0%
DMARD	20,317	7.48	-	100%	76%	4%

The results show important differences in terms of mean QALYs and ICER estimates for both scenarios. For the group of patients with skin disease, etanercept is now subject to extended dominance[67] by adalimumab (i.e. a less strong form of dominance compared to the original analysis). In addition, the ICER for adalimumab, compared to DMARDs, is £23,039 per QALY, appearing marginally more favourable than the original base-case analysis (£25,991 per QALY).

For patients without skin involvement, the mean total QALYs gained were higher for all treatments in comparison to the group of patients with skin involvement, and the cost-effectiveness of the different anti-TNF agents also varied. The ICER for adalimumab increased to £32,461 per QALY compared to DMARDs. Etanercept was no longer dominated, with an ICER of £39,646 compared to adalimumab. Infliximab was dominated by adalimumab.

The manufacturer claimed that caution should be exercised in analysing these results given the correlation between joint outcomes. The manufacturer argued that the sensitivity analysis presented in the original submission, which excluded the costs and utility for PASI outcomes, was a more useful approach to assessing the impact of removing the psoriasis component of the disease. However, the ERG was not seeking to exclude the impact of the psoriasis component but was instead looking at the most appropriate method for exploring the cost-effectiveness in different subgroups. The ERG felt that it was important to maintain the correlation between joint outcomes, whilst also adjusting for the different % of patients

with BSA >3% (in this case 0% and 100%). Simply excluding the costs and utility for PASI outcomes was not in itself sufficient, since this approach would still mean that the response rates themselves were still being adjusted on the basis that 66% of patients had skin involvement.

Exclusion of traditional DMARDs drug costs

The ERG requested a sensitivity analysis of the base-case excluding the drug costs of the "traditional DMARDs" option. This request was based on 2 issues raised by the ERG during their critical review. The first issue related to whether DMARDs were an appropriate comparator or not in the economic analysis (a strict interpretation of the licenses for anti-TNF therapies would suggest that these should be used "end of line" once DMARD therapies have been tried and failed). The second issue related to the the fact that the manufacturer's assumption of adding a weighted cost for DMARDs (including the cost of drug acquisition, monitoring and administration) without considering their actual treatment efficacy might have potentially penalised unduly this treatment strategy, underestimating its cost-effectiveness in the model and hence having a positive impact in the relative cost-effectiveness of the anti-TNF agents *(see Section 5.5.2).*

As part of the critical review, the ERG concluded that the manufacturer could have undertaken additional sensitivity analyses to explore the robustness of the base-case assumptions by assuming different proportions of patients (from 0 to 100%) would continue to receive DMARD therapy. Since the base-case analysis was based on the assumption that 100% would continue to receive DMARD therapy, the ERG wanted to seek clarification on the impact of assuming 0%.

The manufacturer did not perform the sensitivity analysis requested, arguing that *"(…) This analysis has not been conducted because it is considered that for the non-biologic comparison, the most appropriate comparator is a strategy containing conventional DMARDs"*. (ERG clarification-NICE PsA Submission 120107.doc). Instead of providing the ERG with the results of the requested sensitivity analysis the manufacturer predicted that *"The costs associated with conventional DMARDs will not have a major impact on the comparative cost-effectiveness of adalimumab versus etanercept or infliximab"* (ERG clarification-NICE PsA Submission 120107.doc). In the absence of any additional analysis exploring the robustness of the base-case results, the ERG cannot confirm or refute this statement.

Use of alternative HAQ rebound scenario

The ERG requested a sensitivity analysis using an alternative and more conservative assumption of the HAQ rebound after treatment failure with anti-TNF agents, in order to test the robustness of the reported base-case results to this assumption. The manufacturer did not provide the requested analysis, on the basis that *"(…)the immediate rebound to gain assumption on stopping anti-TNF therapy is not unduly optimistic and is a reasonable attempt at modelling the impact of treatment with anti-TNF agents in preventing joint erosion."* . The manufacturer's response also appeared to suggest that the current model may not be sufficiently flexible to explore this issue, commenting that *"We have been unable to calculate an accurate estimate of the scenario assuming rebound to natural history in the current model."*

However, the manufacturer explicitly recognized that under a more conservative rebound assumption, such as the rebound back to natural history scenario explored in the NICE guidance for etanercept and infliximab[23], *"(…) is likely that the benefits will diminish by a similar magnitude to previous estimates from the York model (15-16%)."* (ERG clarification-NICE PsA Submission 120107.doc). It is difficult to quantify the impact of this hypothetical reduction of health benefits on the relative ICERs for all strategies under analysis, however, just as a point of reference, the cost-effectiveness results for the base-case analysis of the York model[23] showed a difference in the ICER for etanercept between both rebound scenarios of about 40% (rebound back to natural history £16,801, rebound equal to gain £27,681, both for a lifetime horizon).

Chapter 7
Discussion and conclusions

7.1 Summary of clinical effectiveness issues

The manufacturer's submission was considered to comprise the most relevant clinical effectiveness evidence for the purpose of this STA. The clinical efficacy data used in the submission were limited, being largely derived from just two RCTs in 415 patients, with only 204 patients having received adalimumab, and two uncontrolled long-term open-label studies. More trial data would have been useful to assess the efficacy of adalimumab in PsA; unfortunately, no other trials relevant to the decision problem are available.

The limited data available indicate that adalimumab is efficacious in the treatment of PsA with beneficial effects on both joint and psoriasis symptoms and on functional status, in patients who had an inadequate response to previous treatment with NSAIDs or DMARDs. Furthermore, the improvements in both joint and psoriasis symptoms appeared to be maintained for up to 88 weeks in an open-label long-term extension study. In contrast to clinical trials of adalimumab in RA, the response rates observed with adalimumab used in combination with methotrexate were similar to those achieved with adalimumab alone. Adalimumab was generally well-tolerated in the clinical trials of PsA, with similar incidences of adverse events as with placebo. Overall the adverse event profile appears to be similar to that associated with use of the drug in RA. However, the long-term safety of adalimumab in PsA is undetermined; therefore review and further investigations of safety are warranted.

Although the pivotal trial results for the primary outcomes appear robust, the ERG felt that there are several areas of uncertainty regarding the clinical efficacy with respect to the decision problem considered in the submission. The participants in the pivotal RCTs were not entirely representative of the population for which adalimumab is currently licensed as neither population was made up exclusively of patients who had failed to respond to at least two DMARDs (40% & 51% for M02-518 and M02-570, respectively. Nevertheless, independent expert clinical advice given to the ERG suggests that the participants in these trials represented a population with relatively severe PsA similar to those currently being treated in UK clinical practice. In addition, a large proportion of the patients receiving adalimumab in these two studies were also receiving concomitant DMARD therapy (51% and 65%, respectively). Adalimumab is currently only licensed as a monotherapy not in combination with DMARDs for the treatment of PsA. There are currently no studies directly

comparing adalimumab with other anti-TNF agents or other established therapies in the treatment of PsA, therefore the relative efficacy is unclear. A further area of uncertainty exists regarding the potential of adalimumab to trigger the development of autoimmune antibodies.

[redacted] The formation of antibodies to adalimumab may have significant clinical implications as immunogenicity may be associated with a shortened duration of clinical response.

7.2 Summary of cost effectiveness issues

The manufacturer's submission was also considered to comprise the most relevant source of cost-effectiveness evidence related to the use of adalimumab for PsA. The ERG noted a number of strengths in the overall approach employed by the manufacturer in their cost-effectiveness analysis. In particular the attempt used to quantify the differential impact of the alternative treatments in terms of their impact on both the psoriasis and arthritis components of PsA were considered to address one of the limitations of existing cost-effectiveness studies in this area. However, the ERG identified a number of issues which appeared to compromise the validity of the model results, including:

- the failure to consider all relevant evidence (the exclusion of 12-week trial data) in estimating response rates for the economic model;
- the assumption of exchangeability of absolute response rates (after adjusting for differences between the proportion of patients with skin involvement) breaking the randomised comparisons in the individual trials;
- the adjustment used by the manufacturer to estimate 12-week response parameters from 24-week trial results;
- the exclusion of a potentially relevant comparator (palliative care)
- the assumptions concerning long-term HAQ progression in patients receiving conventional DMARDs and the assumption used to model the impact of treatment failure.

While the manufacturer attempted to address a number of these issues as part of additional work undertaken to address the clarification points raised by the ERG, the subsequent results were not considered sufficient to resolve the issues raised by the ERG. After reviewing the requested re-analysis the ERG felt that the results lacked face validity bringing into question the robustness of the evidence synthesis approach and/or assumptions used by the manufacturer and as such the subsequent cost-effectiveness estimates may not be valid.

7.3 Implications for research

In order to allow an accurate assessment of the clinical and cost effectiveness of adalimumab in PsA there is clearly a need for further research to clarify those areas of uncertainty outlined in this report. Efficacy trials conducted in the specific populations for which adalimumab is licensed are required (i.e. patients with active and progressive PsA that has responded inadequately to at least two DMARDs). In addition, these studies should be of adequate duration (>1 year) and directly compare adalimumab with other treatments for PsA. Additional information should also be collected on adalimumab in combination with other therapies.

Studies examining the genetic and immune factors involved in the cause and development of psoriatic arthritis would be helpful to establish which patients are likely to derive most benefit from anti-TNF agents and enable therapy to be targeted most appropriately. Furthermore, the adequate duration of therapy needs to be established and outcomes after stopping therapy investigated.

References

1. Abbott Laboratories. Adalimumab (Humira®) for the Treatment of Moderate to Severe Psoriatic Arthritis. Manufacturer's submission to the NICE STA process. November 2006.
2. Abbott Laboratories. Humira® (adalimumab) Summary of Product Characteristics. Electronic Medicines Compendium. http://emc.medicines.org.uk
3. Helliwell P, Marchesoni A, Peters M, et al. A re-evaluation of the osteoarticular manifestations of psoriasis. Br J Rheumatol 1991; 30(5):339-45.
4. Committee for medicinal products for human use (CHMP). Guideline on clinical investigation of medicinal products for the treatment of psoriatic arthritis. Adopted by CHMP 14 December 2006. Date for coming into effect 1 July 2007. http:/www.emea.europa.eu/pdfs/human/ewp/043804enfin.pdf (accessed 08/01/2007).
5. Kane D, Stafford L, Bresnihan B, FitzGerald O. A prospective, clinical and radiological study of early psoriatic arthritis: an early synovitis clinic experience. Rheumatology (Oxford) 2003; 42(12):1460-8.
6. Queiro-Silva R, Torre-Alonso JC, Tinture-Eguren T, Lopez-Lagunas I. A polyarticular onset predicts erosive and deforming disease in psoriatic arthritis. Ann Rheum Dis 2003; 62(1):68-70.
7. Anderson PJ. Tumor necrosis factor inhibitors: clinical implications of their different immunogenicity profiles. Semin Arthritis Rheum 2005; 34(5 Suppl1):19-22.
8. O'Neill T, Silman AJ. Psoriatic arthritis. Historical background and epidemiology. Baillieres Clin Rheumatol 1994; 8(2):245-61.
9. Brockbank J, Gladman D. Diagnosis and management of psoriatic arthritis. Drugs 2002; 62(17):2447-57.
10. Porter GG. Psoriatic arthritis. Plain radiology and other imaging techniques. Baillieres Clin Rheumatol 1994; 8(2):465-82.
11. Gladman DD. Natural history of psoriatic arthritis. Baillieres Clin Rheumatol 1994; 8(2):379-94.
12. McHugh N. Other seronegative spondyloarthropathies. Medicine 2006; 34(10):401-4.
13. Oriente P, Biondi-Oriente C, Scarpa R. Psoriatic arthritis. Clinical manifestations. Baillieres Clin Rheumatol 1994; 8(2):277-94.
14. Gladman DD, Antoni C, Mease P, et al. Psoriatic arthritis: epidemiology, clinical features, course, and outcome. Ann Rheum Dis 2005; 64 Suppl 2:ii14-7.
15. Gladman DD. Traditional and newer therapeutic options for psoriatic arthritis: an evidence-based review. Drugs 2005; 65(9):1223-38.
16. Gladman DD. Effectiveness of psoriatic arthritis therapies. Semin Arthritis Rheum 2003; 33(1):29-37.
17. Helliwell PS, Taylor WJ. Classification and diagnostic criteria for psoriatic arthritis. Ann Rheum Dis 2005; 64 Suppl 2:ii3-8.
18. Taylor W, Gladman D, Helliwell P, et al. Classification criteria for psoriatic arthritis: development of new criteria from a large international study. Arthritis Rheum 2006; 54(8):2665-73.
19. Gordon KB, Ruderman EM. The treatment of psoriasis and psoriatic arthritis: an interdisciplinary approach. J Am Acad Dermatol 2006; 54(3 Suppl 2):S85-91.
20. Galadari H, Fuchs B, Lebwohl M. Newly available treatments for psoriatic arthritis and their impact on skin psoriasis. Int J Dermatol 2003; 42(3):231-7.
21. British National Formulary No 52. March 2006, Analgesics. BMJ Publishing Group Ltd. London.

22. Pipitone N, Kingsley GH, Manzo A, et al. Current concepts and new developments in the treatment of psoriatic arthritis. Rheumatology (Oxford) 2003; 42(10):1138-48.
23. Woolacott N, Bravo Vergel Y, Hawkins N, et al. Etanercept and infliximab for the treatment of psoriatic arthritis: a systematic review and economic evaluation. Health Technol Assess 2006; 10(31):iii-iv, xiii-xvi, 1-239.
24. Mease P. Psoriatic arthritis update. Bull Hosp Jt Dis 2006; 64(1-2):25-31.
25. National Institute for Health and Clinical Excellence (NICE). Etanercept and infliximab for the treatment of adults with psoriatic arthritis; Technology Appraisal Guidance number 104. July 2006 http://www.nice.org.uk/TA104.
26. Kyle S, Chandler D, Griffiths CE, et al. Guideline for anti-TNF-alpha therapy in psoriatic arthritis. Rheumatology (Oxford) 2005; 44(3):390-7.
27. Telephone conference call with Dr Robert Chalmers, Consultant Dermatologist, Dermatology Centre, University of Manchester, Manchester and Dr Ian Bruce, Senior Lecturer and Consultant Rheumatologist, ARC Epidemiology Unit, University of Manchester, Manchester. dates 1912//06 and 30/01/07.
28. Wong K, Gladman DD, Husted J, et al. Mortality studies in psoriatic arthritis: results from a single outpatient clinic. I. Causes and risk of death. Arthritis Rheum 1997; 40(10):1868-72.
29. eMIMs. Humira. http://www.healthcarerepublic.com/mims/Drug/Search/Results/DrugDetails/b2bda6c0b53889e2a8da6469b57ed72e/ (accessed 12/01/2007).
30. Helliwell PS. GRAPPA--Group for Research and Assessment of Psoriasis and Psoriatic Arthritis, Stockholm, May/June 2006. J Rheumatol 2007; 34(1):214-9.
31. Mease PJ, Gladman DD, Ritchlin CT, et al. Adalimumab for the treatment of patients with moderately to severely active psoriatic arthritis: results of a double-blind, randomized, placebo-controlled trial. Arthritis Rheum 2005; 52(10):3279-89.
32. Genovese MC, Mease PJ, Thomson GTD. Adalimumab efficacy in patients with psoriatic arthritis who failed prior DMARD therapy [abstract no. FRI0187]. Ann Rheum Dis 2005; 64 Suppl. III: 313. EULAR Annual European Congress of Rheumatism; 2005 8-11 Jun; Vienna
33. Mease PJ, Gladman DD, Ritchlin CT. Clinical efficacy and safety of adalimumab for psoriatic arthritis: 48-week results of ADEPT [abstract no. 500]. Arthritis Rheum 2005; 52 (Suppl 9): 215. Annual Scientific Meeting of the American College of Rheumatology; 2005 Nov 12-17; San Diego
34. Mease PJ, Sharp JT, Ory PA. Adalimumab treatment effects on radiographic progression of joint disease in patients with psoriatic arthritis: results from ADEPT [abstract no. FRI0212]. Annual European Congress of Rheumatology, Jun 8-11 2005, Vienna. Ann Rheum Dis 2005 Jul; 64 Suppl. III: 320
35. van den Bosch F, Reece R, Manger B, et al. Adalimumab (HUMIRA®) is Effective and Safe in Treating Psoriatic Arthritis(PsA) in Real-Life Clinical Practice: Preliminary Results of the STEREO Trial. Annual Scientific Meeting of the American College of Rheumatology; 2006 Nov 10-15; Washington D.C. .
36. Antoni C, Krueger GG, de Vlam K, et al. Infliximab improves signs and symptoms of psoriatic arthritis: results of the IMPACT 2 trial. Ann Rheum Dis 2005; 64(8):1150-7.
37. Mease PJ, Kivitz AJ, Burch FX, et al. Etanercept treatment of psoriatic arthritis: safety, efficacy, and effect on disease progression. Arthritis Rheum 2004; 50(7):2264-72.
38. Mease PJ, Goffe BS, Metz J, et al. Etanercept in the treatment of psoriatic arthritis and psoriasis: a randomised trial. Lancet 2000; 356(9227):385-90.
39. British Society for Rheumatology (BSR). Statement on Adalimumab for Psoriatic Arthritis. http://www.rheumatology.org.uk/guidelines/clinicalguidelines/adalimpsajan06 January 2006.
40. Scottish Medical Consortium (SMC). Adalimumab 40mg pre-filled syringe for subcutaneous injection.

http://www.scottishmedicines.org.uk/updocs/Adalimumab%20(Humira)%20218-05.pdf (accessed 7/12/06).
41. Wales Medicines Partnership (WMP) assessment report: Adalimumab (Humira). June 2006 http://www.wales.nhs.uk/sites3/Documents/371/Enc%202%20WMP%20Adalimumab%20appraisal%2Epdf (accessed 7/12/06).
42. A prospective study for the effects of treatment with adalimumab in patients with psoriasis and psoriatic arthritis. ISRCTN23328456. http://www.controlled-trials.com/ISRCTN23328456/.
43. A Canadian Open-Label Study to Evaluate the Safety and Effectiveness of Adalimumab When Added to Inadequate Therapy for the Treatment of Psoriatic Arthritis (PsA) (ACCLAIM). http://clinicaltrials.gov/show/NCT00427362.
44. An Observational Study to Evaluate the Cost of the Treatments of the Refractory Psoriatic Arthritis to the Conventional Therapy. http://clinicaltrials.gov/show/NCT00303186.
45. Bansback NJ, Ara, R., Barkham, N., Brennan, A., Fraser, A.D., Conway, P. et al. Estimating the cost and health status consequences of treatment with TNF antagonists in patients with psoriatic arthritis. Rheumatology 2006; 45(8):1029-38.
46. Spanish Registry for Adverse Events of Biological Therapies in Rheumatic Diseases (BIOBADASER), www.biodabaser.ser.es/index_en.html. 2005.
47. Nixon RM, Bansback, N., Brennan, A. Using mixed treatment comparisons and meta-regression to perform indirect comparisons to estimate the efficacy of biologic treatments in rheumatoid arthritis. Statistics in Medicine 2006; 26(6):1237-1254.
48. Flendrie M, Creemers, M.C., Welsing, P.M., den Broeder A.A., van Riel, P.L. Survival during treatment with tumour necrosis factor blocking agents in rheumatoid arthritis. Ann Rheum Dis 2003; 62(Suppl 2:ii30-3).
49. Brazier J, Roberts, J., Deverill, M. . The estimation of a preference-based measure of health from the SF-36. Journal of Health Economics 2002; 21(2):271-292.
50. Gray AM, Rivero-Arias, O., Clarke, P.M. Estimating the association between SF-12 responses and EQ-5D utility values by response mapping. Med Decis Making 2006; 26(1):18-29.
51. Menter M, Gordon, K., Leonardi, C., Heffernan, L., Chen, D. Adalimumab efficacy and safety results in patients with moderate to severe chronic plaque psoriasis: subanalysis of patients with and without PsA. American College of Rheumatology (ACR). San Diego, California, USA, 2005.
52. Montly Index of Medical Specialties (January 2006).
53. Netten A, Curtis, L. . Unit costs of health and social care. University of Kent 2004.
54. Department of Health. Reference costs. 2003.
55. Barton P, Jobanputra, P., Wilson, J., Bryan, S., Burls, A. The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis. Heatlh Technology Assessment 2004; 8(11).
56. Kobelt G, Eberhardt, K., Jonsson, L., Jonsson, B. Economic consequences of the progression of rheumatoid arthritis in Sweden. Arthritis & Rheumatism 1999; 42(2):347-56.
57. Yelin E, Wanke, L.A. An assessment of the annual and long-term direct costs of rheumatoid arthritis: the impact of poor function and functional decline. Arthritis & Rheumatism 1999; 42(6):1209-18.
58. Drummond MF, et al. Methods for the Economic Evaluation of Health Care Programmes: Oxford Medical Publications, 2005.
59. Kaltwasser JP, Nash P, Gladman D, et al. Efficacy and safety of leflunomide in the treatment of psoriatic arthritis and psoriasis: a multinational, double-blind, randomized, placebo-controlled clinical trial. Arthritis Rheum 2004; 50(6):1939-50.

60. Salvarani C, Macchioni P, Olivieri I, et al. A comparison of cyclosporine, sulfasalazine, and symptomatic therapy in the treatment of psoriatic arthritis. J Rheumatol 2001; 28(10):2274-82.
61. Fraser AD, van Kuijk AW, Westhovens R, et al. A randomised, double blind, placebo controlled, multicentre trial of combination therapy with methotrexate plus ciclosporin in patients with active psoriatic arthritis. Ann Rheum Dis 2005; 64(6):859-64.
62. Antoni CE, Kavanaugh A, Kirkham B, et al. Sustained benefits of infliximab therapy for dermatologic and articular manifestations of psoriatic arthritis: results from the infliximab multinational psoriatic arthritis controlled trial (IMPACT). Arthritis Rheum 2005; 52(4):1227-36.
63. Wyeth Pharmaceuticals. An appraisal submission for the National Institute for Clinical Excellence: Etanercept (Enbrel*) to establish the clinical and cost-effectiveness of the use of etanercept for the treatment of active and progressive psoriatic arthritis in patients who have inadequate response to standard treatment. Maidenhead, 2004.
64. Brennan A, Bransback, N.J., Reynolds, A., Conway, P. Modelling the cost-effectiveness of etanercept in adults with rheumatoid arthritis in the UK. Rheumatology 2004; 43:62-72.
65. Bojke L, Claxton, L., Bravo Vergel, Y. Using expert elicitation to resolve issues of structural uncertainty in decision analytic models. SMDM, 2006.
66. Woolacott N, Hawkins, N., Mason, A., Kaith, A., Khadjesari, Z., Bravo Vergel, Y., Misso, K., Light, K., Chalmers, R., Sculpher, M., Riemsma, R. Efalizumab and infliximab for the treatment of psoriasis Health Technology Assessment 2007 (in press).
67. Johannesson M, Weinstein, M.C. . On the decision rules of cost-effectiveness analysis. Journal of Health Economics 1993; 12:459-467.

Appendices

Appendix 1. Search strategy undertaken by ERG for Adalimumab STA for the clinical effectiveness literature review.

Searches were undertaken on the following resources to identify relevant clinical effectiveness data: MedLine, Embase, Science Citation Index, Cochrane Library, EULAR, ACR, BSR, FDA, EMEA and NICE. All databases were searched from their inception to the date of the search. No language or other restrictions were applied to the study selection. The bibliographies of all included studies and the manufacturer's submission were reviewed to identify any further relevant studies. Due to the paucity of efficacy data available trials reported as abstracts were included in the review.

Inclusion criteria:

Participants: Adults with PsA

Interventions*:* Adalimumab administered by subcutaneous injection for the treatment of PsA.

Comparator: Placebo or any other active agent.

Outcomes: No restrictions applied (outcomes included; ACR, TSS, PsARC, PASI, HAQ, SF-36 and FACIT)

Design: Randomised controlled trials (RCTs) and open-label extension studies were included in the evaluation of efficacy.

Exclusion criteria:

Participants: Juveniles with PsA (< 16 years of age)

Intervention*:* Adalimumab for other indications (e.g. rheumatoid arthritis and akylosing spondylitis)

Study selection: Peer review panel

Study selection: The quality of included studies was assessed with regard to study design, adequacy of randomisation, allocation concealment, comparison of baseline characteristics between treatment arms, loss-to-follow up and use of intention-to-treat analysis.

Database Searches:

Database: MedLine (MEZZ)

Host: Dialog DataStar

Date search run: 14/12/06

Date span of search: MedLine 1950-to-date (including Old MedLine and MedLine in Process)

Search String	Description	Hits
1	CLINICAL-TRIAL-PHASE-III-PUBLICATION-TYPE.DE. OR CONTROLLED-CLINICAL-TRIAL-PUBLICATION-TYPE.DE. OR CONTROLLED-CLINICAL-TRIALS.DE. OR CLINICAL-TRIALS-PHASE-III.DE. OR RANDOMIZED-CONTROLLED-TRIALS.DE. OR RANDOMIZED-CONTROLLED-TRIALS.DE.	5118
2	PLACEBO.TI,AB,DE.	116744
3	CLINICAL NEAR TRIAL	577427
4	ADALIMUMAB OR HUMIRA OR D2E7 OR ‘D2’ ADJ 'E7'	495
5	ARTHRITIS-PSORIATIC.MJ.	1237
6	SPONDYLARTHRITIS#.W..DE. OR SPONDYLARTHROPATHIES#.W..DE.	13430
7	PSORIATIC NEXT ARTHRITIS	2436
8	1 OR 2 OR 3	631012
9	5 OR 6 OR 7	14485
10	8 AND 9 AND 4	**33**

Database: Embase (EMZZ)

Host: Dialog DataStar

Date search run: 14/12/06

Date span of search: 1974-to-date

Search String	Description	Hits
1	Clinical-Trial#.DE.	439267
2	Multicenter-Study.MJ. OR Phase-1-Clinical-Trial.MJ. OR Phase-2-Clinical-Trial.MJ. OR Phase-3-Clinical-Trial.MJ. OR Phase-4-Clinical-Trial.MJ. OR Randomized-Controlled-Trial.MJ.	3028
3	Clinical NEAR trial	495405
4	Placebo#.W..DE.	105239
5	Adalimumab OR humira OR DE27 OR ‘DE’ ADJ '27'	1879
6	Psoriatic-Arthritis.MJ.	2097
7	Psoriatic NEXT arthritis	3913
8	1 OR 2 OR 3 OR 4	550915
9	6 OR 7	3913
10	Spondyloarthropathy.W..MJ.	1252
11	9 OR 10	4937
12	8 AND 11 AND 5	187
15	(Adalimumab OR humira OR D2E7 OR ‘D2’ ADJ 'E7').TI,AB,DE.	1767
16	8 AND 11 AND 15	103
17	(Adalimumab OR humira OR D2E7 OR ‘D2’ ADJ 'E7').TI.	170
18	8 AND 11 AND 17	**59**

Database: ISI Proceedings (Index to Scientific & Technical Proceedings)

Host: ISI Web of Knowledge

Date search run: 14/12/06

Date span of search: 1990 - present

Search String	Description	Hits
1	Adalimumab	286
2	Psoriatic arthritis	381
3	1 AND 2	23

Database: The Cochrane Central Register of Controlled Trials (Clinical Trials)

Host: http://www3.interscience.wiley.com

Date search run: 14/12/06

Date span of search: unrestricted

Search String	Description	Hits
1	Adalimumab OR Humira OR DE27 OR DE ADJ '27'	21
2	Psoriatic arthritis	96
3	1 AND 2	**1**

Database: European League Against Rheumatism (EULAR)

Host: http://www.abstracts2view.com/eular/

Date search run: 14/12/06

Date span of search: 2002 – 2006

Search String	Description	Hits
1	Adalimumab	299
2	Psoriatic arthritis	938
3	1 AND 2	**54**

Database: American College of Rheumatology (ACR)

Host: http://www.rheumatology.org

Date search run: 14/12/06

Date span of search: 2002 – 2006

Search String	Description	Hits
1	Adalimumab (keyword)	
2	Psoriatic arthritis (text)	
3	1 AND 2	**7**

Database: British Society for Rheumatology (BSR)

Host: http://www.rheumatology.org.uk

Date search run: 14/12/06

Date span of search: 2002 – 2006

Search String	Description	Hits
1	Adalimumab (text)	
2	Psoriatic arthritis (text)	
3	1 AND 2	**5**

Indirect/mixed treatment literature search

A literature search was conducted to identify clinical trials completed since the previous HTA report.

Database: MedLine (MEZZ)

Host: Dialog DataStar

Date search run: 17/01/07

Date span of search: MedLine 01/01/2004-to-date (including Old MedLine and MedLine in Process)

Search string	Description	Hits
1.	PT=RANDOMIZED-CONTROLLED-TRIAL	
2.	RANDOM-ALLOCATION.DE.	
3.	DOUBLE-BLIND-METHOD.DE.	
4.	SINGLE-BLIND-METHOD.DE.	
5.	PT=CLINICAL-TRIAL$ OR PT=CONTROLLED-CLINICAL-TRIAL OR PT=MULTICENTER-STUDY OR PT=RANDOMIZED-CONTROLLED-TRIAL	
6.	PT=CLINICAL-TRIAL-PHASE-I OR PT=CLINICAL-TRIAL-PHASE-II OR PT=CLINICAL-TRIAL-PHASE-III OR PT=CLINICAL-TRIAL-PHASE-IV	
7.	PLACEBO$	
8.	PLACEBOS.W..DE.	
9.	RANDOM$	
10.	EVALUATION-STUDIES#.DE.	

11.	CLINICAL-TRIALS#.DE. OR CLINICAL-TRIALS#.DE. OR CLINICAL-TRIALS#.DE. OR CLINICAL-TRIALS#.DE.	
12.	FOLLOW-UP-STUDIES#.DE.	
13.	RESEARCH-DESIGN#.DE.	
14.	PROSPECTIVE-STUDIES#.DE.	
15.	1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14	
16.	ANIMAL=YES	
17.	HUMAN=YES	
18.	16 AND 17	
19.	17 NOT 18	
20.	15 AND 19	
21.	ARTHRITIS-PSORIATIC#.DE.	
22.	PSORIA$ ADJ ARTHRIT$	
23.	PSORIAS$ ADJ ARTHROPATH$	
24.	22 OR 23	
25.	21 OR 24	
26.	SULPHASALAZINE	
27.	SULFASALAZINE	
28.	SULFASALAZINE#.W..DE.	
29.	METHOTREXATE	
30.	METHOTREXATE#.W..DE.	
31.	MTX	
32.	CICLOSPORIN$	
33.	CYCLOSPORIN$	
34.	CYCLOSPORINE#.W..DE.	
35.	NEORAL	
36.	CSA	
37.	CYA	
38.	CYC-A	
39.	SANDIMMUM	
40.	CYCLOSPORINS#.W..DE.	
41.	AURANOFIN	
42.	AURANOFIN#.W..DE.	
43.	INTRAMUSCULAR$ ADJ GOLD	
44.	INTRA ADJ MUSCULAR$ ADJ GOLD	
45.	INTRA-MUSCULAR$ ADJ GOLD	
46.	IMI ADJ GOLD	
47.	INJECT$ ADJ GOLD	
48.	IM ADJ GOLD	
49.	IM ADJ GOLD	
50.	GOLD ADJ PREPARATION	
51.	GOLD ADJ SALT	

52.	PERORAL ADJ GOLD	
53.	PARENTERALLY ADJ GOLD	
54.	INTRAMUSCULAR$ ADJ ADMINISTRATION$ ADJ GOLD	
55.	INTRA ADJ MUSCULAR$ ADJ ADMINISTRATION$ ADJ GOLD	
56.	INTRA-MUSCULAR$ ADJ ADMINISTRATION$ ADJ GOLD	
57.	INJECTIONS-INTRAMUSCULAR#.DE.	
58.	GOLD#.W..DE.	
59.	57 AND 58	
60.	AZATHIOPRINE	
61.	AZATHIOPRINE#.W..DE.	
62.	AZA	
63.	PENICILLAMINE	
64.	PENICILLAMINE#.W..DE.	
65.	D-PENICILLAMINE	
66.	D ADJ PENICILLAMINE	
67.	ENKEPHALIN-D-PENICILLAMINE-2-5#.DE.	
68.	DPA	
69.	LEFLUNOMIDE	
70.	HYDROXYCHLOROQUINE	
71.	HYDROXYCHLOROQUINE#.W..DE.	
72.	HCQ	
73.	HXCHL	
74.	SALAZOPYRIN	
75.	SALICYLAZOSULPHAPYRIDINE OR SALICYLAZOSULFAPYRIDINE	
76.	SASP	
77.	PLACEBO$	
78.	PLACEBOS#.W..DE.	
79.	26 OR 27 OR 28 OR 29 OR 30 OR 31 OR 32 OR 33 OR 34 OR 35 OR 36 OR 37 OR 38 OR 39 OR 40	
80.	41 OR 42 OR 43 OR 44 OR 45 OR 46 OR 47 OR 48 OR 49 OR 50 OR 51 OR 52 OR 53 OR 54 OR 55 OR 56	
81.	59 OR 60 OR 61 OR 62 OR 63 OR 64 OR 65 OR 66 OR 67 OR 68 OR 69 OR 70 OR 71 OR 72 OR 73 OR 74 OR 75 OR 76 OR 77 OR 78	
82.	79 OR 80 OR 81	
83.	20 AND 25 AND 82	159

Database: Embase (EMZZ)

Host: Dialog DataStar

Date search run: 17/01/07

Date span of search: 01/01/2004-to-date

Search string	Description	Hits

1.	PT=RANDOMIZED-CONTROLLED-TRIAL	
2.	RANDOM-ALLOCATION.DE.	
3.	DOUBLE-BLIND-METHOD.DE.	
4.	SINGLE-BLIND-METHOD.DE.	
5.	PT=CLINICAL-TRIAL$ OR PT=CONTROLLED-CLINICAL-TRIAL OR PT=MULTICENTER-STUDY OR PT=RANDOMIZED-CONTROLLED-TRIAL	
6.	PT=CLINICAL-TRIAL-PHASE-I OR PT=CLINICAL-TRIAL-PHASE-II OR PT=CLINICAL-TRIAL-PHASE-III OR PT=CLINICAL-TRIAL-PHASE-IV	
7.	PLACEBO$	
8.	PLACEBOS.W..DE.	
9.	RANDOM$	
10.	EVALUATION-STUDIES#.DE.	
11.	CLINICAL-TRIALS#.DE. OR CLINICAL-TRIALS#.DE. OR CLINICAL-TRIALS#.DE. OR CLINICAL-TRIALS#.DE.	
12.	FOLLOW-UP-STUDIES#.DE.	
13.	RESEARCH-DESIGN#.DE.	
14.	PROSPECTIVE-STUDIES#.DE.	
15.	1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14	
16.	ANIMAL=YES	
17.	HUMAN=YES	
18.	16 AND 17	
19.	17 NOT 18	
20.	15 AND 19	
21.	ARTHRITIS-PSORIATIC#.DE.	
22.	PSORIA$ ADJ ARTHRIT$	
23.	PSORIAS$ ADJ ARTHROPATH$	
24.	22 OR 23	
25.	21 OR 24	
26.	SULPHASALAZINE	
27.	SULFASALAZINE	
28.	SULFASALAZINE#.W..DE.	
29.	METHOTREXATE	
30.	METHOTREXATE#.W..DE.	
31.	MTX	
32.	CICLOSPORIN$	
33.	CYCLOSPORIN$	
34.	CYCLOSPORINE#.W..DE.	
35.	NEORAL	
36.	CSA	
37.	CYA	
38.	CYC-A	
39.	SANDIMMUM	

40.	CYCLOSPORINS#.W..DE.	
41.	AURANOFIN	
42.	AURANOFIN#.W..DE.	
43.	INTRAMUSCULAR$ ADJ GOLD	
44.	INTRA ADJ MUSCULAR$ ADJ GOLD	
45.	INTRA-MUSCULAR$ ADJ GOLD	
46.	IMI ADJ GOLD	
47.	INJECT$ ADJ GOLD	
48.	IM ADJ GOLD	
49.	IM ADJ GOLD	
50.	GOLD ADJ PREPARATION	
51.	GOLD ADJ SALT	
52.	PERORAL ADJ GOLD	
53.	PARENTERALLY ADJ GOLD	
54.	INTRAMUSCULAR$ ADJ ADMINISTRATION$ ADJ GOLD	
55.	INTRA ADJ MUSCULAR$ ADJ ADMINISTRATION$ ADJ GOLD	
56.	INTRA-MUSCULAR$ ADJ ADMINISTRATION$ ADJ GOLD	
57.	INJECTIONS-INTRAMUSCULAR#.DE.	
58.	GOLD#.W..DE.	
59.	57 AND 58	
60.	AZATHIOPRINE	
61.	AZATHIOPRINE#.W..DE.	
62.	AZA	
63.	PENICILLAMINE	
64.	PENICILLAMINE#.W..DE.	
65.	D-PENICILLAMINE	
66.	D ADJ PENICILLAMINE	
67.	ENKEPHALIN-D-PENICILLAMINE-2-5#.DE.	
68.	DPA	
69.	LEFLUNOMIDE	
70.	HYDROXYCHLOROQUINE	
71.	HYDROXYCHLOROQUINE#.W..DE.	
72.	HCQ	
73.	HXCHL	
74.	SALAZOPYRIN	
75.	SALICYLAZOSULPHAPYRIDINE OR SALICYLAZOSULFAPYRIDINE	
76.	SASP	
77.	PLACEBO$	
78.	PLACEBOS#.W..DE.	
79.	26 OR 27 OR 28 OR 29 OR 30 OR 31 OR 32 OR 33 OR 34 OR 35 OR 36 OR 37 OR 38 OR 39 OR 40	
80.	41 OR 42 OR 43 OR 44 OR 45 OR 46 OR 47 OR 48 OR 49 OR 50 OR 51 OR 52 OR 53 OR 54 OR 55 OR 56	

81.	59 OR 60 OR 61 OR 62 OR 63 OR 64 OR 65 OR 66 OR 67 OR 68 OR 69 OR 70 OR 71 OR 72 OR 73 OR 74 OR 75 OR 76 OR 77 OR 78	
82.	79 OR 80 OR 81	
83.	20 AND 25 AND 82	81

The following databases were searched for current/ongoing research: Current Controlled Trials register (searched across multiple registers, including, ISRCTN, MRC NHS, and the National Institutes of Health registers), and Scirus, using the free text term psoriatic arthritis.

Appendix 2. Structured critical appraisal of submitted clinical efficacy evidence.

All studies included in the clinical evidence section of the Abbott submission were subjected to a detailed critical appraisal. Additional trial data presented within the submission are included where appropriate.

Appendix 2.1. M02-518 (ADEPT)[31]

CRITICAL APPRAISAL

Name of Trial: Adalimumab for the treatment of patients with moderately to severely active psoriatic arthritis. Results of a double-blind, randomised, placebo-controlled trial (ADEPT)[31]

Reference: Mease P J, Gladman D D, Ritchlin C T et al. Arthritis Rheum 2005; 52 (10): 3279-89

Question: How does adalimumab compare with regards to efficacy and safety with placebo in the treatment of active psoriatic arthritis?

Summary: This trial showed that in the short term (24 weeks), adalimumab was statistically significantly superior to placebo in active psoriatic arthritis patients with regards to ACR20 improvement and the change in score of structural damage on radiographs of hands and feet at 24 weeks. This trial does not tell us how adalimumab performs in the longer term, or how it compares to other anti-TNF drugs in the same class, or even how it compares to DMARDs. The license for adalimumab states that it "is indicated for the treatment of active and progressive psoriatic arthritis in adults when the response to previous disease-modifying anti-rheumatic drug therapy has been inadequate". This population was not the population studied in the trial so it is unclear how adalimumab performs in patients who have tried and failed, on DMARDs.

Did the study ask a clearly focussed question?

Yes - This study was designed to assess the safety and efficacy of adalimumab when compared to placebo in patients with active psoriatic arthritis (PsA).

Was the study design appropriate?

Yes – This was a 24-week, double-blind, randomised, parallel, placebo-controlled trial. Patients were stratified according to their methotrexate use, degree of psoriasis involvement and site and then randomised in a 1:1 ratio to either adalimumab or placebo. All patients who completed the 24-week protocol were eligible for long term treatment in an open-label extension study.

Patients were included in the study if they were at least 18 years old, had moderately to severely active PsA and had either active psoriatic skin lesions or a documented history of psoriasis. Patients also had to have a history of an inadequate response or intolerance to non-steroidal anti-inflammatory drug therapy. Patients were excluded if they had had treatment within 4 weeks of the baseline visit with ciclosporin, tacrolimus, disease-modifying anti-rheumatic drugs (DMARDs) other than methotrexate or oral retinoids. Other exclusions included certain topical psoriasis treatments, concurrent methotrexate and/or corticosteroid treatment at certain dosages and anti-TNF treatment at any time.

Patients were randomised to receive subcutaneous (s/c) injections of either placebo or 40mg adalimumab every other week. After week 12, patients who failed to have at least a 20% decrease in both swollen and tender joint counts on two consecutive visits could receive rescue therapy with either corticosteroids or DMARDs.

The primary efficacy endpoints were the American College of Rheumatology 20% improvement response (ACR20) at week 12 and the change in modified total Sharp score of

structural damage on radiographs of the hands and feet at week 24. Secondary efficacy endpoints included the ACR20 response rate at week 24 and the ACR 50% and 70% response rates (ACR50 and ACR70) at weeks 12 and 24. For patients with psoriasis affecting at least 3% of the body surface area, the Psoriasis Area and Severity Index (PASI) assessed the response of psoriasis in each group; rated as 50% improvement (PASI50) and 75% improvement (PASI75). Other endpoints included response rates on the modified psoriatic arthritis response criteria, the disability index of the health assessment questionnaire and the short form 36 health survey at weeks 12 and 24.
Safety was evaluated in terms of adverse events reported by the patients. Pair wise comparisons of the percentages of patients experiencing adverse events between the adalimumab and placebo treatment groups were performed using Fisher's exact tests. All statistical tests were 2-sided, at s α level of 0.05.

Were participants appropriately allocated to intervention and control groups?
Can't tell – No details are given about the method of randomisation employed and it is not clear how participants were allocated to the intervention and control groups. The trial simply states that patients were stratified according to their methotrexate use, degree of psoriasis involvement and site and were randomised in a 1:1 ratio by site to receive either adalimumab or placebo. The baseline characteristics of both patient groups were comparable and there were no significant differences between the treatment groups.

Were participants, staff and study personnel 'blind' to participants study group?
Yes – Blinding was not specifically mentioned however, the placebo group patients received s/c placebo injections as well as the adalimumab group receiving their s/c injections, so there was an attempt to blind the patients. When one of the primary endpoints, the sharp score, was assessed by the use of radiographs, two readers, who were blinded to the treatment and film order, reviewed the radiographs.

Were all of the participants who entered into the trial accounted for at its conclusion
Yes – All patients who received at least one dose of study treatment were included in the intention to treat analysis. Of the 315 patients randomised to receive treatment, 162 were assigned to the placebo group and 153 to the adalimumab group. Two patients assigned to the adalimumab group never received study medication and were excluded from all analyses. Of the placebo group; 149 completed the study (the 13 patients who discontinued were all accounted for apart from one lost to follow up). Of the 153 adalimumab patients; 140 completed the study (the 11 patients who discontinued were all accounted for).
For the second primary endpoint, 17 patients (10 placebo and 7 adalimumab) did not have radiographs obtained at both time points and several methods of imputation were used to test the effects of the missing data. The difference in change in Sharp score was statistically significant with each imputation method used. The imputation method was not explained in detail in the paper, it stated simply that several sensitivity analyses were used to evaluate any effect of the missing data, giving some brief examples.

Were the participants in all groups followed up and data collected in the same way?
Yes – Study visits were conducted at baseline, weeks two and four, and then every four weeks until week 24. It was not stated in the trial whether the treatment was the same at each of the visits.

Was the study large enough?
Yes - The sample size was based on anticipated changes in the modified total Sharp scores. A sample size of 150 per treatment group resulted in 80% power to detect an effect size of 0.325. Baseline modified total sharp scores were stated as 19.1 +/- 35.5 (placebo, n=161) and 22.7 +/- 46.0 (adalimumab, n=150), but no further values were provided. All patients who received at least 1 dose of study treatment were included in the data (intention to treat) analysis (162 placebo patients and 151 adalimumab patients).

How are the results presented and what is the main result?

Patients treated with adalimumab demonstrated significantly higher ACR response rates than those treated with placebo at all time points. The ACR20 response rate at week 12 (the primary endpoint) was 58% for the adalimumab group and 14% for the placebo treated group (between-group difference 44%, 95% confidence interval (CI) 33-54%, p<0.001). At week 24, the ACR20 response rates were 57% and 15% in the adalimumab and placebo groups respectively (between-group difference 42%, 95% CI 31-52%, p<0.001). At weeks 12 and 24, the ACR50 response rates were significantly higher in the adalimumab group compared to the placebo group and a significant difference in ACR70 was also seen. ACR20, 50 and 70 response rates did not differ between patients taking adalimumab combined with methotrexate (response rates of 55%, 36% and 17% respectively) and those taking adalimumab alone (61%, 36% and 23% respectively). The manufacturer provided additional information in response to the points for clarification raised by the evidence review group regarding the response rates at week 12 and 24 based in the subgroups of patients with psoriasis <3% or >=3% at baseline (Table 2.). ACR and PsARC responses were not significantly different between the two subgroups.

Table 1. Response parameters at 12 and 24 weeks for study M02-518 based on the subgroups of patients with psoriasis <3% or >=3% at baseline.

12-week response rates from M02-518 (ADEPT) trial

Response	Placebo		Adalimumab	
	Psoriasis <3% BSA	Psoriasis >=3% BSA	Psoriasis <3% BSA	Psoriasis >=3% BSA
ACR20	15%	13%	56%	60%
ACR50	4%	3%	37%	36%
ACR70	1%	0%	21%	19%
PsARC	26%	26%	62%	61%

24-week response rates from M02-518 (ADEPT) trial

Response	Placebo		Adalimumab	
	Psoriasis <3% BSA	Psoriasis >=3% BSA	Psoriasis <3% BSA	Psoriasis >=3% BSA
ACR20	16%	14%	60%	53%
ACR50	5%	6%	42%	36%
ACR70	2%	0%	22%	23%
PsARC	26%	20%	64%	56%

The second primary endpoint considered the change in modified total Sharp score of structural damage on radiographs of the hands and feet at week 24. The mean change in the Sharp score in patients who had both baseline and week 24 radiographs was -0.2 for adalimumab patients compared with 1.0 for placebo patients (p<0.001).

The PASI75 response rate at 24 weeks was 59% in the adalimumab group and 1% in the placebo group (n=69 per group, p<0.001). Disability was measured by the health assessment questionnaire (HAQ DI) and was -0.4 +/- 0.5 in the adalimumab group vs. -0.1 +/- 0.5 in the placebo group at week 12 (p<0.001). Further analysis in the trial looked at the response of psoriatic skin disease to treatment and quality of life responses.

How safe were the regimens?

Most adverse events were similar in the placebo and adalimumab groups with the most common being upper respiratory tract infections (14.8% vs. 12.6%), nasopharyngitis (9.3% vs. 9.9%), injection site reactions (3.1% vs. 6.6%) and headache (8.6% vs. 6.0%)

respectively. Seven placebo and five adalimumab patients experienced serious adverse events, four patients (three in the adalimumab group) prematurely discontinued treatment due to adverse events. Additionally, two adalimumab patients discontinued treatment due to abnormal laboratory results.

How precise are the results?

The primary endpoint of the ACR20 at week 12 was found to be statistically significant for adalimumab patients compared to those on placebo (58% vs. 14%). Confidence intervals were quoted as being the 'between-group differences' and were between 33 and 54% and the p value was <0.001. For the second primary endpoint, the change in modified total Sharp score was shown to be significantly better for adalimumab patients than placebo (-0.2 and 1.0 respectively, p<0.001). No confidence intervals were quoted here.

Can the results be applied to the local population?

In terms of the population studied, the geographical locations of the trial were in keeping with the general population of the United Kingdom (UK). Current practice in the UK is to use a disease modifying anti-rheumatic drug (DMARD) after non-steroidal anti-inflammatory drugs (NSAIDs), and this trial studied patients who had a history of intolerance or ineffectiveness to an NSAID but DMARDs were not mentioned in this context. How applicable this would be to a local population is unclear – would it be expected that most patients would be trialled on two DMARDs before trying an anti-TNF drug? If this is the case, this trial could not be extrapolated to this population. The manufacturer provided additional information regarding previous DMARD use in response to the points for clarification raised by the evidence review group. This showed that only 40% of adalimumab patients had previously received treatment with at least two or more DMARDs (Table 2.). Half of the patients in each group were taking methotrexate at baseline, and therefore were taking it in combination with adalimumab; it is unclear how common this is in current UK practice

Table 2: Number of different types of previous DMARDs – M02-518

Number of different types of previous DMARDs	Placebo (n=162) N (%)	Adalimumab (n=151) N (%)
0	[redacted]	[redacted]
1	[redacted]	[redacted]
2	[redacted]	[redacted]
3	[redacted]	[redacted]
4	[redacted]	[redacted]
5	[redacted]	[redacted]
6	[redacted]	[redacted]
Any Previous DMARD	[redacted]	[redacted]

('Academic in confidence')

How does adalimumab compare with regards to efficacy and safety with placebo in the treatment of active psoriatic arthritis?

This trial showed that in the short term (24 weeks), adalimumab was superior to placebo in psoriatic arthritis patients with regards to specific measures. It does not tell us how this drug performs in the longer term, or how it compares to other anti-TNF drugs in the same class, or even how it compares to DMARDs. The ACR20, 50 and 70 response rates did not differ between patients taking adalimumab combined with methotrexate compared to those taking adalimumab alone. The license for adalimumab states that it "is indicated for the treatment of active and progressive psoriatic arthritis in adults when the response to previous disease-modifying anti‑rheumatic drug therapy has been inadequate". This population was not the

population studied in the trial so it is unclear how adalimumab performs in patients who have tried and failed, on DMARDs.

Appendix 2.2. M02-570[32]

CRITICAL APPRAISAL

Name of Trial: Adalimumab efficacy in patients with psoriatic arthritis who failed prior DMARD therapy.[32]

References: Genovese MC, Mease PJ, Thomson GTD et al. Proceedings from the Annual European Congress of Rheumatology (EULAR) 2005, Vienna, Austria. FR10187.

Abbott Laboratories Ltd. Adalimumab (Humira®) for the Treatment of Moderate to Severe Psoriatic Arthritis. Manufacturer's submission to the NICE STA process. November 2006

Question: Is adalimumab more effective than placebo in patients with moderately to severely active psoriatic arthritis who had an inadequate response to DMARD therapy?

Summary: In a small (n=100) 12-week trial, adalimumab was more effective in terms of the ACR20 response than placebo in patients with moderately to severely active PsA in whom DMARDs had failed. Several secondary endpoints, including effects on disability, the skin component of psoriasis and ACR50 and ACR70 responses, also suggested that adalimumab was more effective than placebo. However, the quality of this study, and the robustness of the data cannot be assessed due to the lack of full details provided, as this study has not been published in full. This study provides only short-term safety data. In the initial 12-week RCT the incidence of adverse effects occurring in ≥ 5% of patients were similar in the placebo and adalimumab groups, with only aggravation of psoriasis and psoriatic arthritis occurring more frequently with placebo than adalimumab. Since only 56% of adalimumab patients had previously received treatment with at least two or more DMARDs the patient population in this study may not be fully representative of the UK population for whom, according to current guidelines, anti-TNF therapy, such as adalimumab, would be considered.

Did the study ask a clearly focussed question?

Yes – This study was designed to evaluate the efficacy and safety of adalimumab compared with placebo in patients with moderately to severely active PsA who had an inadequate response to DMARD therapy. The population studied, interventions given and outcomes considered are clearly stated. The primary endpoint was the ACR 20 response at week 12. Secondary endpoints were: mean improvement in the Disability Index of the Health Assessment Questionnaire (HAQ) and, in subjects with a psoriasis target lesion, the mean percent reduction in the Target Lesion (TL) score and the Physician's Global Assessment (PGA) for psoriasis (% of patients considered 'clear' or 'almost clear'). ACR50 and ACR70 responses are also presented. Adverse effects (AEs) occurring in ≥ 5% of patients are reported.

Was the study design appropriate?

Yes – This study was a phase III, 12-week randomised, placebo-controlled, double-blind trial. Eligible patients had moderately to severely active PsA (defined as ≥ 3 swollen joints and ≥ 3 tender joints) and had an inadequate response (not defined) to DMARD therapy based on current or historic DMARD treatment. The submission states that patients were permitted to continue therapy with methotrexate, prednisolone or oral corticosteroids. Other inclusion criteria were: ≥ 18 years of age, and presence of active cutaneous plaque psoriasis lesions or documented history of chronic plaque psoriasis. Exclusion criteria included prior anti-TNF therapy. Upon completion of 12 weeks therapy, patients were eligible to enter an open-label extension study (M02-537).

Were participants appropriately allocated to intervention and control groups?
Probably – Patients were randomised to receive adalimumab 40 mg (n = 51) or matching placebo (PbO; n = 49) subcutaneously (SC) every other week (eow) for 12 weeks. No details of the randomisation method used or who performed the randomisation are given in the published abstract. However, the submission states that patients were randomised in blocks of four using an Interactive Voice Response System in a 1:1 ratio. Patients were stratified according to DMARD use (yes/no), but neither details of the method used nor by whom stratification was carried out are stated. The mean numbers of previous DMARDs taken by both groups are presented and are reported to not differ significantly. The type and percentage use of DMARDs at baseline is presented in the submission. The submission also states that subjects were required to maintain baseline DMARD usage and dosage. The baseline demographics of the two groups appear similar, except there were significantly fewer rheumatoid factor-negative patients in the adalimumab arm than the placebo arm (80.4% vs. 98.0%; p = 0.01 (the submission shows that mean C-reactive protein also differed significantly between the two groups (1.0mg/L vs. 1.6mg/L; p=0.05, respectively)). According to two clinical experts, around 15% of patients would be expected to be rheumatoid factor-positive, whereas in the trial about 19% and 2% of the adalimumab and placebo groups, respectively, were. There is a suggestion that rheumatoid factor and C-reactive protein are severity markers in PsA patients, and therefore a higher percentage of the adalimumab group would be classed as having severe disease. However, the experts considered that this would not bias the results significantly; if anything, the effects of adalimumab would be underestimated.[27]

Were participants, staff and study personnel 'blind' to participants study group?
Yes – The initial 12-week RCT was double-blinded, but no details of how blinding was achieved are given in the abstract. However, the submission states that all investigators, study site staff and patients were blinded to treatment administered and that measurement techniques were not subject to observer bias.

Were all of the participants who entered into the trial accounted for at its conclusion?
Yes – A table of the disposition of the patients is presented. Of the 100 subjects entering the study 96 completed the 12 week study (50 (98%) and 46 (94%) for adalimumab and placebo, respectively).

Were the participants in all groups followed up and data collected in the same way?
Probably – Few details are given in the abstract, but enrolment screening included chest X-ray, electrocardiogram, PPD skin test (for tuberculosis), and routine laboratory tests. Study visits were conducted at weeks 2, 4, 8, 12, 14, 18 and 24.

Was the study large enough?
Probably – No power calculation or details of the statistical analyses performed in the RCT are presented in the abstract. However, the submission states that a sample size of 50 patients per treatment group provides >90% power to detect an effects size of 0.05, assuming an ACR20 response of 60% and 25% in the adalimumab and placebo groups, respectively.

How are the results presented and what is the main result?
The results for the two groups at 12 weeks are reported, and p values for the differences between the groups are presented.

- **ACR20 (primary endpoint):** In the adalimumab group, 39% of the 51 patients achieved an ACR20 response compared with 16% of the 49 in the placebo group (p = 0.05). The submission adds that the ACR20 response rate was greater for adalimumab than placebo by week 2, with the difference becoming significant by week 4 (p=0.001).

- **ACR50 and ACR70:** More patients on adalimumab than placebo achieved an ACR50 response (25% vs. 2%, respectively; p = 0.001) and an ACR70 response (14% vs. 0%, respectively; p = 0.05).
- **Health assessment questionnaire:** The mean change in the HAQ score was -0.3 with adalimumab (a negative value denotes improvement) compared with -0.1 with placebo (p = 0.01).
- **Physician's global assessment:** A PGA of clear or almost clear was reported in 40.6% and 6.7% of the adalimumab and placebo groups, respectively (p = 0.01).
- **Target lesion score:** The mean percentage change (a negative value denotes a reduction and hence clinical improvement) in the TL score was greater with adalimumab (-47.0%) than placebo (-1.6%; p = 0.001). The submission presents TL scores as mean change from baseline (-0.3 vs. -3.7; p = 0.001, respectively).
- **Modified PsARC:** Measurement of PsARC at week 12 was included in the submission as an outcome measure; however no PsARC results are presented in the published abstract. This supplemental data presented by the manufacturer as 'academic in confidence' states that [redacted]. It should be noted that since this supplemental data is not in the public domain it cannot be validated externally. Furthermore, that this data is apparently derived from a very small number of patients (n=32 and n=30 for the adalimumab and placebo groups, respectively).
- **SF-36 Domains of physical functioning:** Measurement of SF-36 domain and component summary scores at baseline and week 12 were also included in the submission, but not in the published abstract**.** Adalimumab-treated subjects were stated to have exhibited improvements in the SF-36 domains of physical functioning (p=0.027), bodily pain (p=0.007), mental health (p=0.009), and general health (p=0.017) when compared to placebo-treated subjects at Week 12.
- **Dermatology Life Quality Index (DLQI):** Measurement of DLQI at baseline and week 12 were also included in the submission, but not in the published abstract. Adalimumab-treated subjects were stated to have exhibited greater improvement in the DLQI from baseline to week 12 compared with placebo (mean change -3.4 vs. -1.7, respectively).

How safe were the regimens?

The numbers and percentages of patients in each group with any adverse event (AE) occurring in ≥ 5% of patients during the 12-week RCT are reported with p values for significant differences between the groups. Overall, significantly more adverse effects were reported in the placebo group than the adalimumab group (39 (79.6%) vs. 27 (52.9%); p ≤ 0.01, respectively). The only adverse effects reported to occur significantly more frequently in the placebo group than the adalimumab group were psoriasis aggravated (16.3% vs. 3.9%, respectively, p ≤ 0.05)) and psoriatic arthritis aggravated (14.3% vs. 2.0%, respectively, p ≤ 0.05). When adverse effects relating to disease aggravation were excluded there was still a statistically significant difference between the two groups (p ≤ 0.05).

How precise are the results?

Can't tell – The numbers in each group are small (51 and 49 in the adalimumab and placebo groups, respectively), no power calculations for any of the comparisons are presented in the published abstract, nor are details of the statistical methods used to determine differences between the groups, although p values are reported. Although the abstract data is supplemented with additional information provided by manufacturer in the submission, this supplemental information is not in the public domain and therefore cannot

be validated externally. Since this study has not been fully published and the results are reported only in abstract form there is insufficient data presented to fully assess the quality and validity of this study. Until this study is fully published and the complete data made available for evaluation these results should be interpreted with due caution.

Can the results be applied to the local population?

If the clinical characteristics of the local population with PsA who would be considered appropriate for anti-TNF therapy are the same as those of the trial population in terms of age, disease severity, inadequacy of response to DMARDs and presence or history of chronic plaque psoriasis, then the results of this trial would be expected to be applicable. However, insufficient details are provided to be sure that the trial population is representative of the UK population with PsA for whom anti-TNF therapy, such as adalimumab, would be considered. The British Society for Rheumatology (BSR) has produced guidelines for the use of anti-TNF therapy in PsA.[26] Adalimumab is not currently included in this guideline; however, the BSR has issued a statement supporting its use in accordance with the aforementioned guidelines.[39] An adequate trial (clearly defined in the guidelines) of two standard DMARDs alone or in combination is recommended, and if patients fail such therapy and still have active joint disease (≥ 3 tender joints and ≥ 3 swollen joints on two separate occasions one month apart), then the next step is to use a licensed anti-TNF agent.[26] An inadequate response to DMARDs is not explicitly defined by the authors of this study. The placebo and adalimumab groups had received 2.1 ± 1.3 and 1.7 ± 0.9 (mean ± SD) DMARDs. The manufacturer provided additional information regarding previous DMARD use in response to the points for clarification raised by the evidence review group. This showed that only 56% of adalimumab patients had previously received treatment with at least two or more DMARDs (Table 1.). Therefore the patient population in this study may not fully match the population for whom anti-TNF therapy would be considered in the UK. With respect to disease severity, the patients had ≥ 3 tender joints and ≥ 3 swollen joints, but the authors do not state whether these symptoms occurred on two separate occasions one month apart, so again the patient population in the trial may not match the UK population. It is stated in the submission that patients continued to receive DMARDs during the trial. If this was the case, this would, in the opinion of two clinical experts, reflect current clinical practice, as if a patient was on a DMARD that was inadequate then the next drug would be added to the DMARD to avoid a flare resulting from stopping the DMARD.[27] Furthermore, patients under the age of 18 years were excluded from the trial. In its guidance, the CHMP states that patients between 16 and 18 years should not be excluded from clinical trials.[4] Therefore, this trial provides no information relevant to this younger age group.

Table 1: Number of different types of previous DMARDs – M02-570

Number of different types of previous DMARDs	Placebo (n=49) N (%)	Adalimumab (n=51) N (%)
1	[redacted]	[redacted]
2	[redacted]	[redacted]
3	[redacted]	[redacted]
4	[redacted]	[redacted]
5	[redacted]	[redacted]
6	[redacted]	[redacted]
Any Previous DMARD	[redacted]	[redacted]

('Academic in confidence')

Is adalimumab more effective than placebo in patients with moderately to severely active psoriatic arthritis who had an inadequate response to disease-modifying anti-rheumatic drug (DMARD) therapy?

In a small (100 patients) short-term (12-week) RCT, adalimumab was more effective, in terms of to the primary endpoint, i.e. the ACR20 response, than placebo in patients with moderately to severely active PsA who had failed prior DMARD therapy. Several secondary endpoints, including effects on disability, the skin component of psoriasis and ACR50 and ACR70 responses, also suggest that adalimumab is more effective than placebo. However, it must be noted that the quality of this trial and the robustness of data cannot be assessed, as this study has not been published in full and full details are not provided. In this 12-week RCT the incidences of adverse effects occurring in ≥ 5% of patients were similar in the placebo and adalimumab groups, with only aggravation of psoriasis and psoriatic arthritis occurring more frequently with placebo than adalimumab. Since only 56% of adalimumab patients had previously received treatment with at least two or more DMARDs the patient population in this study may not be fully representative of the UK population for whom, according to current guidelines, anti-TNF therapy, such as adalimumab, would be considered.

Appendix 2.3. M02-537 (open-label extension)[32-34]

CRITICAL APPRAISAL

Names of Trials: Clinical efficacy and safety of adalimumab for psoriatic arthritis: 48-week results of ADEPT[33]
Adalimumab treatment effects on radiographic progression of joint disease in patients with psoriatic arthritis: results from ADEPT[34]
Adalimumab efficacy in patients with psoriatic arthritis who failed prior DMARD therapy[32]

References:
1. Mease P, Gladman D, Ritchlin et al. Presented at ACR, Annual Scientific Meeting, 2005
2. Mease PJ, Sharp JT, Ory P et al. Presented at EULAR, Annual Scientific Meeting, 2005
3. Genovese MC, Mease PJ, Thomson GTD et al. Presented at EULAR Annual Meeting 2005

Question: Are the efficacy and safety of adalimumab maintained in patients with active psoriatic arthritis?

Summary: In patients with moderately to severely active PsA who had failed NSAIDs or DMARDs, the efficacy of adalimumab appeared to be maintained for 48 and 24 weeks, respectively. However, the robustness of these findings cannot be assessed due to the lack of statistical analyses. The adverse effect profile after adalimumab treatment for 48 weeks was similar to that observed after 24 weeks. The majority of the patients in this trial do not match those for whom anti-TNF therapy, such as adalimumab, are currently recommended, i.e. patients who have failed an adequate trial of two standard DMARDs. As PsA is a chronic condition for which long-term treatment is likely to be necessary, longer term efficacy and safety data are needed.

Introduction
This critical appraisal deals with the results of study M02-537, an open-label extensions of two randomised controlled trials (RCTs) designed to assess the efficacy and safety of adalimumab compared with placebo in patients with moderately to severely active psoriatic arthritis:

- A 24-week, open-label extension of a 24-week RCT in patients who had failed
- NSAID therapy, the ADEPT trial (M02-518), which has been published in full[31] and is critically appraised elsewhere. Data from this open-label extension are published in abstract form only.[33, 34]
- A 12-week, open-label extension of a 12-week RCT in patients who had failed disease-modifying anti-rheumatic drug (DMARD) therapy (trial number M02-570), which has been published in abstract form only[32] and is critically appraised elsewhere.

Did the study ask a clearly focussed question?
Yes – The aim was to establish whether the responses to adalimumab after 24 weeks in the ADEPT trial[33, 34] or 12 weeks in the M02-570[32] trial were maintained for a further 24 or 12 weeks, respectively. Patients who completed 24 or 12 weeks of adalimumab or placebo therapy in these two RCTs were eligible to enter an open-label extension study (M02-537), in which those who received adalimumab 40 mg subcutaneously (SC) every other week (eow) for 24 or 12 weeks continued to receive it for a further 24 or 12 weeks, and those who received placebo for 24 or 12 weeks switched to adalimumab 40 mg SC eow for 24 or 12 weeks. According to these abstracts, 285 and 97 patients (382 in total), who completed the ADEPT[33, 34] and MO2-570[32] trials, respectively, entered this open-label extension study. Patients with an inadequate response to adalimumab after 12 weeks of open-label therapy (i.e. at 36 weeks) could increase the dose to 40 mg every week (30 patients did).[33]

The 48-week outcomes reported in the Mease et al ACR abstract are the ACR20/50/70 responses, and the mean change from baseline in the HAQ.[33] In patients with psoriasis covering ≥ 3% of their body surface area (BSA), the psoriasis area and severity scale (PASI) 50%, 75% and 90% improvements, mean percentage change in PASI, and physician's global assessment (PGA) of psoriasis clear or almost clear at 48 weeks are reported.[33] In the Mease et al EULAR abstract,[34] radiological changes due to PsA of the hands and feet at 48 weeks, assessed by the modified Total Sharp Score (mTSS), are reported. Patients who received adalimumab for 48 weeks were analysed as one cohort across 48 weeks and those who received placebo for the first 24 weeks and adalimumab for the second 24 weeks were analysed as separate cohorts in weeks 1-24 and 24-48.[33] In addition, the ACR and PASI responses of patients who received adalimumab for 48 weeks and used or did not use methotrexate (MTX) at baseline were compared.[33] Adverse effects that occurred in ≥ 5% of patients during weeks 24-48 are reported.[33]

In the Genovese et al 2005 abstract,[32] ACR20/50/70 responses, mean change in HAQ from baseline, and in patients with active cutaneous chronic plaque lesions of a documented history of chronic plaque psoriasis, PGA of clear or almost clear, and mean percentage change in target lesion score at week 24 (i.e. after 12 weeks of open-label therapy) are reported.

Was the study design appropriate?
Yes – As patients with severe disease activity cannot be maintained in a placebo-controlled trial for a long period,[4] continuing the studies with a placebo arm for a further 24 or 12 weeks might not have been ethical or possible. Therefore, an open-label design is appropriate, but as studies in their own rights, these open-label extensions were not powered to show benefit or otherwise. No statistical analyses are presented for any of the results, except the comparison of ACR and PASI response rates after adalimumab treatment for 48 weeks in patients who did and did not use MTX at baseline.[33]

Were participants appropriately allocated to intervention and control groups?
Not applicable - As these extensions were open-label studies, all patients received adalimumab 40mg SC eow.

Were participants, staff and study personnel 'blind' to the participants' study group?
No - As all patients received adalimumab in the open-label phase, neither patients nor observers would have been blinded to treatment. However, the two X-ray readers who assessed the X-ray films for mTSS assessment were blind to treatment and film order.[34]

Were all of the participants who entered into the trial accounted for at its conclusion?
Yes - Tables of the disposition of patients are provided in all three abstracts. Of the 285 patients who entered the open-label extension of the ADEPT trial, 272 (95.4%) completed 48 weeks of treatment,[33, 34] and 92 of the 97 (95%) who entered the open-label extension of the M02-570 trial completed 24 weeks of treatment[32] However, it is not clear how many patients originally randomised to either placebo or adalimumab did not complete the open-label phase of trial M02-570[32] or specifically entered or completed the open-label phase of the ADEPT trial.[33, 34]

Were the participants in all groups followed up and data collected in the same way?
Can't tell - Probably, but too few details are provided to be able to tell.

Was the study large enough?
Can't tell – No statistical analyses, except of MTX users and non-users (see above), are presented.

How are the results presented and what is the main result?

In the abstracts, results for the adalimumab and placebo/adalimumab groups are presented in graphical and/or tabular form. The results for the patients who completed 48 and 24 weeks of adalimumab therapy in the open-label extensions of the ADEPT and M02-570 RCTs, respectively, are summarised in Table 1. The responses after adalimumab treatment for 24 and 12 weeks appeared to be maintained after a further 24 and 12 weeks of treatment. The responses of the groups that received placebo in the RCT and then received adalimumab for 24 and 12 weeks, respectively, in the open-label extension studies increased to levels similar to those seen in adalimumab-treated patients (results not shown). ACR and PASI response rates did not differ significantly between patients taking adalimumab combined with methotrexate and those taking adalimumab alone.

Table 1. Summary of results with adalimumab for 48 weeks in the open-label extension of the ADEPT trial and 24 weeks in the open-label extension of the M02-570 trial.

OUTCOME MEASURE	ADEPT TRIAL (M02-518) Mease et al 2005 ACR,[33] Mease et al 2005 EULAR[34]		M02-570 Genovese et al 2005[32]	
	Adalimumab for 48 weeks (n=151)	Adalimumab for 24 weeks (end of RCT) (n=151)	Adalimumab for 24 weeks (n=51)	Adalimumab for 12 weeks (end of RCT) (n=51)
ACR20/50/70 responses (% of patients)	61%/46%/31%[1]	57%/39%/23%[1]	64%/43%/27%[3]	39%/25%/14%[3]
Mean change in HAQ from baseline	-0.4[1]	-0.4[1]	-0.3[3]	-0.3[3]
PASI 50/75/90 responses (% of patients) (n=69)	70%/58%/46%[1]	75%/59%/42%[1]		
Mean % change in PASI (n=69)	-67%[1]	-66%[1]		
PGA clear or almost clear	63%[1] (n=70)	67%[1] (n=70)	56.3%[3] (n=32)	40.6%[3] (n=32)
ACR 20/50/50 MTX (n=77) c.f. MTX non-use (n=74) at baseline	62%/48%/31% c.f. 58%/43%/31% (p > 0.05)[1]			
PASI 50/75/90MTX (n=29) c.f. MTX non-use (n=40) at baseline	70%/69%/55% c.f. 63%/50%/40% (p > 0.05)[1]			
mTSS (n=128)	0.1[2]	-0.1[2]		
Mean % change in target lesion score (n=32)			-58.8%[3]	-47.0%[3]

How safe were the regimens?

The adverse effect profile with adalimumab for 48 weeks appeared similar to that with adalimumab for 24 weeks.[33] Two additional adverse effects in ≥ 5% of patients were reported in the extension to trial M02-570: cough and nasopharyngitis each in 5/97 (5.2%) patients.[32]

How precise are the results?

Can't tell - It is not possible to assess the precision or robustness of the results as no statistical analyses are presented, except for the results for the MTX users and non-users, and no details of the statistical test used to compare the latter are given in the abstracts.

Can the results be applied to the local population?

Probably - In the UK, according to current guidelines, anti-tumour necrosis factor-α (anti-TNF-α) therapy is recommended for patients with psoriatic arthritis who have failed an adequate trial of two standard DMARDs.[6] Therefore, the population of this extension to the ADEPT trial, i.e. patients with moderately to severely active psoriatic arthritis who had an inadequate response to NSAID therapy, does not match the UK patient population for whom anti-TNF-α therapy, such as adalimumab, would be considered. In trial M02-570, how many

patients had failed two or more DMARDs is not stated, so the patients of this open-label extension may not match the population for whom anti-TNF- therapy would be considered in the UK. Fuller discussion of the patient populations of the trials is presented elsewhere in the critical appraisals of the two RCTs.

Are the efficacy and safety of adalimumab maintained in patients with active psoriatic arthritis?

The efficacy of adalimumab appeared to be maintained for 48 and 24 weeks in patients with moderately to severely active psoriatic arthritis who failed NSAID and DMARD therapy, respectively. However, the robustness and precision of these findings cannot be assessed due to the lack of any statistical comparisons of the open-label results. The profiles of adverse effects occurring in ≥ 5% of patients who received adalimumab for 48 and 24 weeks appeared similar. These extension trials provide limited long-term (maximum 48 weeks) efficacy and safety data. As psoriatic arthritis is a chronic disease for which long-term treatment is likely to be necessary, longer term efficacy and safety data are needed. The Committee for Medicinal Products for Human Use considers that although efficacy may be demonstrated in 12- to 24-week trials, maintenance of the effect in longer trials (e.g. one year) should be demonstrated[4] The patients in these two open-label extension studies have been followed up for longer,[1] but the findings have not been published.

The majority of the patient population in these two extension studies are not, according to current BSR guidelines,[26] representative of the UK population for whom anti-TNF- therapy would be considered.

Appendix 2.4. M04-724 (STEREO)[35]

CRITICAL APPRAISAL

Name of trial: STEREO a prospective, multicentre, multinational, open label study (study no M04-724)[35]

References: Adalimumab (Humira®) is effective and safe in treating psoriatic arthritis (PsA) in real-life clinical practice: preliminary results of the STEREO trial. Presented at ACR, Annual Scientific meeting, 2006.

Abbott Laboratories Ltd. Adalimumab (Humira®) for the Treatment of Moderate to Severe Psoriatic Arthritis. Manufacturer's submission to the NICE STA process. November 2006

Question: What is the safety and efficacy of adalimumab in patients with active psoriatic arthritis, when added to insufficient standard therapy, in real life clinical practice?

Summary: Preliminary results show that adalimumab, when added to insufficient standard therapy, improves ACR20 response in patients with acute Psoriatic Arthritis (PsA) in the short term (12 weeks) in real life clinical practice. Results of several other endpoints, including effects on quality of life, the skin component of psoriasis and ACR50 and ACR70 responses, also show improvement. However the robustness of these findings cannot be assessed due to lack of information on statistical tests performed and lack of inclusion of confidence intervals for mean results. It is also unclear as to what proportion of the patient population used in the trial is the same as that for whom the drug would be recommended in current clinical practice. Analysis of preliminary data does not tell us how adalimumab performs in the longer term or how it compares to other anti-TNF drugs in the same class, or even how it compares to standard DMARDs. This trial does not appear to add any further information on the efficacy and safety of adalimumab that is already available from more robust RCT data.

Introduction

This critical appraisal is based on the preliminary data for the STEREO trial, in the form of an abstract. This trial is currently ongoing so results have yet to be published in full.

Did the study ask a clearly focussed question?

Yes – the study assesses the efficacy and safety of adalimumab in patients with active PsA, in real life clinical practice.

Was the study design appropriate?

Probably - This is a 12-week, prospective, open label study. The aim of the study is to add to the already published RCT evidence by examining the efficacy and safety of adalimumab in large numbers of patients with PsA, in real life clinical practices. The open label design may be appropriate in this situation as patients that were included were those that had failed on standard therapy (including failure on other anti-tumour necrosis factor (TNF) agents) and were classified as still having active disease. Hence a placebo arm may have been deemed unethical.

Patients were included if they were over 18 years of age and had active psoriatic arthritis (defined as $\geq$ 3 tender joints, $\geq$3 swollen joints) despite standard psoriatic arthritis therapy and they had to have a history of an inadequate response or intolerance to at least one disease modifying anti-rheumatic drug (DMARD). Patients also had to be able and willing to self-administer subcutaneous (sc) injections. Patients were excluded if they had a history of cancer or lymphoproliferative disease, HIV, Hepatitis B or C, drug or alcohol abuse. It is unclear whether patients with co-morbidities were included in the open-label trial. The

manufacturer's submission states co-morbidities e.g. uncontrolled diabetes, unstable ischaemic heart disease, IBD and chronic leg ulcer as exclusion criteria. However, the objective of trial as stated in the abstract was to asses the efficacy and safety of adalimumab in real-life clinical practices including patients with various co-morbidities.

Patients received adalimumab 40mg sc every other week (eow) for 12 weeks, in addition to their existing therapy.

The key efficacy outcomes were the American College of Rheumatology 20%, 50% and 70% improvement response (ACR20, ACR50, and ACR70) and change in disease activity score 28 (DAS28) at week 12. For skin symptoms the Physicians global assessment for psoriasis (PGA) was measured at week 12. Other endpoints included responses to the health assessment questionnaire (HAQ) and the dermatology life quality index (DLQI). Safety was evaluated in terms of adverse events reported by the patients and routine safety evaluations at weeks 2, 6 and 12 were conducted. The submission stated that Psoriatic Arthritis Response Criteria (PsARC) would also be measured as an outcome measure; however there appear to be no results for this measure at this stage.

Were participants appropriately allocated to intervention and control groups?
Not applicable – all patients received adalimumab.

Were participants, staff and study personnel 'blind' to participants study group?
Not applicable – as this was an open label trial all patients received adalimumab for the duration of the trial and neither patients nor observers were blinded to treatment.

Were all of the participants who entered into the trial accounted for at its conclusion?
Not applicable – The Trial is still ongoing, as of April 2006 253 patients out of a total of 441 had completed week 12.

Were the participants in all groups followed up and data collected in the same way?
Yes – Study visits were conducted at baseline, week 2, week 6 and week 12.

Was the study large enough?
Can't tell – The submission states that a justification of sample size for the trial was included in the clinical study protocol, however no further detail is provided on this in the abstract. 441 patients have been enrolled and this would appear to be a reasonable number.

How are the results presented and what is the main result?
As of April 2006, 253 patients (52% male), of the total 441 enrolled in the trial had completed week 12. Key efficacy outcomes are presented in tabular and graphical form, showing mean changes at week 12, from baseline, where this data is available. Baseline characteristics of patients, including prior exposure to biologics (anti-TNF agent), were included. In addition graphs comparing response rates in patients who had had prior exposure to a biologic (n=47) and biologic naïve patients (n=185) were included for some of the markers (ACR20, 50 and 70, DAS28 and HAQ).

<u>ACR20 (%)</u>
An ACR20 response was achieved by 72% of patients at week 12. (vs. 38% and 60% at week 2 and 6)

<u>ACR50 and 70 (%)</u>
ACR 50 and 70 responses were achieved by 49% and 27% of patients respectively at week 12 (vs. 14% and 2% at week 2 and 35% and 13% at week 6).

ACR 20, 50 and 70 responses were looked at in biologic naïve patients and patients with prior biologic exposure and adalimumab appeared to have a similar effect in both.

Swollen/Tender joint counts (SJC/TJC)
At week 12 mean tender joint counts was found to be 7.3 (range 0-78) compared to a baseline of 17.6 (possible range 0-78). (vs. 11.6 and 8.7 at weeks 2 and 6 respectively). At week 12 mean swollen joint count was found to be 2.3 (possible range 0-76) compared to a baseline of 9.3 (possible range 0-76). (vs. 5.2 and 3.2 at weeks 2 and 6 respectively)

Disease Activity Score DAS28 (mean)
Mean DAS28 scores showed a decrease (2.6) when compared to baseline (4.8) at week 12. (vs. 3.4 and 2.9 at weeks 2 and 6). This decrease was shown in graph form as mean change from baseline, where the mean change at week 12 was -2.2. This was analysed further, to show mean change from baseline in DAS 28 scores for biologic naive patients (-2.2) and patients with prior biologic exposure (-2.1) at week 12.

Health Assessment Questionnaire (HAQ) scores
Mean HAQ scores showed a decrease over the 12 week period (1.20 to 0.86) and this is illustrated as a graph showing mean change from baseline. The minimum clinically important difference (MCID) is quoted as -0.30, with the actual mean difference reaching -0.34 at week 12. This was analysed further to show mean change from baseline in HAQ scores for biologic naïve patients (-0.34) and patients with prior biologic exposure (-0.36) at week 12.

Physician's Global Assessment of psoriasis (PGA) – total nos of clear/almost clear.(%)
The percentage of patients with a PGA of 'clear or almost clear' increased over the 12 week period, from 35% at baseline to 65% at week 12. (vs. 39% at week 2 and 53% at week 6). This was analysed further to show % of patients with a PGA of 'clear or almost clear' in biologic naïve patients (66%) and patients with prior biologic exposure (63%) at week 12 compared to baseline. (biologic naïve = 36%; prior biologics = 39%)

Dermatology Life Quality Index (DLQI)
Data is only available for baseline and week 12. There is no interim data (i.e. for weeks 2 and 6). DLQI scores showed a decrease which suggests an improvement in quality of life. (Mean baseline scores = 6.4 and mean scores at week 12 = 2.8).

How safe were the regimens?
Preliminary safety data for all 441 patients is presented as a summary.
Adalimumab was thought to be well tolerated overall with only 19 patients (4%) experiencing serious adverse effects (SAE). SAE's, as defined by the investigator included abdominal pain, anaemia, dental abscess, urosepsis, fever with reduced general condition, allergic reaction, severe hip pain and hypersomnia. The spectrum of adverse effects was similar to those highlighted in earlier RCTs and withdrawal rates, so far, appear to be fairly low (8%). No new safety concerns for adalimumab have been brought to light because of this trial.

How precise are the results?
Can't tell – It is difficult to assess the robustness or precision of the results as no statistical analyses is presented in preliminary data. An intention to treat analysis was said to have been undertaken but details of this were omitted from the submission or abstract. In addition results are presented as a mean result without any confidence intervals. Numbers of patients with prior biologic exposure (n=47; 17% of total numbers) is small so graphs comparing prior exposure with no exposure need to be interpreted with caution. Study duration for an open label study is short (12 weeks). In order to assess long term safety and efficacy a longer study duration would have been more appropriate as PsA is a chronic condition. Current CHMP guidelines recommend a trial duration period of 6 months to a year, adjusted according to numbers of patients, to evaluate long term safety and efficacy of a drug in

psoriatic arthritis.[4] Measurement of PsARC at week 12 was included in the submission as an outcome measure, however there appear to be no results for this so far. The CHMP recommends use of the PsARC as a primary outcome measure to measure efficacy in trials, in patients with psoriatic arthritis[4].
Preliminary results for adalimumab on disease markers appear to be favourable however these results do not appear to add any further information to already published RCT data.

Can the results be applied to the local population?
Can't tell - In terms of the population studied, the geographical locations of the trial were in keeping with the general population of the United Kingdom (UK). The British Society for Rheumatology (BSR) has produced guidelines for the use of anti-TNF-α therapy in psoriatic arthritis.[26] Adalimumab is not included in this guideline yet, but the BSR has issued a statement supporting its use in accordance with the guideline.[39] An adequate trial (clearly defined in the guidelines) of two standard DMARDs alone or in combination is recommended, and if patients fail such therapy and still have active joint disease (≥ 3 tender joints and ≥ 3 swollen joints on two separate occasions one month apart), then the next step is to use a licensed anti-TNF-α agent.[26] Patients enrolled in this trial were included if they had failed on at least one DMARD. Further information on DMARD use has not been provided, so it is unclear how many patients of the total study population, would be representative of the UK population for whom anti-TNF-α therapy would be considered.

What is the safety and efficacy of adalimumab when added to insufficient standard therapy in patients with active psoriatic arthritis in real life clinical practice?
Preliminary results show that in the short term (12 weeks) adalimumab appears to be effective and safe in patients with active PsA when added to insufficient standard therapy in real life clinical practice. However, current results do not tell us how this drug performs in the longer term or how it compares to other anti-TNF drugs in the same class, or even how it compares to standard DMARDs. As complete data is not available at this stage, it is difficult to assess the percentage of the patient population used in the trial that would closely reflect the population of patients for whom the drug would be recommended.

Appendix 2.5. Meta-analysis (MO2-518 & MO2-570)[1]

CRITICAL APPRAISAL

Meta-analysis: Meta-analysis of 12-week results of M02-518 and M02-570 trials using ACR20, ACR50, ACR70 and PsARC outcomes.

Reference: Abbott Laboratories Ltd. Adalimumab (Humira®) for the Treatment of Moderate to Severe Psoriatic Arthritis. Manufacturer's submission to the NICE STA process. November 2006

Question: How does adalimumab compare with regards to short-term efficacy with placebo in the treatment of active psoriatic arthritis?

Summary: This meta-analysis showed that in the short term (12 weeks), adalimumab was statistically significantly superior to placebo in active psoriatic arthritis patients with regards to ACR20, ACR50, ACR70 and PsARC improvement. This analysis does not tell us how adalimumab performs in the longer term, or how it compares to other anti-TNF drugs in the same class, or even how it compares to DMARDs and is limited both in the extra information it provides over the findings in the individual trials and the very short term duration of therapy. Patient numbers were significantly larger in the M02-518 trial and have subsequently influenced the findings of the analysis.

Did the study ask a clearly focussed question?
Yes – The meta-analysis considered the efficacy of adalimumab compared to placebo after 12 weeks of treatment.

Did the authors look for the appropriate sort of papers?
Yes – A full literature search was conducted by the authors and has been commented upon elsewhere.

Were the important relevant studies included?
Yes – Only two randomised controlled trials were identified as suitable for analysis by the authors and subsequently included in the meta-analysis.

Was the quality of the studies assessed?
Yes – Quality was considered elsewhere in the submission. One study (M02-570) was only available in abstract form which limited the analysis.[32]

Was it reasonable to combine the results of the review?
Yes – Entry criteria to the studies were essentially the same. One important difference was that in M02-570 patients could use DMARDS other than methotrexate which was the only DMARD allowed in M02-518.[31] 12-week results only were considered due to the open label extension after this period in M02-570. Whilst PsARC outcomes for M02-570 were included in the analysis, they were not available in the only published abstract and could not be checked for accuracy; it is unclear where these results originated from.
The main body of the report contains forest plots for fixed effects relative risk suggesting they are the most appropriate results to be considered. Fixed effects analysis assumes the effect of the risk factor is constant across all included studies i.e. it assumes any variability between studies is exclusively because of random sampling variations. However as has already been commented, there were slight differences in entry criteria with regard to DMARD use and the relative risk for the endpoints under consideration are sufficiently different to suggest the random effects model may be more appropriate.

What is the overall result of the review?
Adalimumab was more effective than placebo in the control of the arthritis components of psoriatic arthritis at 12 weeks.

How precise are the results?
For the random effects model the meta analyses results demonstrate at 12 weeks a relative risk for ACR20 of 3.41 (95% CI, 2.10-5.54), for ACR50 of 10.17 (95% CI, 4.79 – 21.60), for ACR70 of 24.74 (95% CI, 4.88 – 125.47) and for PsARC of [redacted] (95% CI, [redacted]). P values for overall effect are stated to be ≤0.0001 for all analyses.

Can the results be applied to the local population?
As per the critical appraisal for the individual trials.

Were all important outcomes considered?
No – Only the outcomes relating to the arthritis component were included in the meta-analysis as M02-570 did not consider outcomes relating to the psoriasis components of the disease.

How does adalimumab compare with regards to short-term efficacy with placebo in the treatment of active psoriatic arthritis?
This meta-analysis showed that in the short term (12 weeks), adalimumab was statistically significantly superior to placebo in active psoriatic arthritis patients with regards to ACR20, ACR50, ACR70 and PsARC improvement. This analysis does not tell us how adalimumab performs in the longer term, or how it compares to other anti-TNF drugs in the same class, or even how it compares to DMARDs and is limited both in the extra information it provides over the findings in the individual trials and the very short term duration of therapy. Patient numbers were significantly larger in the M02-518 trial and have subsequently influenced the findings of the analysis.

Appendix 2.6. Indirect/mixed treatment comparison

Table 2.6.1: Summary of outcome measures:

Study	Intervention	Primary end-points		Secondary end-points								
Mease PJ et al[37] *(28)*	Etan Placebo	**ACR 20**		**PsARC**	ACR 50 and 70	**PASI** 50 and 75	**SF-36**	**HAQ**	TSS at 6 and 12 months	PhGA of psoriasis		
ADEPT[31]	Ada Placebo	**ACR 20** at week 12	TSS at week 24	**PsARC**	ACR 50 and 70	**PASI** 50 and 75	**SF-36**	**HAQ DI**	**DLQI**	**FACIT - F**	ACR 20 at week 24	
IMPACT II[36]	Inflix Placebo	**ACR response** (20,50 and 70)		**PsARC**	Duration of morning stiffness	**PASI** 50, 75 and 90	**SF-36**	Dactylitis	Enthesopathy			
M02-570[32]	Ada Placebo	**ACR 20**		ACR 50 and 70	**HAQ DI**	PhGA of psoriasis	Target lesion assess.					
Mease et al[38] *(27)*	Etan Placebo	**PsARC**	**PASI** 75	**ACR 20, 50 and 70**	% change in PASI	Target lesion assess.						
IMPACT[62]	Inflix Placebo	**ACR 20** at week 16	**PASI** at week 16 and 50	**PsARC**	DAS 28	Enthesitis	PhGA	**HAQ 28**	CRP	ESR	PtGA	
Salvarani[60]	CSA SSZ Placebo	**ACR 20**		ACR 50 and 70	Spondylitis functional index	Pain score	Morning stiffness duration	PtGA	Modified Schober test	**PASI**		
Fraser[61]	CSA Placebo	Tender joint articular index		TJC	SJC	ESR and / or CRP	**PASI**	**HAQ**	Pt pain assess	PGA	X- ray change	PtGA
Kaltwasser et al[59]	Leflunomide Placebo	**PsARC**		**ACR 20**	**PASI**	Target lesion assess	**HAQ**	**DLQI**				

Outcome measures in **bold** meet the inclusion criteria of including $\geq$ 2 of the following:

PsARC PASI ACR response criteria HAQ disability index Health state utility.

Key: Ada – adalimumab 40mg sc every two weeks, Etan – Etanercept 25mg twice weekly, Inflix – Infliximab 5mg/kg IV infusion, CSA – ciclosporin A, SSZ – sulphasalazine, PhGA – physician's global assessment, PtGA – patient's global assessment. For clarity *(27)(28)* refer to the reference numbers cited in the Abbott submission.

Differences in comparator clinical trials at baseline:

Table 2.6.2: Week 12 (or week 14*/16**)

Study	Intervention	Number of patients assessed for PASI (%)	Baseline PASI score	% achieving PASI 75
Mease 28	Etanercept Placebo	66 (65%) 62 (60%)	Baseline not reported	23% (15 pts) 3%
ADEPT	Adalimumab Placebo	69 (43%) 69 (46%)	7.4 8.3	43% (30pts) 0%
**IMPACT II	Infliximab Placebo	83 (83%) 87 (87%)	11.4 10.2	64% (53pts) 2% (2pts)
M02-570	Adalimumab Placebo	Not assessed	Not assessed	Not assessed
Mease 27	Etanercept Placebo	19 (64%) 19 (64%)	10.1 6.0	26% (5pts) 0%
*IMPACT	Infliximab Placebo	42 (81%) 40 (77%)	5.1 4.2	Reported as % improvement or worsening

Appendix 3. - Review of the guidelines for anti-TNF-alpha therapy in PsA[26]

To be considered for anti-TNF-α therapy the patient needs to have active disease, to have failed to respond to adequate trials of at least two standard DMARDs individually or in combination, and to satisfy none of the exclusion criteria then they should be considered for licensed anti-TNF-α therapy. The definitions given in the guidance for some of the terms used in this statement are as follows.

- The DMARDs listed as standard in the guidance are sulphasalazine, methotrexate, ciclosporin or leflunomide.

- The definition of an adequate trial is
  - Treatment for at least 6 months, of which at least two months is at standard dose (unless significant intolerance or toxicity limits the dose)
  - Treatment for <6 months, where treatment is withdrawn because of drug intolerance or toxicity
  - When treatment is withdrawn because of intolerance or toxicity after >2 months should have been at therapeutic doses.

The guidelines do not give specific treatment response criterion for DMARDs in PsA patients. This should be a combined patient and physician decision after full clinical assessment.

- Active disease is defined as three or more tender joints and three or more swollen joints on two separate occasions at least 1 month apart, based on a 78-tender and 76-swollen joint count.
  - It is accepted that there will be patients with severe symptoms and disability who do not fulfil the guideline criteria and will therefore have to be put forward on a named basis until further evidence becomes available.
  - Dactylitis, where present should be counted as one active joint.
  - Enthesitis should be treated as a separate entity (not covered by these guidelines).

- Exclusion criteria have been adapted from those for anti-TNF-α treatment in rheumatoid arthritis and include
  - Pregnant or breastfeeding
  - Demyelinating disease
  - Active infections
  - Malignancy or pre malignancy states (excluding Basal cell carcinoma, malignancies diagnosed and treated more than 10 years previously)
  - Special caution is recommended in
    - Patients with active psoriasis who have received >1000 joules cumulative dosage of PUVA; particularly those patients who have subsequently been treated with ciclosporin for at least 1 yr.
    - HIV-positive/AIDS patients
    - Congestive cardiac failure (CVS.)/cardiovascular disease.

To measure response to therapy the PsARC response criterion is recommended. The response is defined as an improvement in two factors (with at least one being a joint score) with worsening of none of the following four factors:

- Patient global assessment (on a 0-5 Likert scale)
- Physician global assessment (as above) (improvement defined as decrease by at least 1 unit; worsening defined as increase by at least 1 unit)
- Tender joint score

- Swollen joint score (improvement defined as decrease of at least 30%; worsening defined as an increase of at least 30%).

It is also emphasized in the guidance that patient choice is very important and that anti-TNF-α therapy is not mandatory.

The BSR guidelines for use of anti-TNF-α therapy in PsA were released prior to a licence being granted for adalimumab use in PsA. The BSR issued a statement in January 2006 on adalimumab use for PsA which gave the following information. “We believe that there is clear evidence to support the use of adalimumab as a treatment for adult patients with active and progressive PsA, who have had an inadequate response to previous disease modifying agents, and in accordance with the current BSR guideline for the use of anti-TNF- drugs in PsA Whilst there have not been any direct comparisons between anti-TNF- drugs in PsA, adalimumab appears to be as effective as other licensed agents.

Appendix 4 - Review of CHMP Guideline on clinical investigation of medicinal products for the treatment of psoriatic arthritis[4]

Main domains to be assessed in PsA and instruments to be used in each domain are:

Disease activity in PsA

- Peripheral joint disease activity – modified ACR response which has been demonstrated to be a reliable measure of activity in PsA
- Axial inflammation – measures of activity developed for ankylosing spondylitis e.g. BSADAI

Measure of function

- Self-reported questionnaire is the most extended and preferable approach e.g. HAQ which is not specific for PsA but is a reasonable choice especially in trials that focus on patients with a predominant peripheral PsA disease.

Measure of structural joint damage

- Conventional radiographs
- Scoring methods adapted for PsA e.g. modified Steinbrocker scoring method, modified sharp score, Sharp-Van der Heijde modified scoring method
- PARS – psoriatic arthritis ratingen score
- BASRI – Bath ankylosing spondylitis radiology index
- SASS – stoke ankylosing spondylitis spine score (modified version)
- MRI and ultra-sound have not yet been validated but might be useful in the evaluation of enthesopathy

Other domains and instruments which should be assessed:

- **Skin disease activity.** Demonstration of efficacy on skin disease will require separate specific trials, nevertheless the effect of any new therapy for PsA on skin lesions should be assessed (type of psoriasis, body surface area involved and presence of nail lesions)
- **Enthesitis.** Measured using the Maastricht Ankylosing Spondylitis Enthesis Score Index (MASES) or other instruments which have been validated and are reliable.
- **Biological measures of inflammation.** CRP or ESR may be related to the activity of the disease. However there are no data to support them as useful surrogate variables to assess efficacy in PsA.
- **Quality of life.** PsAQoL or SF-36 have both been tested in PsA and were found to be reliable, valid and responsive to change. The effect of arthritis and psoriasis on health related quality of life should be assessed independently.
- **Global assessment of disease activity.** Patient and / or physician's global assessment measured by means of a visual analogue scale.

Appendix 5 - Search strategy - PsA economic evaluations

1. NHS EED (CRD interface)

	Search	Matching records
# 1	MeSH Antirheumatic Agents	125
# 2	adalimumab OR humira OR D2E7	18
# 3	MeSH Methotrexate	70
# 4	methotrexate OR rheumatrex OR trexall OR MTX	186
# 5	leflunomide OR arava	21
# 6	sulfasaline OR sulphasalazine OR azulfidine	15
# 7	infliximab OR remicade	72
# 8	etanercept OR enbrel	57
# 9	DMARD OR "disease modifying antirheumatic drug"	34
# 10	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9	326
# 11	MeSH Arthritis, Psoriatic	9
# 12	"psoriatic arthritis"	15
# 13	#11 or #12	15
# **14**	**#10 and #13**	**12**

Of the12 results 8 are HTA and 4 are DARE- **none are NHS EED records**.
Given the nil result using all of the search terms, a search was conducted using a much broader search strategy designed to capture all NHS EED records relating to psoriatic arthritis;

	Search	Matching records
# 1	MeSH Arthritis, Psoriatic	9
# 2	"psoriatic arthritis"	15
# **3**	**#1 or #2**	**15**

Of the15 results-8 are HTA, 4 are DARE and **3 are NHSEED records**.

2. OHE HEED (October 2006 CD-Rom)
'psoriatic arthritis' or 'psoriasis arthritis' (all fields)

3. EconLit (via WebSpirs)
(psoriatic or psoriasis) AND arthritis (all fields)

4. CAIRS T (NHS EED administrative system)
Limit n
psoriasis(1w)arthritis or psoriatic(1w)arthritis (all fields)

Purpose of search was to identify all papers considered for inclusion on NHS EED rather than just those included on NHS EED.

5. CINAHL (via SiverPlatter) updates 01/09/06-27/10/06
Combined NHS EED economic search filter (see below) with;
(UD = 20060901-20061027)
(psoriasis or psoriatic) and arthritis) in ti,ab
“Arthritis-Psoriatic"/ all subheadings

Purpose of search was to identify any relevant papers from the latest updates available for CINAHL that have not yet been screened for NHS EED and added to the CAIRS T system.

6. MEDLINE and MEDLINE IP (via SilverPlatter) updates 25/09/06-27/10/06
Combined NHS EED economic search filter (see below) with;
(UD = 20060925-20061027)
(psoriasis or psoriatic) and arthritis) in ti,ab
“Arthritis-Psoriatic"/ without-subheadings

Purpose of search was to identify any relevant papers from the latest updates available for MEDLINE that have not yet been screened for NHS EED and added to the CAIRS T system.

NB. An update search of EMBASE was considered but was not necessary as EMBASE updates are added weekly and the latest update available would have been included in the CAIRS T search above (4).

Results

Database (search no.)	Records	After de-dupe	Custom 4 field
NHS EED (1)	3	3	NHSEED 13/12/06
OHE HEED (2)	4	3	HEED 13/12/06
EconLit (3)	0	0	
CAIRS T (4)	12	9	CAIRS T 14/12/06
CINAHL (5)	0	0	
MEDLINE (6)	2	1	MEDLINE 14/12/06

NHS EED economic search filter;

1 economics / all SUBHEADINGS in MJME,MIME
2 explode "costs and cost analysis" / all SUBHEADINGS in MJME,MIME
3 value of life / all SUBHEADINGS in MJME,MIME
4 economics dental / all SUBHEADINGS in MJME,MIME
5 explode "economics hospital" / all SUBHEADINGS in MJME,MIME
6 economics medical / all SUBHEADINGS in MJME,MIME
7 economics nursing / all SUBHEADINGS in MJME,MIME
8 economics pharmaceutical / all SUBHEADINGS in MJME,MIME
9 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8
10 (econom* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic*) in ti,ab

11 (expenditure* not energy) in ti,ab
12 (value near1 money) in ti,ab
13 budget* in ti,ab
14 #10 or #11 or #12 or #13
15 #9 or #14
16 letter in pt
17 editorial in pt
18 historical-article in pt
19 #16 or #17 or #18
20 #15 not #19
21 ANIMALS in TG
22 HUMANS in TG
23 #21 not (#21 and #22)
24 #20 not #23
25 (metabolic near cost) in ti,ab
26 ((energy or oxygen) near cost) in ti,ab
27 #24 not (#25 or #26)
28 addiction in so
29 Am-J-Manag-Care in so
30 Am-J-Public-Health in so
31 Ann-Intern-Med in so
32 Arch-Intern-Med in so
33 BMJ in so
34 Br-J-Gen-Pract in so
35 CMAJ in so
36 Cochrane-Database-Syst-Rev in so
37 Control-Clin-Trials in so
38 Health-Aff in so
39 Health-Econ in so
40 Health-Serv-J in so
41 Health-Serv-Res in so
42 Health-Technol-Assess in so
43 Int-J-Qual-Health-Care in so
44 Int-J-Epidemiol in so
45 Int-J-Technol-Assess-Health-Care in so
46 JAMA in so
47 J-Adv-Nurs in so
48 J-Epidemiol-Community-Health in so
49 J-Health-Econ in so
50 J-Health-Polit-Policy-Law in so
51 J-Health-Serv-Res-Policy in so
52 J-Med-Ethics in so
53 J-Public-Health-Med in so
54 J-Stud-Alcohol in so
55 J-R-Soc-Med in so
56 Lancet in so
57 Med-Care in so
58 Med-Decis-Making in so
59 Milbank-Q in so
60 N-Engl-J-Med in so
61 Nurs-Stand in so
62 Nurs-Times in so
63 Palliat-Med in so
64 Pharmacoeconomics in so
65 Prev-Med in so

66 Prof-Nurse in so
67 Public-Health in so
68 Qual-Assur-Health-Care in so
69 Qual-Life-Res in so
70 Soc-Sci-Med in so
71 Stat-Methods-Med-Res in so
72 Value-Health in so
73 world-health-forum in so
74 #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39 or #40 or #41 or #42 or #43 or #44 or #45 or #46 or #47 or #48 or #49 or #50
75 #51 or #52 or #53 or #54 or #55 or #56 or #57 or #58 or #59 or #60 or #61 or #62 or #63 or #64 or #65 or #66 or #67 or #68 or #69 or #70 or #71 or #72 or #73
76 #74 or #75
77 #27 not #76

Appendix 6 - Points for clarification raised by the ERG

Section A. Clarification on effectiveness data

A1. General Points

- The diagnostic criteria for psoriatic arthritis have not been made clear in the submission. Clear diagnostic criteria are required to identify the patient population and give an indication of incidence – is this available from the manufacturer?

- Clarify the definition of inadequate or failure of previous DMARD therapy used for the trials?

- Are there data available on patients not having used DMARDs and those having failed 2 or more DMARDs?

- Pg 10 – Ref 2 accessed in Feb 2005 should this be 2006?

- Pg 18 – section 5.2.2 Inclusion and exclusion criteria of relevant randomised controlled trials (RCTs). Exclusion criteria: 'M02-537 is an extension of the two relevant RCTs M02-517 and M02-518' M02-517 is not discussed in the document, should this read M02-570?

- Pg 19 – Discrepancy in number of RCTs retrieved for more detailed evaluation (n = 39, with 37 excluded) but total number of trials listed in box is 35. Please provide a corrected diagram.

A2. Issues related to study M02-518

- Please provide tables similar to table 5.9.2.1 for response parameters at 12 and 24 weeks for study M02-518 based on the subgroups of patients with psoriasis <3% or >=3% at baseline.

- Please provide additional data on the mean (SE) HAQ change at 12 and 24 weeks for the subgroups of patients based on previous DMARD use (<2 and >=2) and for patients with psoriasis <3% and >=3%.

- Please report the mean (SE) HAQ change for PsARC responders/non-responders.

- Please state the number of patients in each group that received rescue therapy.

- Pg 33 - Table 5.4.2 – please state the numbers of patients assessed for modified PsARC and PGA.

- Pg 35 - Table 5.4.6 ACR and PASI response rates and HAQ at weeks 12 and 24 shows no data for week 12. This table also shows 74 patients in each subgroup (total 148 for adalimumab) assessed for PASI, yet only 69 patients at baseline were eligible for evaluation. Please clarify and also provide p values for this data set

- Pg57 - Table 5.9.2.1 shows the numbers of patients with respect to history of treatment failure with at least 2 DMARDs – does this include patients on concomitant methotrexate?

- Please state the % of patient in this study that have never received a DMARD.

A3. Issues related to study M02-570

- Please provide tables similar to table 5.9.2.1 for response parameters at 12 for study M02-570 based on the subgroups of patients according to DMARD use at baseline (yes/no) and based on psoriasis <3% or >=3% at baseline. Please report the mean (SE) HAQ change at 12 weeks for these additional subgroups.

- Please report the mean (SE) HAQ change for PsARC responders/non-responders at 12 weeks.

- Pg 21 – Table 5.3.1.2 Study Design: M02-570
  Supporting medications:
  Prednisolone at a dose of < 10mg daily
  Oral corticosteroids $\underline{\geq}$ prednisolone 10mg/day
  Is this correct?

- Please state the % of patients in this study that had failed at least 2 DMARDS.

- Subgroup analysis according to concomitant DMARD use would be useful.

A4. Issues related to study M02-573

- Pg 39 – In the Mease abstract - Clinical efficacy and safety of adalimumab for psoriatic arthritis: 48 week results of ADEPT (not provided by Abbott) it states that 285 patients continued into study MO2-537, yet table 5.4.13 shows inclusion of 298 subjects – please clarify.

- Pg 40 – Figure 5.4.3 does the data from the placebo group refer to those given placebo for 12 weeks then adalimumab for the remainder of the study? This section needs clarification.

A5. SAFETY

- No safety data is provided for treatment beyond 24 weeks. Please provide tabulated safety data from study M02-537.

- Pg 50 - Table 5.7.1 - provide a definition of (any) severe adverse event.

- Are there any data on the appearance of antibodies to adalimumab?

- Are there any post-trial data on side effects when adalimumab is used in clinical practice for psoriatic arthritis?

Section B: Indirect/Mixed treatment comparisons

B1.

- Please report change in HAQ scores (e.g. mean (SE)) reported in the studies included in Table 5.6.3.

- Present response results (as in Table 5.6.3 and including change in HAQ scores) for the 12 week trials excluded from the cost-effectiveness study reported in Table 5.6.2.

- Please undertake a full synthesis of PsARC response and other outcome parameters at 12 weeks for etanercept, infliximab and adalimumab using data from all available studies (including both 12 and 24 week studies).

- Please provide the response data that are entered into the model, including the adjustment for the differing skin involvement in the patient populations of the trials.

Section C. Economic Analysis

C1. Please supply further information on the use of Microsoft Solver to predict the correlation between ACR/PASI/PsARC responses. It is not transparent how the estimated responses in Tabels A5.3 to A5.6 have been calculated. In addition please provide the Excel sheets for these calculations.

C2. Please explain further how you used the above information to "identify the most likely distribution of responses" for the rest of treatment options, and how you sample the type of response for the next timepoint from the joint distribution of ACR, PASI and PsARC. In addition, please clarify:

- How do you generate the aggregate responses for etanercept, infliximab, conventional DMARDs and MTX reported in the tables in Appendix 5

- What is the uncertainty around those estimates

- Where exactly these efficacy estimates are used in the model

C3. Please confirm that you are not using data from the adalimumab trial M02-570 to inform the model. If that's correct, could you elaborate further on the reasons why this trial was excluded from the economic analysis?

C4. In order to test the robustness of the model, could you please:

- Run the original model and present results for the following subgroups:
    1. Use of previous DMARDs at baseline
    2. Patients with (BSA ≤ 3%) and without skin involvement (BSA > 3%)

- Re-run the base case analysis based on a synthesis of 12-week PsARC data from trials MO2-518 and MO2-570, or at least present an additional sensitivity analysis using PsARC estimates at 12 weeks from the M02-570 trial.

- Re-run the base case analysis based on a complete synthesis of all 12-week PsARC data for all comparators (based on all trials of at least 12 weeks duration).

- Re-run the base case analysis without the "conventional DMARDs" drug cost.

C5. Please elaborate further on the differences in terms of patient characteristics between the Leeds cohort study and the Toronto dataset and discuss how these relate to the characteristics of the patients recruited into the adalimumab trials.

C6. Could you please be explicit about which are the two DMARDs failed before the hypothetical patient enters the model? Could you please exclude the corresponding drug costs from the weighted average cost used for the "conventional DMARD" option?

C7. Please clarify whether the withdrawal rates calculated based on the BIOBADASER study were estimated using the sub-sample of PsA patients contained in that registry. Please clarify whether the withdrawal rates calculated based on the Flendrie et al. study were estimated using PsA or RA patients.

C8. The model assumes that HAQ worsening (rebound effect) occurs immediately at the point of withdrawal and is equivalent to the initial HAQ gain. Please run the base case analysis for the more conservative assumption of rebound back to natural history (i.e. when a patient fails therapy, their HAQ returns to what it would have been had they not been treated).

C9. Please supply further clarification on the methods used to estimate EQ5D scores from SF-12 responses. In particular, please clarify whether you are building uncertainty between those estimates.

C10. Please provide a table summarising utility values for trials MO2-518 and MO2-570 based on SF-6D and EQ-5D scores. In particular can you provide the following information (e.g. mean (SE)) for both treatment arms.

- Baseline utilities (all patients and for subgroups defined according to previous DMARD use and for patients with <3% and >=3% psoriasis at baseline)
- Change in utilities based on PsARC response/non-response at 12 weeks
- As above but also for subgroups defined according to previous DMARD use and for patients with <3% and >=3% psoriasis at baseline.

C11. Please explain what is the rationale behind the finding that after excluding the skin component of the analysis adalimumab still dominates etanercept, when the latter shows a better PsARC response rate at 24 weeks (70% for etanercept compared to 60% for adalimumab).