Evidence Review Group's Report

Title: Golimumab for the treatment of psoriatic arthritis

Produced by CRD and CHE Technology Assessment Group, University of York, Heslington, York, YO10 5DD

Authors

Huiqin Yang, Research Fellow, CRD

David Epstein, Honorary Visiting Research Fellow, CHE

Laura Bojke, Research Fellow, CHE

Dawn Craig, Research Fellow, CRD

Kate Light, Information Specialist, CRD

Ian Bruce, Professor of Rheumatology and Honorary Consultant Rheumatologist, University of Manchester

Mark Sculpher, Professor of Health Economics, CHE

Nerys Woolacott, Senior Research Fellow, CRD

Correspondence to Dr Huiqin Yang, Research Fellow, CRD

Date completed 11th August 2010

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 09/120/01.

Declared competing interests of the authors

Mark Sculpher has a financial interest in a consulting company which has undertaken work for Abbott, Schering-Plough and Wyeth, but not relating to psoriatic arthritis, and he has not personally participated in this consultancy work. He has personally undertaken paid consultancy for some of the comparator manufacturers, again not relating to psoriatic arthritis.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

Submitted to NICE 11th August 2010

This report should be referenced as follows:

CRD and CHE Technology Assessment Group. Golimumab for the treatment of psoriatic arthritis: a single technology appraisal. Centre for Reviews and Dissemination/ Centre for Health Economics, 2010.

Contributions of authors

Huiqin Yang wrote the clinical effectiveness sections of the report. David Epstein, Laura Bojke and Dawn Craig wrote the cost effectiveness sections of the report and performed the economic analyses and evidence synthesis. Kate Light wrote the sections on the search strategies. Ian Bruce commented on the report and provided clinical advice. Nerys Woolacott and Mark Sculpher commented on and revised the report.

Table of Contents

List of Tables

List of Figures

List of abbreviations

ACR	American College of Rheumatology response criteria
ADEPT	Adalimumab Effectiveness in Psoriatic Arthritis Trial
AE	Adverse event
BAD	British Association of Dermatologists
BNF	British National Formulary
BSA	Body surface area
BSR	British Society of Rheumatologists
CEAC	Cost-effectiveness acceptability curve
CI	Confidence interval
CrI	Credible Interval
DMARD	Disease modifying anti-rheumatic drug
ERG	Evidence Review Group
EQ-5D	EuroQol-5D
HAQ	Health Assessment Questionnaire
HRQoL	Health-related quality of life
ICER	Incremental cost-effectiveness ratio
IMPACT	Infliximab Multinational Psoriatic Arthritis Controlled Trial
ITT	Intention-to-treat
MS	Manufacturer's submission
MTC	Mixed treatment comparison
N/A	Not applicable
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
NSAID	Nonsteroidal anti-inflammatory drug
OMERACT	Outcome Measures in Rheumatoid Arthritis (Rheumatology) Clinical trials
PASI	Psoriasis Area and Severity Index
PRESTA	Psoriasis Randomized Etanercept Study in Subjects with Psoriatic Arthritis
PsA	Psoriatic arthritis
PSA	Probabilistic sensitivity analysis
PsARC	Psoriatic Arthritis Response Criteria
QALYs	Quality adjusted life years
RCT	Randomised Controlled Trial
RR	Relative risk
SPC	Summary of product characteristics
STA	Single Technology Appraisal
TNF	Anti-tumour necrosis factor

1. SUMMARY

1.1. Scope of the submission

This report presents the ERG's assessment of the manufacturer's (Schering-Plough) submission to NICE on the use of golimumab for the treatment of active and progressive psoriatic arthritis (PsA) in patients who have responded inadequately to previous disease-modifying anti-rheumatic drugs (DMARDs). The report includes an assessment of both the clinical and cost effectiveness evidence submitted by the company.

The manufacturer's evaluation of clinical efficacy included evidence relating to monthly golimumab therapy versus placebo, a comparison of the relative efficacy between anti-tumour necrosis factor (TNF) agents which included evidence relating to the other three relevant comparators (etanercept, infliximab and adalimumab), and a decision analysis assessing the cost-effectiveness of golimumab compared with etanercept, infliximab, adalimumab and palliative care.

1.2. Summary of submitted clinical effectiveness evidence

The main clinical effectiveness data were derived from a single phase III randomised controlled trial (RCT), GO-REVEAL, that compared golimumab with placebo for treating active and progressive PsA patients who were symptomatic despite the use of current or previous DMARDs or nonsteroidal anti-inflammatory drugs (NSAIDs). The 14 week data from this trial showed that, compared with placebo, golimumab 50 mg significantly improved joint disease response as measured by ACR 20 (RR 5.73, 95% CI 3.24 to 10.56) and PsARC (RR 3.45, 95% CI 2.49 to 4.87), and skin disease response as measured by PASI 75 (RR 15.95, 95% CI 4.62 to 59.11). The 24 week absolute data showed that these treatment benefits were maintained. There was a statistically significant improvement in patients' functional status as measured by HAQ change from baseline at 24 weeks (-0.33, $p<0.001$), thereby achieving the minimum clinically significant threshold for PsA (-0.3).[1] Golimumab 100 mg significantly achieved a similar magnitude of treatment effects at 14 and 24 weeks. The open-label extension data showed that these beneficial effects were also maintained at 52 and 104 weeks.

In the absence of head-to-head comparisons of the relative efficacy between different anti-TNF agents, the manufacturer conducted a mixed treatment comparison (MTC) analysis to estimate the relative efficacy of these four anti-TNF agents. The results showed that infliximab appears to be the

most effective of the four anti-TNF agents: infliximab was associated with the highest probabilities of response in terms of joint and skin disease outcomes. Golimumab achieved the third highest PsARC response (joint disease), and golimumab and etanercept had the lowest response for skin disease in terms of PASI change from baseline. In those patients who achieved a PsARC response, the highest mean improvement in the functional status (HAQ) was seen with etanercept (-0.703), and the lowest mean improvement in HAQ was seen with golimumab (-0.424). For all four anti-TNF agents, the changes in HAQ for those patients who did not achieve a PsARC response were below the minimum clinically significant threshold (-0.3).[1]

Short-term radiographic data from the GO-REVEAL trial indicated that golimumab 50 mg significantly slowed joint disease progression during the 24 weeks (p=0.01). There was a lack of follow-up radiographic data to determine whether these effects persisted in the longer term.

The limited available evidence for the safety evaluation from the single GO-REVEAL trial suggested that the most frequently reported adverse events associated with golimumab therapy were infections and infestations, upper respiratory tract infection and nasopharyngitis. Serious adverse events including serious infection and malignancy were rare. No active tuberculosis in any treatment arm was observed.

1.3. Summary of submitted cost effectiveness evidence

The manufacturer's submission (MS) presented a decision model to compare etanercept, infliximab, golimumab, adalimumab versus palliative care for patients with psoriatic arthritis. In the base-case model, 73% of the cohort of patients were assumed to have significant psoriasis (>3% body surface area). Estimates of the effectiveness of anti-TNF agents in terms of PsARC, HAQ change and PASI change were obtained from a MTC analysis of RCT data.

Patients in the model were assumed to continue with biologic therapy after 12 weeks if they achieve PsARC response. HRQoL and costs were a function of HAQ and PASI score. The acquisition costs of anti-TNF agents were estimated from the British National Formulary.[27] The acquisition cost of golimumab was assumed equal to that of adalimumab.

The original MS base-case model was revised following requests for clarifications from the ERG. The revised MS model amended the functional form of the utility algorithm linking HAQ and PASI to HRQoL. The revised model also assumed that infliximab was administered without vial sharing.

The revised decision model found that the ICER of golimumab versus palliative care was just under £20,000 per QALY. However, the comparison to palliative care does not meet the NICE requirement for an incremental cost-effectiveness analysis to be conducted, in which each strategy should be compared with the next best alternative.

1.4. Commentary on the robustness of submitted evidence

1.4.1. Strengths

The manufacturer's systematic review of the literature used appropriate search methods to identify the relevant evidence of golimumab for the treatment of active and progressive PsA despite the previous use of DMARDs. The key findings were derived from a single double-blind phase III RCT (GO-REVEAL) which was conducted in a relevant population and the dosing regimen (including dose adjustment) for the golimumab 50 mg group was generally reflective of clinical practice. The results from the 14 week data analyses of this trial were considered to be robust.

Regarding the manufacturer's MTC analysis, the degree of clinical heterogeneity between the included trials in terms of joint and skin disease severity and functional status was considered reasonable. Therefore, the assumption of exchangeability between the trials for the purposes of the MTC analysis was acceptable. The ERG also considered the statistical approach in the manufacturer's MTC analysis to be reliable.

The MS report of the economic model was generally clearly written. The MS responded to all requests for clarifications and amendments to the economic analyses. The output of the spreadsheet model corresponded with the results reported in the MS. The model took account of all the important elements of the decision problem, in terms of the rules for the continuation of biologic therapy, the natural history of arthritis and psoriasis in these patients, the treatment effects, the relationship between psoriasis, arthritis and HRQoL, and the costs.

1.4.2. Weaknesses

The manufacturer did not adequately apply the intention-to-treat approach for the efficacy analysis in the MS. The ERG requested the clarification of such analyses presented in the MS. Based on these data further provided by the manufacturer, it appears that the intention-to-treat analysis was appropriately performed for most efficacy outcomes. However, it should be noted that analyses of PASI 50 and PASI 90 at 14 weeks and all the PASI outcomes at 24 weeks were not performed on the basis of intention-to-treat analysis. Therefore, such analyses may have potentially compromised the internal validity of the results in terms of these skin disease outcomes.

There was a further concern about the robustness for the analyses on the 24 week data in the GO-REVEAL trial. Whilst the analyses at 24 weeks involved all the intention-to-treat data from the randomisation, it appears that these intention-to-treat analyses failed to adjust the treatment contamination due to patients' crossing-over at week 16. Therefore, this may have threatened the internal validity of trial results for all the efficacy and safety outcomes at 24 weeks.

In terms of the safety evaluation, the manufacturer did not present data to facilitate a comparison between the adverse events of golimumab with those of the comparator anti-TNF agents. The longer-term follow-up safety data (e.g. at 52 and 104 weeks) from the GO-REVEAL trial were not available. Furthermore, the manufacturer failed to consider adverse event data of golimumab from controlled studies in other conditions such as rheumatoid arthritis and ankylosing spondylitis.

Regarding the economic evaluation, there was concern about the robustness of the estimation of costs associated with psoriasis which was based on a survey of 22 dermatologists. The MS stated that, based on the survey, the cost per PASI point was [redacted] if phototherapy is excluded and [redacted] if phototherapy is included as a treatment for psoriasis. This implies that reducing PASI from, say, 9.9 to 3.3 (a reduction of 6.6 points estimated for infliximab) would reduce the expected cost of treating psoriasis per year by [redacted] if phototherapy was used and by [redacted] if phototherapy is not used. However, the MS provided insufficient detail of these calculations for the ERG to check whether these costs were valid or not. No estimates of variability or sampling uncertainty were provided. The manufacturer provided raw data on request for clarification, but did not show the unit costs or detail of how the results of the survey were synthesised and how the mean cost per PASI point was calculated. However, the raw data did show that there was considerable heterogeneity in resource use among the clinicians surveyed.

The MS did not correctly calculate the ICERs used to compare the cost-effectiveness of the treatments. The MS did not exclude extendedly dominated alternatives. The ERG recalculated the ICERs using the results of the MS model. The corrected ICER from the MS model for etanercept versus palliative care is about £17,000 per QALY. According to the MS model, with the ICERs correctly calculated, other anti-TNF agents (golimumab, adalimumab and infliximab) are not cost-effective, because they are either dominated or extendedly dominated by etanercept.

1.4.3. Areas of Uncertainty

Whilst MTC analyses provide evidence of the relative efficacy of these anti-TNF agents, those findings may be considered more uncertain than would be provided in head-to-head RCTs. In terms of the results of manufacturer's MTC analyses, the credible intervals of most outcomes for all four anti-TNF agents overlapped each other. In particular, there were substantial uncertainties for the estimates of PASI change from baseline due to a small sample size of patients evaluable for psoriasis. Furthermore, no comparisons of anti-TNF agents in the MTC were performed for the treatment duration beyond 14 weeks, since only 12-14 week data from the included trials were used to establish the relative efficacy between these anti-TNF agents. Additionally, there were no comparisons for disease progression (radiographic data) in the MTC analyses.

Despite most patients in the included trials of MTC being under licensed conditions, it was noteworthy that the majority of patients had previously received at least one DMARD, and no trial specified the failure to respond to at least two DMARDs (patients whom the current BSR guidelines and NICE guidance for etanercept, infliximab and adalimumab consider eligible for the biologic treatment) as a recruitment criterion. Thus, the trial participants were likely to have less severe disease of PsA compared to those patients receiving biologic treatment in routine practice. As trial participants were not precisely representative of the active and progressive PsA population recommended for anti-TNF agents by the current guidelines, it remained unclear that the beneficial effects observed in these trial participants were similar in those treated in routine clinical practice.

In terms of the safety evaluation, the longer-term follow-up safety data at 52 and 104 weeks from the GO-REVEAL trial would be valuable, though the manufacturer stated that these data were not available at the time of the ERG's request.

A key area of uncertainty is whether the anti-TNF agents should be considered equally clinically effective, that is, to treat them as a class. This was the position adopted by the recent guidance issued by NICE regarding the previous appraisal of etanercept, adalimumab and infliximab for psoriatic arthritis.[2] If all anti-TNF agents are considered equally effective (in terms of PsARC, HAQ and PASI responses) then etanercept, adalimumab and golimumab have very nearly equal costs and equal QALYs and all have an ICER of about £15,000 per QALY versus palliative care. Infliximab has higher acquisition costs and would be dominated by other biologic strategies under an assumption of equal effectiveness.

Other areas of uncertainty are: alternative estimates of clinical effectiveness in terms of PsARC; HAQ change and PASI change from the ERG evidence synthesis; alternative estimates of cost of

administration of drugs; alternative values for NHS cost of psoriasis, measured by PASI; alternative utility functions; and trying a higher dose of golimumab for patients who do not achieve adequate response at 12 weeks, according to the licence.

To address these uncertainties, the ERG conducted sensitivity analyses using the spreadsheet model provided in the MS. None of these sensitivity analyses changed the conclusion that golimumab is extendedly dominated by etanercept. In the scenario using the ERG estimates of clinical effectiveness in the MS model, the ICER of etanercept increased to just over £20,000 per QALY and the ICER of infliximab was £56,000 per QALY.

Further analyses were also conducted using the ERG model developed by York Assessment Group during the recent appraisal of etanercept, infliximab and adalimumab. These analyses were used to validate the MS model by comparing the results to an independently constructed model. The MS model and the ERG alternative model have a broadly similar structure and data inputs and gave similar results.

The licence for golimumab indicates that patients over 100 kg in weight who fail to respond to golimumab 50 mg at 3 months can be trialled on a higher dose of 100 mg. A full economic analysis of this option could not be undertaken because of a lack of clinical data for this subgroup of patients. The ERG notes that, if patients are titrated and maintained on a higher dose, the additional acquisition costs will be around £2145 per 3 months. However, the clinical adviser to the ERG suggests that, in practice, this scenario is unlikely because of the additional cost; eligible patients are more likely to be tried on an alternative biologic.

A remaining source of uncertainty is the annual cost of treating psoriasis. Although the MS conducted a survey of dermatologists and presented the raw data from the survey, there was no detail of the statistical method used to calculate the mean costs from the raw data and, therefore, the ERG could not validate the calculations. However, the ERG conducted sensitivity analysis on the PASI cost using the ERG model. Doubling or halving the [redacted] did not materially affect the results of the ERG model, indicating that this is not a key parameter for the decision, at least in patients who do not have severe psoriasis.

1.5. Key Issues

The data from the GO-REVEAL trial provide evidence to suggest that golimumab appears to be an efficacious treatment for active and progressive PsA patients despite the use of previous DMARDs or

NSAIDs. The effect sizes of point estimates of joint and skin disease response and functional status were moderate to large, implying that these treatment effects could be clinically significant.

The main limitation of the efficacy evaluation of golimumab, however, is that there were limited efficacy data available. The analyses for efficacy outcome were limited to only one RCT (GO-REVEAL) with limited sample size. In particular, few patients provided data on the psoriasis response to golimumab treatment.

The radiographic outcomes in the GO-REVEAL trial were evaluated over the short follow-up period of 24 weeks, which were often considered inadequate to assess radiographic changes in response to the treatment. There was a lack of long-term efficacy data of radiographic assessments. Given the fact that the treatment effect on the joint disease is more accurately reflected by the more objective radiographic measure, radiographic long-term data could be valuable to provide more generalisable estimates of the treatment effect in responding to golimumab therapy.

The ERG further considered the evidence for the safety evaluation of golimumab to be inadequate. The evidence was exclusively based on 24 week data from the single RCT with PsA patients (GO-REVEAL). The manufacturer failed to provide longer term data or to consider adverse event data of golimumab from controlled studies in other conditions such as rheumatoid arthritis and ankylosing spondylitis. Whilst the adverse effects profile of golimumab appears similar to other anti-TNF agents, the longer-term safety profile of golimumab remained uncertain. Given these limitations and uncertainties, the manufacturer's conclusion that golimumab is a safe treatment option similar to other anti-TNF agents may be premature and may not be reliable.

Despite the claim made by the manufacturer that golimumab is a cost-effective treatment option, the manufacturer's own model showed that golimumab is not cost-effective, when the ICERs are correctly calculated. None of the sensitivity analyses carried out by the manufacturer or the ERG regarding uncertainty in the estimates of clinical effectiveness, the acquisition and administration cost of drugs, the cost of treating psoriasis and the utility functions estimated to generate health outcomes changed this conclusion.

However, a key area in determining the cost-effectiveness of anti-TNF agents is whether they should be considered equally clinically effective, that is, to treat them as a class. This was the position adopted in the guidance issued by NICE following the previous appraisal of etanercept, adalimumab and infliximab for psoriatic arthritis.[2] If all anti-TNF agents are considered equally effective (in terms

of PsARC, HAQ and PASI responses) then etanercept, adalimumab and golimumab have very nearly equal costs and equal QALYs and all have an ICER of about £15,000 per QALY versus palliative care. Infliximab has higher acquisition costs and is dominated by other biologic strategies if they are considered equally effective. The possibility that all anti-TNF agents are equally clinically effective was addressed in the manufacturer's submission in response to a request by the ERG for this sensitivity analysis to be carried out.

2. BACKGROUND

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

The manufacturer provided a brief but accurate summary of the key issues relating to active and progressive PsA.

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

The manufacturer's submission (MS) provided an accurate overview of the treatment pathway for active and progressive PsA patients, based on the current clinical pathway recommended by British Society of Rheumatology (BSR).[3] The manufacturer provided details of which anti-TNF agents (etanercept, infliximab and adalimumab) are currently used for the treatment of active and progressive PsA in UK clinical practice. However, the manufacturer did not give details of what proportion of patients receive each biologic in routine practice.

3. CRITIQUE OF MANUFACTURER'S DEFINTION OF DECISION PROBLEM

3.1. Population

In the statement of the decision problem, the manufacturer specified the relevant population as people with active and progressive PsA who have responded inadequately to previous disease-modifying anti-rheumatic drugs (DMARDs). This exactly reflects the population specified in the NICE scope. The key trial which comprised the direct trial evidence of efficacy of golimumab (GO-REVEAL[4, 5]) was relevant to this decision problem, being limited to patients with a diagnosis of PsA for at least 6 months prior to first study agent administration who had active PsA despite the use of current or previous DMARDs or NSAIDs.

The manufacturer's MTC analyses included trials with patients with active and progressive PsA who have responded inadequately to previous DMARDs. All these included trials were relevant to the scope specified by NICE.

3.2. Intervention

The intervention specified in the manufacturer's decision problem is golimumab. Golimumab is licensed for the treatment of active and progressive PsA,[6] and the final scope defined by NICE specifies golimumab to be the intervention of interest. The manufacturer's evaluation of clinical efficacy and cost effectiveness adequately addressed this intervention specified by the NICE scope, although it did not specify the dose of golimumab. The MS presented data on therapy initiated with golimumab 100 mg, which is not reflective of the product licence. The current licensed dose of golimumab is 50 mg subcutaneously administered once a month.[6] For patients weighing more than 100 kg who do not achieve an adequate clinical response after three or four doses, an increase of the dose to 100 mg once a month may be considered. Continued therapy should be reconsidered in those who show no evidence of therapeutic benefit after receiving three to four additional doses of 100 mg.

3.3. Comparators

The decision problem again reflects exactly the NICE scope and specifies the following comparators to be of interest: a) alternative TNF-α inhibitors, and b) conventional management strategies for

active and progressive psoriatic arthritis that has responded inadequately to previous DMARD therapy excluding TNF-α inhibitors.

The key trial (GO-REVEAL[4, 5]) of golimumab efficacy employed placebo as the comparator. However, it should be noted that concomitant therapies (e.g. DMARDs, NSAIDs) were used in both treatment and placebo groups in this trial. The patients in the placebo group who had responded inadequately to previous DMARDs were maintained with conventional management strategies such as DMARDs and NSAIDs. As the majority of patients included in the placebo group received these concomitant therapies, the treatment effect observed in the placebo group was largely reflective of conventional management strategies for the treatment of active and progressive PsA.

The manufacturer's MTC analyses included trials evaluating three alternative TNF-α inhibitors (etanercept, infliximab and adalimumab). All the three anti-TNF agents are licensed for the treatment of active and progressive PsA patients who have inadequately responded to previous DMARDs. Without head-to-head comparisons between golimumab and the alternative TNF-α inhibitors, using MTC to indirectly estimate the relative efficacy between these anti-TNF agents was an appropriate way for the manufacturer to adequately address the comparators of alternative TNF-α inhibitors defined by the NICE scope.

3.4. Outcomes

In the statement of the decision problem, the manufacturer's submission addressed each of the following outcomes: pain and other symptoms; functional capacity; effect on concomitant skin condition; joint damage; disease progression (e.g. imaging); adverse effects of treatment; health-related quality of life. The primary outcomes in the only trial of golimumab (GO-REVEAL[4, 5]) were proportion of patients achieving an American College of Rheumatology (ACR) 20 response at week 14, and change from baseline in the PsA modified van der Heijde-Sharp (vdH-S) score at week 24 (these latter data were not presented in the MS but were provided after being requested by the ERG). Secondary outcomes in the direct efficacy comparisons were ACR 20 response at week 24, Psoriatic Arthritis Response Criteria (PsARC) response at week 14 and 24, and Psoriasis Area and Severity Index (PASI) 75 improvement at week 14 in patients with ≥ 3% body surface area (BSA) psoriasis at baseline. The physical functional status was measured by Health Assessment Questionnaire (HAQ) at week 24. The health-related quality of life was measured by the Short Form 36 Health Survey (SF-36) at week 14 (data were not presented in the MS). The safety outcome was the incidence of adverse events.

The primary outcomes in the MTC were PsARC response, change in HAQ score conditional on PsARC response to treatment, and change in PASI score for patients with BSA≥3% at baseline. The primary measure of cost-effectiveness was incremental cost per quality-adjusted life year (QALY) gained.

3.5. Timeframe

A timeframe was not specified in the NICE scope nor in the decision problem. Length of follow-up in the trial (GO-REVEAL) appeared to be adequate to observe the clinically meaningful changes in the efficacy outcomes (e.g. anti-inflammatory response, skin lesion response) of golimumab for patients with active and progressive PsA. These clinical efficacy endpoints were measured at 14 and 24 weeks, and then followed until 52 weeks. The data at 52 weeks follow-up were not presented in the MS.

In terms of the radiographic outcome, measuring radiographic changes for joint lesion response at 24 weeks is not considered adequate. Although this allows for the evaluation of the rapid onset of biologic therapies, clinical advice to the ERG suggests that observing meaningful changes in joint disease through radiographic measures usually requires one year follow-up.

3.6. Other relevant factors

N/A

4. CLINICAL EFFECTIVENESS

4.1. Critique of manufacturer's approach

4.1.1. Description of manufacturer's search strategy and comment on whether the search strategy was appropriate

The submission was checked against the Single Technology Appraisal (STA) specification for manufacturer/sponsor submission of evidence Update October 2009.

The manufacturer's submission described the search strategies used to identify relevant studies of golimumab, etanercept, infliximab, adalimumab and psoriatic arthritis, and full details of the search strategies used in each section were reported in the appendices or in the clarifications provided. Overall, most of the search strategies employed were fit for purpose, with the exception of those for sections 6.4 and 6.5 where no strategies were provided.

4.1.1.1. Search strategy for clinical evidence

Search strategy for section 5.1, clinical evidence

The submission gave detailed descriptions of the search strategies and met NICE requirements. It included the specific databases searched (MEDLINE, MEDLINE In-Process, EMBASE and The Cochrane Library); the date the searches were run; the complete strategies used and the results for each set. The date spans of searches were noted for EMBASE but not for MEDLINE or Medline in Process. The service providers were not noted for any of the searches. The searches were updates and expansions of those conducted by Rodgers et al.[7]

On the whole the strategies were followed faithfully, though all registry numbers for the drugs were omitted. The trials filter used for the MEDLINE searches differed from the one used in the searches performed by Rodgers et al.[7] The searches were run to retrieve material published since June 2006 for the drugs covered by that review. The strategies were expanded to include terms for golimumab. The drug registry numbers (CNTO-148 or CNTO148 or CNTO 148) were not included.

In addition, the submission states that ClinicalTrials.gov was searched. No search strategy was provided for this. Internet sources were searched for searched for information on adverse effects. The sources were not listed.

The search strategy for section 5.1, clinical evidence, was fit for purpose.

Search strategy for section 5.7, indirect/mixed treatment comparison

The submission indicated that the search strategy for this section was the same as that for the direct comparisons (5.1).The search strategy for section 5.1, indirect/ mixed treatment comparison, was fit for purpose.

Search strategy for section 5.8, non-RCT evidence

The submission indicated that the search strategy for this section was the same as that for the direct comparisons (5.1). Since this strategy included an RCT filter it is likely that much (though not all) non-RCT material would have been missed but the ERG is satisfied that this would not affect the outcome of the review.

Search strategy for section 5.9, adverse events

The submission gave detailed descriptions of the search strategies and met NICE requirements. It included the specific databases searched (MEDLINE, MEDLINE In-Process, EMBASE and The Cochrane Library); the date the searches were run; the complete strategies used. The date spans of searches were noted for EMBASE but not for MEDLINE or Medline in Process; the service providers were not noted for any of the searches. The searches were updates and expansions of those conducted by Rodgers et al.[7]

On the whole the strategies were followed faithfully, though all registry numbers for the drugs were omitted and this may possibly have led to trials being missed. In addition, the search term “treatment emergent$” was reported as “emergency treatment”. Potentially, this could lead to relevant material being missed.

In previous sections, golimumab was appropriately searched without date restrictions. For this section the new drug was searched using the same date restrictions as the other drugs, so only material since June 2009 will have been retrieved.

Internet sources were searched for information on adverse effects. The sources were not listed.

4.1.1.2. Search strategy for cost effectiveness

Search strategy for section 6.1, cost effectiveness

The databases searched for the cost effectiveness section included; MEDLINE, MEDLINE In-Process, EMBASE and NHS EED. EconLIT was the only database required by NICE that was not searched and this was due to lack of access.

The submission gave detailed descriptions of the search strategies and met NICE requirements. It included the specific databases searched; the service providers used; the dates when searches were conducted; the date spans of the searches; and the complete strategies used. These strategies were not based on Rodgers et al.[7] The strategies generally included a suitable combination of free text terms and subject headings, though the EMBASE strategy did not make use of available subject heading terms for the various drugs. In all cases the terms were combined appropriately.

The search strategy for section 6.1, cost-effectiveness, was appropriate.

Search strategy for section 6.4 – Measurement and valuation of health effects
Search strategy for section 6.5 – Resource identification, measurement and valuation.
For both of these sections, the report refers to Rodgers et al.[7] No search strategies were provided so it is impossible to say which, if any, searches were done.

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

Clinical efficacy: The evaluation of clinical efficacy included randomised controlled trials (RCTs) (including open-label extensions) evaluating golimumab for the treatment of active and progressive PsA patients with an inadequate response to previous standard therapy (at least one DMARD), and reporting relevant efficacy and quality of life outcomes (PsARC, PASI, HAQ, Dermatology Life Quality Index **(**DLQI), EQ-5D, and SF-36). Although the MS did not differentiate clearly between the intervention and comparators in the inclusion criteria, the eligible comparators appeared to be alternative anti-TNF agents (etanercept, infliximab or adalimumab) and palliative care including DAMRDs and NSAIDs (see Section 3.3 for the justification of this comparator).

The ERG noted that eligible outcomes did not include ACR 20 and change in radiographic outcomes. However, these two were primary outcomes in the GO-REVEAL trial and ACR 20 was the primary outcome for most trials being included in the MTC.

Safety evaluation: The inclusion/exclusion criteria in the MS for the evaluation of safety did not appear to correspond with the synthesis of safety data presented. The ERG requested further details in the points for clarification. In their response the manufacturer stated that 'The study selection criteria

refer to both efficacy and safety searches.' However, the synthesis of adverse effects in the MS comprised a summary of adverse effects from the GO-REVEAL trial[4, 5] and a table of summary information from six systematic reviews, this latter having been taken from the Rodger et al. report.[7] The list of references supplied by the manufacturer in response to the request by the ERG did not appear to comprise a list of included studies. It was clear that the synthesis of adverse effects data was not derived from studies included on the list.

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

Table 4.1: The included studies of the evaluation of clinical efficacy

Study	Study Design	Intervention and comparator
GO-REVEAL [4, 5]	RCT	golimumab versus placebo
Mease 2000 [8]	RCT	etanercept versus placebo
Mease 2004 [9]	RCT	etanercept versus placebo
IMPACT [10]	RCT	infliximab versus placebo
IMPACT 2[11]	RCT	infliximab versus placebo
ADEPT[12]	RCT	adalimumab versus placebo
Genovese 2007 [13]	RCT	adalimumab versus placebo

Direct trial evidence of the efficacy of golimumab: The MS included only a single phase III RCT (GO-REVEAL[4, 5]) comparing golimumab with placebo for treating active and progressive PsA patients who were symptomatic despite the use of current or previous DMARDs or NSAIDs (see Table 4.1).

Mixed treatment comparison: Apart from the GO-REVEAL trial[4, 5] (golimumab versus placebo), six additional trials were included in the MTC analyses (see Table 4.1). These included two RCTs (Mease 2000[8] and Mease 2004[9]) comparing etanercept with placebo, two RCTs (IMPACT[10] and IMPACT2) comparing infliximab with placebo, and two RCTs (ADEPT[12] and Genovese 2007[13]) comparing adalimumab with placebo.

Safety evaluation: As stated in section 4.1.2 the synthesis of adverse effects in the MS comprised a summary of adverse effects from the GO-REVEAL trial[4, 5] and a table of summary information from six systematic reviews, this latter having been taken from the Rodger et al. report.[7] The synthesis of adverse effects data did not appear to have been derived from the studies 'included' for safety evaluation.

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

Clinical efficacy: Based on the study selection flow diagram (p.28), there were 43 references being included in the evaluation of clinical efficacy and the list of these studies was supplied after being requested by the ERG. Based on these references (clarification response A12), it appears that all relevant studies have been included in the evaluation of direct trial evidence of the efficacy of golimumab and MTC analyses. The ERG noted, however, that one trial (the PRESTA trial[14]) was included on the reference list (clarification response A12) despite not meeting the inclusion criteria for the review of clinical efficacy nor being included in the evaluation of clinical efficacy in the MS.

Safety evaluation: Regarding the intervention being appraised only adverse events data from one trial (GO-REVEAL[4, 5]) were presented for the safety evaluation and only data up to week 24 were presented. The manufacturer could not comply with the ERG's request for the 52 week data because the analysis of these data has not yet been completed. The manufacturer failed to consider adverse event data of golimumab from controlled studies with other conditions such as rheumatoid arthritis and ankylosing spondylitis.

4.1.5. Description and critique of manufacturer's approach to validity assessment

Clinical efficacy: The MS used a modified published validity assessment checklist to assess the validity of the GO-REVEAL trial[4, 5] as well as other six trials being included in the MTC analyses (MS, p.66-67). The criteria used were appropriate, including randomisation method, concealment of allocation, blinding, drop-outs, similarity in terms of prognostic factors at baseline, measuring more outcomes than reported, and intention-to-treat analysis. For further discussion of study quality, see Section 4.2.

Safety evaluation: As above the same criteria employed, the MS appropriately used the modified published validity assessment checklist to assess the validity of the GO-REVEAL trial[4, 5] being included in the evaluation of safety.

4.1.6. Description and critique of manufacturer's outcome selection

Direct trial evidence of the efficacy of golimumab: The GO-REVEAL trial[4, 5] was powered to detect a significant difference between the placebo and combined golimumab groups in the primary efficacy outcome of ACR 20 response at week 14. The ACR 20 is generally accepted to be the minimal clinically important difference that indicates some response to a particular treatment; the

choice of this primary outcome appeared to be appropriate. The change from baseline in the PsA modified van der Heijde-Sharp (vdH-S) score at week 24 was also used in the GO-REVEAL trial[4, 5] as the primary outcome for radiographic assessment. However, this radiographic scoring method has not been validated in large PsA populations and measuring radiographic data on progression of joint disease at 24 weeks is a short time over which to identify a clinically significant effect of treatment: at least one-year follow-up is considered necessary.

The secondary outcomes included: ACR 20 response at week 24, PsARC response at week 14 and 24, PASI 75 improvement at week 14 in patients with ≥ 3% body surface area (BSA) psoriasis skin involvement at baseline, and HAQ score at week 24. The manufacturer also stated physical component summary score of SF-36 at week 14 as one secondary outcome, but the data were not presented in the MS.

Mixed treatment comparison: The outcomes of interest in the manufacturer's MTC analyses were PsARC response, change in HAQ score given PsARC response to treatment, and change in PASI in patients with BSA≥3% at baseline.

It should be noted that all included trials except for Mease 2000[8] used ACR 20 response as the primary outcome, and these trials were often powered to detect a significant difference for this outcome between treatment groups. However, the manufacturer's MTC analysis did not synthesise the outcome of ACR 20. In addition, PASI was chosen as the primary measure of skin disease response in the MTC. The ERG considered this outcome to be an appropriate measure for the skin disease response, as recommended by the British Association of Dermatologists (BAD) guidelines.[15]

Safety evaluation: The incidence of specific adverse events (including upper respiratory tract infection, nasopharyngitis, headache, diarrhoea, hypertension, infections and infestations) were reported for each arm at 16 and 24 weeks of the GO-REVEAL trial.[4, 5] Relative risks were provided to estimate the comparative risks between treatment and placebo groups in the GO-REVEAL trial,[4, 5] though 95% confidence intervals were not provided in the MS. However, the long-term follow-up adverse events data (52 weeks follow-up data) were not presented in the MS; the ERG requested these data, the manufacturer did not provide the data because these were not available at the time of request. Furthermore, the incidence of serious adverse events was not adequately reported in the MS. The incidence of malignancy was reported, but the incidence of serious infection and tuberculosis was not reported. These data were provided after being requested by the ERG.

4.1.7. Describe and critique the statistical approaches used

Direct trial evidence of the efficacy of golimumab and safety evaluation: The manufacturer provided tabulated data of the GO-REVEAL trial[4, 5] for the evaluation of efficacy and efficacy, but a narrative summary was not provided. In the MS the data were reported as being derived from an intention-to-treat (ITT) analysis, but following the ERG's request for clarification it became clear that this was not the case. A complete new table of results for the GO-REVEAL trial[4, 5] were submitted by the manufacturer in their clarification (see clarification response A3 and Appendix 1 of the report)

Mixed treatment comparison: In the absence of head-to-head comparisons between alternative anti-TNF agents, the manufacturer used the MTC approach to estimate the relative efficacy between the four anti-TNF agents of interest. A complete Bayesian MTC analysis incorporating the entire available network of direct and indirect evidence is a robust approach to estimate the relative efficacy between the alternative treatments.[16] All the four anti-TNF agents being evaluated in the included trials had a common comparator of placebo, allowing the network between golimumab, etanercept, infliximab and adalimumab to be established. Thus, the MTC network in the MS was appropriately constructed.

In terms of the MTC analysis, the absolute probabilities or mean changes from baseline for each biologic for joint disease, functional status and skin disease outcomes were appropriately presented in the MS (p.86). The MS combined evidence using Bayesian evidence synthesis methods. A brief outline of the methods used has been presented in Table 4.2 and are discussed below. The outcomes synthesised were PsARC, HAQ conditional on PsARC response and absolute PASI change from baseline. The manufacturer included seven RCTs in their synthesis and chose to synthesise using the latest available endpoint. A number of assumptions were required to facilitate modelling, in brief these included:

1. The change in HAQ from baseline was modelled conditional on PsARC response

2. PASI is modelled as an aggregate across patients with or without a PsARC response

3. The model uses absolute changes in HAQ and PASI. Where trials only report the relative change in PASI (e. g. average 54% improvement) or "response criteria" such as PASI 50, PASI 75, etc., the absolute changes have to be inferred.

4. PASI is only modelled for the subset of patients with initial BSA $\geq$ 3%.

5. All patients with BSA >3% are assumed to have identical PASI baseline values equal to the mean PASI baseline score reported for this subgroup in the trial

6. If the trial does not report the baseline PASI for a group, it is assumed to be equal to the average score reported in the other trials

7. The PASI change is not correlated with the PASI baseline score

8. The PASI change and HAQ change are not correlated in the BSA > 3% group

9. The HAQ change is conditional on PsARC response

10. Where trials do not report the HAQ outcomes separately by PsARC response group, it has been assumed that the HAQ change for the PsARC non-responders is equivalent to the average HAQ change in non-responders seen in other trials, and the HAQ change for the PsARC responders is inferred to match the reported mean HAQ change

11. The HAQ change from baseline to the last RCT controlled data point up to week 24 is the main outcome of interest and is the main determinant of the outcomes of the economic model

12. The HAQ change is not correlated with baseline HAQ score

13. The HAQ change is assumed identical for the subgroups with or without BSA≥ 3% at baseline

For PsARC response the MS model used a fixed effect meta-analysis, which incorporated 12 or 14 week outcome data. HAQ conditional on a PsARC response was modelled using two linked meta-analysis which estimated the probability of response and then the mean reduction in HAQ score conditional on that response. The use of the latest available outcome data meant that some 24 week HAQ outcomes were incorporated into the model.

To enable them to include all seven trials the MS assumed that for the one trial where the data were not stratified by responder/non-responder[8] that the HAQ change for the PsARC non-responder was equivalent to the average HAQ change in the non-responders as seen in other trials, and that the HAQ change for the PsARC responders could be inferred to match the reported mean HAQ change.

The MS elected to use absolute PASI change as their main outcome, on the basis that this was the most appropriate outcome for the economic modelling. No synthesis was undertaken on the ACR outcomes. The annotated WinBUGS code, assumptions and data were all presented. However, the manufacturer did not present the estimated relative risks (RRs) and weighted mean differences of each outcome, relative to placebo, on the basis of results of MTC analyses. Based on the results provided by manufacturer after the ERG's request (clarification response A18), there was consistency between the results from the MTC and standard meta-analyses for the outcomes of PsARC response and HAQ change unconditional on response.

Table 4.2: Overview of evidence synthesis model

MS evidence synthesis model overview	
Interventions	Etanercept, Infliximab, Adalimumab, Golimumab
Studies used in the analysis	IMPACT,[10] IMPACT 2,[11]Mease 2000,[8] Mease 2004,[9] ADEPT,[12] Genovese 2007,[13] Woolacott 2006,[17] GO-REVEAL[4, 5]
Outcomes of interest	
PsARC	12 or 14 weeks data from all trials
HAQ	HAQ at 12 and 24 weeks for Adalimumab/ 14 or 16 weeks for Infliximab/ 12 weeks for Etanercept and 14 weeks for Golimumab. (Conditional on PsARC response).
PASI 50/75/90	Estimated absolute PASI change from baseline. Incorporated data from week 24 for Adalimumab/ week 14 or 16 for Infliximab/ week 24 for Etanercept and week 14 for Golimumab
ACR 20/50/70	Not estimated.
Model	Two linked meta-analysis: estimating the change in HAQ from baseline conditional on PsARC response. Absolute change in PASI was modelled.
Results Reported	Incremental HAQ change given PsARC response in treatment, Incremental HAQ change given PsARC non-response in treatment, Incremental HAQ change given PsARC response in placebo, Incremental HAQ change given PsARC non-response in placebo.

The skin disease response from the MTC analyses in the MS depended on the assumption that the change in HAQ and change in PASI were independent. In order to establish whether this assumption was valid, the ERG requested further data on the number of patients with psoriasis at baseline who achieved PsARC response with and without achieving PASI 75 response in the form of 2x2 tables for each biologic. The ERG also asked the manufacturer to perform the statistical test to confirm that the differences in mean PASI change were the same in PsARC responders and non responders. The data from the GO-REVEAL trial[4, 5]further provided by the manufacturer (clarification response A15) generally supported the assumption that PASI 75 response was independent of PsARC response in patients on golimumab. However, this does not provide information on whether psoriasis and arthritis responses might be independent for other anti-TNF agents.

The ERG further noted that the manufacturer used the last randomised endpoint before week 24 to measure the change in HAQ and PASI from baseline, based on the assumption that changes in HAQ and PASI were same at the follow-up of 12 weeks and 24 weeks. In order to support this assumption, the ERG requested the manufacturer to re-estimate the meta-analysis using the data from the data point closest to 12 weeks. The re-estimated analyses using the data from the data point closest to 12 weeks (clarification response A17) did not materially alter the results. This supported the assumption

that changes in HAQ and PASI were the same at the follow-up of 12 weeks and 24 weeks (see Appendix 2).

Based on the above evidence presented, the ERG considered the statistical approaches used in the MTC analyses in the MS to be reliable.

4.1.8. Summary Statement

Although the manufacturer's search strategies were appropriate and likely to have identified all the evidence relevant to the decision problem, the ERG had concerns about how the studies were selected in the MS. In terms of the evaluation of clinical efficacy, it appears that all relevant studies have been included. However, there was a lack of clarity regarding the study selection for the safety evaluation, which was not based on the 'included' studies but on the Rodger et al. report.[7]

In the MS, appropriate criteria were used to assess the study validity. The outcome selection in both direct efficacy evidence of golimumab and MTC analyses was generally considered appropriate. However, there was no detailed narrative summary for the evaluation of efficacy and adverse events, though the tabulated data of the GO-REVEAL trial[4, 5] were provided. On the basis of the further evidence provided by the manufacturer, the ERG considered the statistical approaches used in the MTC analysis to be reliable.

4.2. Summary of submitted evidence

4.2.1. Summary of results

Based on the limited available evidence, the manufacturer concludes that golimumab is associated with greater efficacy than placebo for the treatment of active and progressive PsA in patients who have responded inadequately to previous DMARDs. This was based on the data from the only trial of golimumab (GO-REVEAL[4, 5]), which showed a statistically significant improvement on patients' joint disease as measured by ACR 20 and PsARC and skin disease as measured by PASI 75 at both 14 and 24 weeks, compared with placebo. There was also a significant improvement of patients' functional status (HAQ) at 24 weeks. The data of open-label extension from the GO-REVEAL trial[4, 5] showed that these beneficial effects were maintained at 52 and 104 weeks.

Based on the results of MTC analyses, the manufacturer states that the efficacy of golimumab is comparable to other TNF-α inhibitors including infliximab, adalimumab and etanercept.
In terms of the safety of golimumab, the manufacturer concludes that golimumab is a safe treatment option similar to other TNF-α inhibitors on the basis of the safety evaluation results from the GO-REVEAL trial.[4, 5]

4.2.2. Critique of submitted evidence syntheses

4.2.2.1. Direct trial evidence of the efficacy of golimumab (GO-REVEAL)

Table 4.3 summarises the characteristics of the Go-REVEAL trial. In the GO-REVEAL trial,[4, 5] randomisation and concealment of allocation were adequate. Blinding of patients, investigators and assessors was also adequately performed, thereby minimising the potential bias during the data collection and analyses. This trial was adequately powered to detect a significant difference of the primary outcome between the treatment and placebo groups.

In the GO-REVEAL trial,[4, 5] at 16 weeks, for those patients with less than 10% improvement from baseline in both the swollen and tender joint counts in the lower dose treatment arm, the dose was increased from golimumab 50 mg to 100 mg. The dosing regimen (including the dose adjustment) for the golimumab 50 mg group was generally representative of the routine clinical practice,[6] except it did not meet the criterion specified by the summary of product characteristics (SPC)[6] that increasing the dose from 50 mg to 100 mg should be specifically considered on those patients weighing more than 100 kg. However, it should be noted that the dosing regimen used in the golimumab 100 mg group was not representative of the current clinical practice, because the dose of golimumab 100 mg is not directly licensed as a starting dose for the treatment of active and progressive PsA.[6]

Table 4.3 Characteristics of the GO-REVEAL Trial (Adapted from the manufacturer's submission)

Study details and design	Participant details	Intervention/outcome/analyses details
GO-REVEAL trial [4, 5] **Study location:** 58 centres: 36 in North America (18 in the US and 18 in Canada), 22 in Europe (5 in Belgium, 10 in Poland, 3 in Spain, and 4 in the UK) **Funding** Schering-Plough **Study design** Multicenter, randomized, double-blind, placebo-controlled study was designed to assess the efficacy, safety, and clinical pharmacology of golimumab 50 mg or 100 mg administered subcutaneously q4 weeks in adult subjects with active PsA. Stage 1: Double-blind RCT (up to week 24) Stage 2: Open-label follow-up (from week 24 to week 52) **Method of randomisation** Subjects were to be randomized in a 1:1.3:1.3 ratio to 1 of 3 treatment groups: placebo, golimumab 50 mg& golimumab 100 mg. In order to ensure relatively even treatment balance within sites, within baseline MTX usage (yes/no), and within the study overall, subject allocation to a treatment group was performed using an adaptive stratified randomization design.	**Participants** Patients with a diagnosis of psoriatic arthritis (PsA) for at least 6 months prior to first study agent administration who had active PsA despite current or previous disease-modifying anti-rheumatic drug (DMARD) or nonsteroidal anti-inflammatory drug (NSAID) therapy, and who had not previously been treated with anti-tumour necrosis factor (TNF) α therapy. **Age** *Mean age (Standard deviation (SD))* Golimumab 50 mg: 45.7 (10.7) Golimumab 100 mg: 48.2 (10.9) Placebo: 47.0 (10.6) Gender, Male (%) Golimumab 50 mg: 61% Golimumab 100 mg: 59% Placebo: 61% Psoriatic arthritis history *Duration of psoriatic arthritis, Mean years (SD)* Golimumab 50 mg: 7.2 (6.8) Golimumab 100 mg: 7.7 (7.8) Placebo: 7.6 (7.9) Psoriasis history *Duration of psoriasis, Mean years (SD)* Golimumab 50 mg: 17.7 (11.9) Golimumab 100 mg: 18.4 (12.7) Placebo: 19.0 (12.9) Psoriasis Evaluation *Patients with ≥ 3% body surface area affected with psoriasis, n (%)* Golimumab 50 mg: 109 (75%) Golimumab 100 mg: 108 (74%) Placebo: 79 (70%) **Concomitant therapy during trial**	**Intervention 1: 50 mg golimumab** No. randomised: 146 No. completed at week 16: 139 No. completed at week 24: 137 **Intervention 2: 100 mg golimumab** No. randomised: 146 No. completed at week 16: 144 No. completed at week 24: 142 **Comparator: placebo** No. randomised: 113 No. completed at week 16: 103 No. completed at week 24: 101 **Primary Outcomes** 1) Proportion of subjects achieving an American College of Rheumatology (ACR) 20 response at week 14 2) Change from baseline in the PsA modified van der Heijde-Sharp (vdH-S) score at week 24 **Secondary outcomes:** 1) ACR 20 response at week 24 2) Psoriasis Area and Severity Index (PASI) 75 improvement at Week 14 in a subset of subjects with ≥ 3% body surface area (BSA) psoriasis skin involvement at baseline 3) HAQ score at week 24 4) Physical component summary score of SF-36 at week 14 5) Additional secondary endpoints : safety, and tolerability of golimumab & the pharmacokinetics/pharmacodynamics of golimumab dose groups **Sample size calculation** The planned sample size (n = 396; 110 in the placebo group and 286 in the combined golimumab group) provided > 98% power to detect a significant difference (α= 0.05) between the placebo and combined golimumab groups in the primary efficacy end point, assuming equal proportions of patients in each group received MTX at baseline and the following proportions of patients achieved an ACR 20 response at week 14: 15% of patients receiving placebo, 25% of patients receiving placebo plus MTX, 42% of patients receiving both golimumab doses combined, and 42% of patients receiving both golimumab doses

	Corticosteroids (%): Golimumab 50 mg: 13% Golimumab 100 mg: 18% Placebo: 17% **NSAIDS (%):** Golimumab 50 mg: 75% Golimumab 100 mg: 75% Placebo: 78% **MTX (%):** Golimumab 50 mg: 49% Golimumab 100 mg: 47% Placebo: 48%	combined plus MTX. Statistical analyses Treatment group differences were assessed with a 2-side Cochran-Mantel-Haenszel test for discrete variables or 2-side analysis of variance on van der Waerden normal scores for continuous parameters. All analyses included treatment and patients' use of MTX at baseline as factors and were performed at a 0.05 level of significance. **ITT analysis** The analyses were performed on an intention-to-treat basis

Efficacy after 14 weeks treatment

The results of the GO-REVEAL trial[4, 5] are summarised in Table 4.4. These results are not those presented in the MS but newly analysed ITT results submitted by the manufacturer in their clarification letter. The ERG had noted that the intention-to-treat analysis was not adequately applied in the efficacy analysis in the MS and requested the clarification of such analyses presented in the MS. In the new analyses provided by the manufacturer, the intention-to-treat analysis was appropriately performed for most efficacy outcomes. Given the evidence that treatment effect was maintained during the follow-up in the trial, using the last observation carried forward method to impute missing data was considered appropriate. However, it should be noted that analyses of PASI 50 and PASI 90 at 14 weeks and all the PASI outcomes at 24 weeks were not performed on the basis of intention-to-treat analysis. Therefore, these analyses may have potentially compromised the reliability of the results in terms of skin disease outcomes.

Table 4.4 Efficacy data of golimumab in the GO-REVEAL trial

Duration	Outcomes	Golimumab 50 mg	Golimumab 100 mg	Placebo	Golimumab 50 mg: RR or mean difference (95% CI)	Golimumab 100 mg: RR or mean difference (95% CI)
14 weeks	PsARC	107/146 (73.3%)	105/146 (71.9%)	24/113 (21.2%)	3.451 (2.49 - 4.87)	3.386 (2.43 - 4.80)
	ACR 20	74/146 (50.7%)	66/146 (45.2%)	10/113 (8.8%)	5.727 (3.24 - 10.56)	5.108 (2.86 - 9.48)
	ACR 50	44/146 (30.1%)	41/146 (28.1%)	2/113 (1.8%)	17.027 (4.81 – 63.32)	15.866 (4.47 – 59.11)
	ACR 70	18/146 (12.3%)	25/146 (17.1%)	1/113 (0.9%)	13.932 (2.46 – 81.82)	19.349 (3.48 – 112.44)
	HAQ change from baseline, mean (SD)	n/a	n/a	n/a	-	-
	PASI 50*	63/106 (59.4%)	83/107 (77.6%)	7/73 (9.6%)	6.198 (3.22 – 12.7)	8.089 (4.38 – 16.04)
	PASI 75*	44/109 (40.4%)	63/108 (58.3%)	2/79 (2.5%)	15.945 (4.62 – 59.11)	23.042 (6.85 – 84.59)
	PASI 90*	22/106 (20.8%)	26/107 (24.3%)	0/73 (0.0%)	Inf (4.21 – Inf)	Inf (4.95 – Inf)
24 weeks	PsARC	102/146 (69.9%)	124/146 (84.9%)	33/113 (29.2%)	2.392 (1.81 – 3.20)	2.908 (2.28 – 3.68)
	ACR 20	76/146 (52.1%)	89/146 (61.0%)	14/113 (12.4%)	4.202 (2.60 – 7.03)	4.920 (3.09 – 8.13)
	ACR50	47/146 (32.2%)	55/146 (37.7%)	4/113 (3.5%)	9.094 (3.62 – 23.94)	10.642 (4.27 – 27.85)
	ACR70	27/146 (18.5%)	31/146 (21.2%)	1/113 (0.9%)	20.897 (3.77 – 121.19)	23.993 (4.35 – 138.68)
	HAQ change from baseline, mean (SD)	0.33 ± 0.55 p < 0.001	0.39 ± 0.50 p < 0.001	- 0.01 ± 0.49	-	-
	PASI 50*	77/102 (75.5%)	87/106 (82.1%)	6/73 (8.2%)	9.185 (4.69 – 19.45)	9.986 (5.21 – 20.76)
	PASI 75*	57/102 (55.9%)	70/106 (66.0%)	1/73 (1.4%)	40.794 (7.86 – 232.88)	48.208 (9.44 – 274.39)
	PASI 90*	33/102 (32.4%)	34/106 (32.1%)	0/73 (0.0%)	Inf (6.65 – Inf)	Inf (6.59 – Inf)
	vdH-S score change from baseline, mean (SD)	-0.16 ± 1.31 p=0.011	-0.02 ± 1.32 p=0.086	0.27 ± 1.26	-	-

*reported for patients with at least 3% BSA psoriasis. n/a: not available

The ERG further noted that there was an imbalance in patients' characteristics at baseline between the treatment and placebo group. A lack of balance at baseline between the treatment groups for variables of significant predictors of treatment response can influence the outcomes.[18] [redacted].

Polyarticular arthritis is a significant predictor of the treatment response and patients with polyarticular arthritis are often associated with a poor response. However, it should be also noted that there was a noticeable balance in the mean swollen and tender joint counts at baseline between the treatment and placebo group. Therefore, overall, given that joint disease response depends in part on the number of joints involved at outset, there was baseline comparability of joint disease severity between the treatment and placebo groups.

There were [redacted] between the treatment and placebo group in the GO-REVEAL trial at 14 weeks: placebo [redacted], golimumab 50 mg [redacted] and golimumab 100 mg [redacted] Although there was little concern about [redacted][19] [redacted].

The data from the GO-REVEAL trial[4, 5] demonstrated a significant improvement in PsA patients for all joint disease outcomes at 14 weeks (see Table 4.4). Compared with placebo, golimumab 50 mg was associated with a statistically significant reduction in joint symptoms assessed by ACR 20 (RR 5.73, 95% CI 3.24 to 10.56) and PsARC (RR 3.45, 95% CI 2.49 to 4.87) and on skin disease in terms of PASI 50 (RR 6.20, 95% CI 3.22 to 12.7) and PASI 75 (RR 15.95, 95% CI 4.62 to 59.11).

At 14 weeks, golimumab 100 mg also achieved a statistically significant improvement on both joint and skin diseases in terms of ACR 20, PsARC, PASI 50 and PASI 75.

The manufacturer did not present the mean HAQ score from baseline at 14 week in their revised data table (clarification response A3), though these values were originally presented in the MS. No justification for this exclusion was provided.

Efficacy after 24 weeks treatment

The GO-REVEAL trial[4, 5] maintained randomisation for 24 weeks but the interpretation of the 24 week results was complicated by the trial design, which permitted upward titration of the dose at week

16 such that patients in the placebo group could switch to 50 mg golimumab and patients in the 50 mg golimumab group could have their dose increased to 100 mg. This dose adjustment is in accordance with the SPC for golimumab[6] (see Section 4.2.2.1).

The data for all measures of the joint and skin disease were similar to those observed at the earlier 14 weeks follow-up, suggesting that the benefits of golimumab were maintained up to 24 weeks of treatment (see Table 4.4). However, the efficacy results at 24 weeks were confounded by the analyses, which did not appropriately handle crossing-over data: the analyses appeared to ignore the treatment contamination due to dose/treatment changes. In the trial, 50% of patients in the placebo group crossed over to golimumab 50 mg treatment and 20% of patients in the golimumab 50 mg group crossed over to golimumab 100 mg treatment. Whilst the analyses at 24 weeks involved all the intention-to-treat data (i.e. it included these crossing-over data) from the randomisation, it appears that these intention-to-treat analyses failed to adjust for the treatment contamination due to patients' crossing-over at week 16. Therefore, the failure to adjust this treatment contamination in analyses may have threatened the internal validity of trial results for all the outcomes at 24 weeks.

It should be further noted that the absolute values for the golimumab 50 mg arm at 24 weeks were reflective of the use of this biologic in clinical practice, as the dosing regimen (including the dose adjustment) was generally in line with the product licence.[6] However, the relative effects (golimumab versus placebo) may have been confounded by the treatment contamination due to the failure to adjust for crossing-over data.

At 24 weeks, there were statistically significant improvement in mean HAQ score from baseline of 0.33 for golimumab 50 mg ($p<0.001$), and 0.39 for golimumab 100 mg ($p<0.001$). These indicated beneficial effect of golimumab therapy on patients' functional status.

Short-term radiographic measures of vdH-S score indicated that golimumab 50 mg can slow disease progression in the short term at 24 weeks with a significant reduction from baseline of 0.16 ($p=0.01$), though this significant impact was not observed in the golimumab 100 mg group ($p=0.09$). Due to a lack of follow-up radiographic data, it was unclear whether these effects persisted in the longer term.

Additionally, an increased rate of drop-outs was observed at 24 weeks: placebo 10.6%, golimumab 50 mg 6.2% and golimumab 100 mg 2.7%.

Longer term efficacy

[redacted]

Summary of efficacy of golimumab

Overall, the analyses on the 14 week data in the GO-REVEAL trial[4, 5] were considered reliable. Longer term data suggested that the treatment effects were maintained in those who continued therapy or had an increase in the dose of golimumab 50 mg to 100 mg. However, there was lack of the robustness for the analyses on the 24 week data in terms of the beneficial effect of golimumab therapy relative to placebo because the analyses failed to adjust the treatment contamination due to patients' crossing-over.

4.2.2.2. Relative efficacy between anti-TNF agents

Trial design and quality of evidence

In the absence of head-to-head trials of the four anti-TNF agents, the relative efficacy evidence for each biologic agent was investigated by means of a MTC. The analyses used all the available data comprising seven RCTs: one trial of golimumab and two trials for each agent of etanercept, infliximab and adalimumab.

In the MTC, the included trials other than GO-REVEAL (see 4.2.2.1 details of quality of the GO-REVEAL trial[4, 5]) were generally of good quality. Randomisation, blinding, concealment of allocation and intention-to-treat analyses were adequate in most trials. The validity of the MTC of meta-analysis is built on the assumptions that no important differences exist between trials in terms of baseline characteristics such as disease severity.[20] The population characteristics of all seven trials are summarised and compared in Table 4.5. The trials are similar in terms of

patients' joint disease severity at baseline (e.g. mean tender joint count, mean swollen joint count); this is important to ensure the validity of the PsARC outcome in the MTC.

Despite some differences in the mean HAQ score at baseline between the included trials, there was a high variability of these HAQ values (high standard deviation) and, thus, it is very likely that differences in mean HAQ scores were not significant. Although there was a concern about the correlation between baseline HAQ scores and absolute HAQ changes in these PsA patients, given such a high variability of these HAQ values, the ERG considered the exchangeability of mean HAQ scores across the included trials in the MTC analysis to be acceptable.

Whilst there were slight differences in mean PASI score at baseline for those patients evaluable for psoriasis endpoints between included trials, these mean PASI scores were also associated with a large variability. Again, this did not suggest significant heterogeneity in terms of patients' skin disease severity at baseline between included trials in the MTC.

There were, however, differences in terms of proportions of patients evaluable for psoriasis endpoints at baseline between included trials. Where reported, the IMPACT 2 trial,[11] GO-REVEAL trial,[4, 5] and Mease 2004[9] had [redacted]%, [redacted]%, 63% of patients evaluable for psoriasis endpoints at baseline, respectively. In contrast, the IMPACT trial[10] and ADEPT trial[12] had only [redacted]% and 45% of patients evaluable for psoriasis endpoints. Thus, there might be potential interactions between different patient samples and the treatment effect when estimating the relative efficacy in terms of skin disease responses.

Despite some limitations mentioned above, overall, in the MTC analyses in the MS, the degree of clinical heterogeneity between the included trials in terms of joint and skin disease severity and functional status was reasonable. Therefore, the assumption of exchangeability between the trials for the purposes of the MTC was acceptable.

Table 4.5 Summary of trial population characteristics

	Golimumab			Etanercept				Infliximab				Adalimumab			
	GO-REVEAL[4, 5]			Mease 2000[8]		Mease 2004[9]		IMPACT[10]		IMPACT 2[11]		ADEPT[12]		Genovese 2007[13]	
	[redacted]			Etanercept (n=30)	Placebo (n=30)	Etanercept (n=101)	Placebo (n=104)	[redacted]		[redacted]		Adalimumab (n=151)	Placebo (n=162)	Adalimumab (n=51)	Placebo (n=49)
Age in years Mean (SD)	[redacted]			46.0 (30.0-70.0)†	43.5 (24.0-63.0)†	47.6 (18-76)†	47.3 (21-73)†	[redacted]		[redacted]		48.6 (12.5)	49.2 (11.1)	50.4 (11.1)	47.7 (11.3)
Male (%)	[redacted]			53	60	57	45	[redacted]		[redacted]		56	55	57	51
Duration of PsA (years) Mean (SD)	[redacted]			9.0 (1-31)†	9.5 (1-30)†	9.0 (-)†	9.2 (-)†	[redacted]		[redacted]		9.8 (8.3)	9.2 (8.7)	7.5 (7.0)	7.2 (7.0)
Duration of psoriasis (years) Mean (SD)	[redacted]			19.0 (4-53)†	17.5 (2-43)†	18.3 (-)†	19.7 (-)†	[redacted]		[redacted]		17.2 (12.0)	17.1 (12.6)	18.0 (13.2)	13.8 (10.7)
Number of prior DMARDS Mean (SD)	[redacted]			1.5	2.0	1.6	1.7	[redacted]		[redacted]		1.5	1.5	1.7	2.1
Proportion of patients with numbers of previous DMARDs*	[redacted]			-	-	27% = 0 40% = 1 20% = 2	21%=0, 50% =1 19% =2	[redacted]		[redacted]		-	-	-	
Concomitant therapies during study (%)	[redacted]							[redacted]		[redacted]					
Corticosteroids				20	40	19	15					-	-	-	-
NSAIDs				67	77	88	83					-	-	73	86
Methotrexate				47	47	45	49					51	50	47	47
Hydroxycloroquine				-	-	-	-					-	-	16	16
Sulfasalazine				-	-	-	-					-	-	8	14
Leflunomide				-	-	-	-					-	-	6	4
Other DMARD				-	-	-	-					-	-	2	6

Type of PsA (%) DIP joints in hand and feet Arthritis mutilans Polyarticular arthritis Asymmetric peripheral arthritis Ankylosing arthritis Spondylitis with peripheral arthritis	[redacted]	[redacted]	[redacted]	- - - - - -	- - - - - -	51 1 86 41 3 -	50 2 83 38 4 -	[redacted]	[redacted]	[redacted]	[redacted]	- 1 64 25 1 -	- 0 70 25 0 -	- 0 82 10 2 -	- 0 84 14 2 -
Tender Joint Count Mean (SD)	[redacted]	[redacted]	[redacted]	22.5 (11, 32)*	19.0 (10, 39)*	20.4 (-)*	22.1 (-)*	[redacted]	[redacted]	[redacted]	[redacted]	23.9 (17.3)	25.8 (18.0)	25.3 (18.3)	29.3 (18.1)
Swollen Joint Count Mean (SD)	[redacted]	[redacted]	[redacted]	14.0 (8, 23)*	14.7 (7, 24)*	15.9 (-)*	15.3 (-)*	[redacted]	[redacted]	[redacted]	[redacted]	14.3 (12.2)	14.3 (11.1)	18.2 (10.9)	18.4 (12.1)
HAQ (0-3) Mean (SD)	[redacted]	[redacted]	[redacted]	1.3 (0.9, 1.6)*	1.2 (0.8, 1.6)*	1.1 (-)*	1.1 (-)*	[redacted]	[redacted]	[redacted]	[redacted]	1.0 (0.6)	1.0 (0.7)	0.9 (0.5)	1.0 (0.7)
Number (%) of patients evaluable for PASI at baseline	[redacted]	[redacted]	[redacted]	19 (63%)♦	19 (63%)♦	66 (65%)♦	62 (60%)♦	[redacted]	[redacted]	[redacted]	[redacted]	70 (46%)♦	70 (43%)♦	-	-
PASI (0-72) at baseline among patients evaluable for PASI Mean (SD)	[redacted]	[redacted]	[redacted]	10.1 (2.3-30.0) †	6.0 (1.5-17.7) †	9.6 (-)	11.1(-)	[redacted]	[redacted]	[redacted]	[redacted]	7.4 (6.0)	8.3 (7.2)	-	-

†median (range)
* median (25th, 75th percentile)
♦ Patients with ≥3% BSA psoriasis at baseline
‡Patients with a baseline PASI score ≥2.5

Results of evidence synthesis

Table 4.6 presents the results of the manufacturer's MTC analyses. The data used from the GO-REVEAL trial[4, 5] in the MTC were 14 week data. As discussed in the Section of 4.2.2.1, the analyses on 14 week data from this trial were considered robust. This also ensured the robustness of the results from the MTC analyses.

Table 4.6: Results of MTC analyses in the MS

Outcomes	Placebo	Infliximab		Etanercept		Adalimumab		Golimumab	
	Manufacturer result	Manufacturer result	Rank	Manufacturer result	Rank	Manufacturer result	Rank	MS	Rank
PsARC response Mean (SD), 95% CrI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HAQ change from baseline, in PsARC responders Mean (SD), 95% CrI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HAQ change from baseline, in PsARC non-responders Mean (SD), 95% CrI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PASI change from baseline, in patients ≥3% BSA psoriasis at baseline Mean (SD), 95% CrI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

PsARC response

The results of PsARC response in the MS (p.86) were in the form of probability of response. These results for all four anti-TNF agents were statistically significant with moderate to large effect sizes (see Table 4.6). Among all the four anti-TNF agents, infliximab achieved a highest probability of PsARC response on the joint disease at 12 or 14 weeks ([redacted]). Golimumab achieved a third highest probability of PsARC response ([redacted]).

Changes in HAQ

The results of HAQ conditional on PsARC response in the MS (p.86) were presented as absolute changes in HAQ from baseline. For the HAQ data in the GO-REVEAL trial,[4, 5] it was unclear where these data came from, as the manufacturer failed to present these data in their revised data table (point clarification A3).

For patients who achieved a PsARC response, statistically significant reductions in mean HAQ scores at 12 -14 weeks were observed in all four anti-TNF agents (see Table 4.6). Etanercept was associated with the largest estimate of improvement in patients' functional status in terms of mean HAQ change from baseline ([redacted]), and golimumab was associated with the lowest estimate of improvement in patients' functional status ([redacted]), though the credible interval overlapped those of the other anti-TNF agents. The mean HAQ reductions for all four anti-TNF agents [redacted][1]

For patients who did not achieve a PsARC response, statistically significant reductions in mean HAQ at 12-14 weeks were only observed for etanercept ([redacted]) and infliximab ([redacted]). The mean HAQ reductions of all four anti-TNF agents were [redacted][1]

PASI

The results of skin disease response in the MS (p.86) were presented as absolute changes in PASI from baseline. In patients evaluable for psoriasis endpoints, statistically significant reductions in PASI from baseline at 14 or 24 weeks were observed in all four anti-TNF agents (see Table 4.6). Infliximab achieved the largest estimate of improvement on skin disease in terms of mean PASI reduction from baseline ([redacted]). Golimumab achieved a lower estimate of improvement on skin disease ([redacted]). The credible intervals for all four anti-TNF agents overlapped between each other, and there were substantial uncertainties for these estimates due to a small sample size of patients evaluable for psoriasis

Generalisability

Despite most patients in the included trials being under licensed conditions, it was noteworthy that the majority of patients had previously received at least one DMARD, and no trial specified the failure to

respond to at least two DMARDs (patients whom the current BSR guidelines and NICE guidance for etanercept, infliximab and adalimumab consider eligible for the biologic treatment) as a recruitment criterion. Thus, the trial participants in the MTC analysis were likely to have less severe disease of PsA compared to those patients receiving biologic treatment in routine practice. Given this consideration, trial participants were not precisely representative of the active and progressive PsA population recommended for anti-TNF agents by the current guidelines. Therefore, it remained unclear that the beneficial effects observed in these trial participants in the MTC analysis were similar in those treated in routine clinical practice.

4.2.2.3. Safety evaluation

The evidence in the MS for the safety of golimumab was exclusively drawn from one RCT (GO-REVEAL[4, 5]). The study quality of GO-REVEAL trial[4, 5] is discussed in Section 4.2.2.1 of this report. The safety data from this trial were analysed on the basis of treated patients, i.e. patients who received at least one study agent administration. This analysis approach on the basis of patients who received the treatment was considered to be rigorous.

In the GO-REVEAL trial,[4, 5] the most frequently reported adverse events associated with golimumab therapy included infections and infestations, upper respiratory tract infection and nasopharyngitis. Compared with placebo, an increased risk of upper respiratory tract infection at 16 weeks was observed in the group of golimumab 50 mg (RR 2.14) and golimumab 100 mg (RR 1.57). An increased risk of nasopharyngitis was also observed in both golimumab 50 mg and 100 mg groups at week 16, with RR 1.17 and RR 2.54, respectively. There was also an increased risk of hypertension (RR 5.85) in each golimumab group at week 16. The 24 week data showed similar results for the outcomes of upper respiratory tract infection and nasopharyngitis; however, it should be noted that these results were confounded by the treatment contamination as the analyses failed to adjust this issue due to crossing-over. The manufacturer failed to provide 95% confidence intervals of these data after the ERG's request; this made it difficult to assess the significance level of these comparative safety data of golimumab relative to placebo. Only limited information on serious adverse events was included in the MS: malignancy was reported for three patients receiving golimumab 100 mg through 24 weeks, including two cases of basal cell malignancies and one case of prostate cancer. After the 24 weeks (open label extension), there was one treatment-related death due to small cell lung cancer and one case of liver histoplasmosis.

The incidence of serious adverse events was not adequately reported in the MS but the manufacturer provided these data after the ERG's request. Additional adverse effect data provided by the manufacturer in their clarification letter is given in Appendix 3 and summarised in Table 4.7.

Table 4.7: Summary of additional adverse effects data submitted in clarification letter (serious adverse effects, serious infections, tuberculosis, and adverse effects leading to discontinuation)

Adverse event	**Placebo**	**Golimumab 50 mg**	**RR (95% CI) golimumab versus placebo**
Week 16	**n = 113**	**n = 146**	
Patients with ≥ 1 AE	63/113 (55.8%)	85/146 (58.2%)	1.044 (0.85-1.30)
Patients with ≥ 1 serious AE	6/113 (5.3%)	3/146 (2.1%)	0.387 (0.11-1.39)
Patients with ≥ 1 serious infections	3/113 (2.7%)	1/146 (0.7%)	0.258 (0.04-1.78)
AE leading to discontinuation	4/113 (3.5%)	2/146 (1.4%)	0.387 (0.08-1.78)
Subjects with ≥ 1 injection site reactions	3/113 (2.7%)	5/146 (3.4%)	1.290 (0.35-4.84)
Subjects with tuberculosis	0/113 (0.0%) -	0/146 (0.0%)	-
Week 24	**n = 113 (includes 50 patients who switched to 100 mg dose plus 11 who discontinued treatment)**	**n = 146 (includes 28 patients who switched to 100 mg dose plus 11 who discontinued treatment)**	
Patients with ≥ 1 AE	67/113 (59.3%)	99/146 (67.8%)	1.144 (0.95-1.38)
Patients with ≥ 1 AE of severe intensity	12/113 (10.6%)	8/146 (5.5%)	0.516 (0.22-1.19)
Patients with ≥ 1 serious AE	7/113 (6.2%)	3/146 (2.1%)	0.332 (0.09-1.15)
Patients with ≥ serious infections	4/113 (3.5%)	1/146 (0.7%)	0.193 (0.03-1.27)
AE leading to discontinuation	5/113 (4.4%)	2/146 (1.4%)	0.310 (0.07-1.36)
Subjects with ≥ 1 injection site reactions	3/113 (2.7%)	7/146 (4.8%)	1.806 (0.52-6.35)
Subjects with tuberculosis	0/113 (0.0%) -	0/146 (0.0%)	-

At week 16 the proportion of patients with at least one adverse event, a serious adverse event, serious infection, or injection site reaction was not greater in the golimumab 50 mg group than the placebo group. None of the differences of these outcomes between the treatment and placebo group were statistically significant, but the trial was not powered to detect significant differences in these rare adverse events and the time period was short for the detection of serious and uncommon adverse effects. The data for golimumab at 24 weeks cannot be compared with placebo because the placebo group includes 50 patients who switched to 100 mg dose. The absolute rates of adverse events in the golimumab 50 mg group (which best reflects the use of golimumab in clinical practice) were high (67.8%) but the rates of serious adverse events and infections and discontinuations due to adverse events were low. No active tuberculosis was reported in any arms of the GO-REVEAL trial.[4, 5]

Data taken from an analysis of all golimumab doses in the GO-REVEAL trial[4, 5](n= 343) 64.7% experienced at least one adverse event, 5.0% experienced a severe adverse event, 2.0% a serious adverse event, 0.6% a serious infection, and 2.3% had an adverse event which resulted in discontinuation of therapy and 4.1% had an injection site reaction. In terms of the injection site reaction, at 16 weeks, the rates of this adverse event were: golimumab 50 mg 3.4%, golimumab 100 mg 3.4% and placebo 2.7%. The 24 week data showed a similar incidence of injection site reactions. These data suggested no additional burden of injection site reactions on patients receiving golimumab compared with placebo. The MS did not present data to facilitate a comparison between the adverse events of golimumab with those of the comparator biological for other adverse events.

Further longer-term follow-up safety data (e.g. at 52 and 104 weeks) from this trial would be valuable but were not available. There were no adverse event data of golimumab from controlled studies in other conditions such as rheumatoid arthritis and ankylosing spondylitis.

Given these limitations and uncertainties, the manufacturer's conclusion that golimumab is a safe treatment option similar to other anti-TNF agents may be premature and may not be reliable.

4.2.3. Summary

The manufacturer's submission on the clinical direct efficacy of golimumab was solely based on a single RCT (GO-REVEAL[4, 5]) in which the dosing regimen (including the dose adjustment) for the golimumab 50 mg group was generally representative of the routine clinical practice,[6] except it did not meet the criterion specified by the SPC that increasing the dose from 50 mg to 100 mg should be specifically considered on patients with more than 100 kg weight. However, the use of golimumab 100 mg in the GO-REVEAL trial was not reflective of clinical practice, as this dose is not licensed as a starting dose.[6] The limited available evidence indicates that golimumab appears to be an efficacious treatment for active and progressive PsA patients despite the use of previous DMARDs or NSAIDs.

The efficacy data from the relevant trial (GO-REVEAL[4, 5]) showed that golimumab 50 mg significantly improved patients' skin and joint disease outcomes, and functional status. At 14 weeks, compared with placebo, golimumab 50 mg significantly improved joint disease response as measured by ACR 20 (RR 5.73, 95% CI 3.24 to 10.56) and PsARC (RR 3.45, 95% CI 2.49 to 4.87), and skin disease response as measured by PASI 75 (RR 15.95, 95% CI 4.62 to 59.11). The effect sizes of point estimates were moderate to large, implying that these treatment effects could be clinically significant. The 24 week absolute data showed that these benefits were maintained. There was a significant improvement in patients' functional status as measured by HAQ change from baseline at 24 weeks

(-0.33, p<0.001), thereby achieving the minimum clinically significant threshold for PsA of -0.3.[1] The open-label extension data showed that the beneficial effects were also maintained at 52 and 104 weeks. Golimumab 100 mg achieved a similar magnitude of treatment effects during the follow-up. However, the ERG noted the limitations of the methodological quality in this trial. In particular, there was lack of robustness of the analyses of the 24 week data in terms of treatment effects relative to placebo, because of a failure to adjust the treatment contamination due to patients' crossing over between treatment arms.

In the absence of head-to-head comparison on the relative efficacy between different anti-TNF agents, a MTC analysis was performed to estimate the relative efficacy between the four anti-TNF agents in the MS. The degree of clinical heterogeneity between the included trials in terms of joint and skin disease severity and functional status was considered reasonable. Therefore, the assumption of exchangeability between the trials for the purposes of the MTC analysis was acceptable. The ERG also considered the statistical approach in the manufacturer's MTC analysis to be reliable.

The results of manufacturer's MTC showed that infliximab appears to be most effective of the four anti-TNF agents; infliximab was associated with the highest probabilities of response in terms of joint and skin disease outcomes. Golimumab achieved the third highest PsARC response in joint disease. Golimumab and etanercept had the lowest response in skin disease in terms of PASI change from baseline. In those patients who achieved a PsARC response, the highest mean improvement in the functional status (HAQ) was seen with etanercept ([redacted]), whilst the lowest mean improvement in HAQ was seen with golimumab ([redacted]). For all four anti-TNF agents, the changes in HAQ for those patients who did not achieve a PsARC response were [redacted]).[1]

The limited available evidence on the safety evaluation of golimumab suggested that the most frequently reported adverse events associated with golimumab therapy were infections and infestations, upper respiratory tract infection and nasopharyngitis. Serious adverse events including serious infection and malignancy were rare. No active tuberculosis in any treatment arm was observed. The evidence, however, was exclusively based on 24 week data from a single RCT with PsA patients (GO-REVEAL[4, 5]). The manufacturer failed to provide longer term data or to consider adverse event data of golimumab from controlled studies in other conditions such as rheumatoid arthritis and ankylosing spondylitis. Whilst the short-term adverse effects profile of golimumab

appears similar to other anti-TNF agents, the longer-term safety profile of golimumab remains uncertain.

5. ECONOMIC EVALUATION

This section focuses on the economic evidence submitted by the manufacturer in their initial report. 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 to highlight key assumptions and possible limitations. Section 6 presents a description of the additional information provided by the manufacturer following ERG points of clarification and a critique of this by the ERG, alongside additional work undertaken by the ERG to address any remaining uncertainties.

5.1. Overview of manufacturer's economic evaluation

This section provides a structured critique of the economic evaluation reported in the MS to NICE. The manufacturer's initial economic submission included:

1. A description of the databases and websites searched in the literature review (MS, p.193-195, Appendix 9.10). A description of the systematic search strategy used to identify existing cost-effectiveness studies for TNF-α inhibitors in psoriatic arthritis patients with active PSA (≤3 tender joints and ≤3 swollen joints) who have failed to respond to adequate treatment (>6 months) with non-biologic DMARDs.

2. A report on the *de novo* economic evaluation conducted by the manufacturer. The report described the technology; comparators and patient population; the categories of resource use costed; the resource use and unit cost assumptions and sources; the base-case results; and sensitivity analysis (MS, p.111-161).

3. An Excel-based model comprising the manufacturer's electronic economic model.

The manufacturer conducted a literature search to identify published cost-effectiveness studies (CEA) for TNF-α inhibitors in the treatment of PsA. A brief overview of the search was described in the MS (p.98) and search strategies are presented separately in Appendix 9.10 of the MS (p.193-195).

The databases searched for the cost effectiveness section included those defined by NICE in the specification for manufacturer/sponsor submission of evidence: MEDLINE, MEDLINE In-Process, EMBASE and NHS EED. Due to unavailability of access, EconLIT was not searched. The submission gave detailed descriptions of the search strategies and met NICE requirements. It included the specific databases searched; the service providers used; the dates when searches were conducted; the date spans of the searches; and the complete strategies used. The number of records identified for each search set was also presented.

The strategies and thesaurus terms used in each database varied but were appropriate to each database searched. In addition, the filter used to identify study types in the searches were also appropriate to each database searched. The ERG considered the search strategy for Section 9.10 (cost-effectiveness) to be appropriate.

The submission (MS p.99, Table B19) identified five cost-effectiveness analyses. Details of these five cost-effectiveness evaluations were summarised in the MS (p.99-110) and quality assessed (p.196-204, Appendix 11). The cost effectiveness analyses included the current assessment group model (Rodgers et al.[7]), the previous assessment group model,[17] two previous manufacturers' submissions (infliximab[21] and adalimumab[22]) and one published model.[23]

Following the literature search, the manufacturer developed a *de novo* economic model. An overall summary of the manufacturer's approach and signposts to the relevant sections in the MS are reported in Table 5.1.

Table 5.1: Summary of the manufacturer's economic evaluation (and signposts to MS)

	Approach	Source / Justification	Signpost (location in MS)
Model	Cohort model with common baseline HAQ and PASI (varied in sensitivity analysis). Lifetime treatment assumed	Not justified	pg. 111-119
States and events	Initial response according to PsARC determined. HAQ and PASI tracked over time, accounting for withdrawals.	Response as measured by clinical trials. Withdrawals based on Rodgers et al.[7]	pg. 114-117
Comparators	Golimumab, adalimumab, infliximab, etanercept and palliative care	Scope as specified by NICE	pg. 118
Natural History	Mean HAQ increase (worsening) of 0.0719 per year for patients on palliative care	Taken from Rodgers et al.[7]	pg. 123
Treatment effectiveness	As measured by relevant trials. Evidence synthesis model used.	Data from 7 trials (IMPACT,[10] IMPACT 2,[11] Mease 2000,[8] Mease 2004,[9] GO-REVEAL,[4, 5] Genovese 2007[13] and ADEPT[12]) were used in the evidence synthesis model.	
Adverse events	Not included	No relevant studies were identified	pg. 133, 144
Health related QoL	QALYs are measured as a function of HAQ and PASI. In the base case, an algorithm estimating the utilities based on HAQ and PASI in Rodgers et al.[7] was used	The GO-REVEAL data collected SF36 HRQOL data at each follow up. Regression was used to predict 'utility' from HAQ and PASI scores using this individual patient data. Two alternative methods to generate values for the utilities were explored: the Gray algorithm (selected as the base-case) and the Brazier algorithm.	pg. 128-132
Resource utilisation and costs	Resources use associated with drug acquisition, administration and monitoring taken from Rodgers et al.[7] Costs of nurse time for administration of golimumab, etanercept and adalimumab was included. Costs for hospital visits were taken from NHS reference costs. The acquisition cost of golimumab is assumed to be the same as adalimumab (assuming 50 mg dose). The ongoing costs of managing PsA were estimated as a function of HAQ. Costs of psoriasis, as a function of PASI, were derived from a questionnaire involving 35 dermatologists.	Rationale for including nurse time for administration of golimumab, etanercept and adalimumab was not included. Finished consultant episode (FCE) costs were taken from the reference costs, and are therefore likely to include drug costs. These were derived from the Kobelt et al.2002 study.[24] A copy of the questionnaire is presented. No details on how the questionnaire responses were synthesized were provided.	pg. 136-144
Discount rates	3.5% for QALYs and costs	NICE reference case	pg. 117
Sub groups	A sub-group analysis was conducted exclusively on patients presumed to have predominantly rheumatic disease (and no psoriasis). A subgroup analysis only including psoriatic arthritis patients with significant psoriasis was also modelled.	The proportions were estimated from clinical trials. No impact on PASI was assumed for those with predominantly rheumatic disease.	pg. 157
Sensitivity analysis	One way sensitivity analysis conducted on 17 key parameters and assumptions. Probabilistic sensitivity was also conducted.	Some justification was given on the choice of alternative parameter values and assumptions used in the sensitivity analysis. No justification was given for the exclusion of parameters and assumptions not subject to sensitivity analysis.	pg. 151-152 Table B30

5.1.1. Natural history

The patient cohort within the model was those with active and progressive PsA who have responded inadequately to DMARDs. Patients were assumed to have the same baseline characteristics as those observed in the GO-REVEAL trial (Section 5.3.2). The mean age was [redacted]; mean patient weight was assumed to be [redacted], baseline HAQ was [redacted], baseline PASI was [redacted] for those who have ≥3% body surface area psoriasis skin involvement and the proportion with psoriasis was [redacted]. For patients with no clinically significant psoriasis component to their disease, only the change in HAQ was modelled. Baseline (natural history) HAQ progression was 0.0719 taken from Rodgers et al.[7]

5.1.2. Treatment effectiveness within the submission

The intervention evaluated in the manufacturer's model is golimumab 50 mg (once a month). Golimumab is compared with other TNF-α inhibitors and palliative care; therefore the following treatment alternatives were compared:

- Golimumab: 50 mg given once a month, on the same date each month.
- Infliximab: 5 mg/kg given as an intravenous infusion over a 2 hour period followed by additional 5 mg/kg infusion doses at 2 and 6 weeks after the first infusion, then every 8 weeks thereafter.
- Adalimumab: 40 mg administered every other week as a single dose via subcutaneous injection.
- Etanercept: 25 mg administered twice weekly, or 50 mg administered once weekly.
- Palliative care comprising DMARDs

The primary outcome used in the modelling in the MS was QALYs, estimated as a function of both HAQ and PASI. An evidence synthesis model was used to determine the probability of PsARC response to anti-TNF agents at around 12 weeks, the associated HAQ for PsARC responders and non-responders and the average change in PASI from baseline, for each biologic drug. The evidence synthesis model used to generate HAQ and PASI changes at 12 weeks. An overview of the evidence synthesis model and the data used are presented in section 4.1.7.

Not all RCTs presented the HAQ change from baseline by PsARC responders and non-responders, as required by the decision model. Where RCTs only presented the average change in HAQ from baseline on biologic therapy and placebo (not distinguished by responder and non-responder). In these cases it was assumed that the HAQ change in non-responders was equal to the average value from the

other RCTs for that drug (MS: pg 127). Given these data and assumptions the HAQ change for responders in that RCT can be inferred.

For PsARC responders who continue with biologic therapy after 12 weeks, a further HAQ reduction was also assumed for the 2nd cycle (weeks 13-24) of -0.063 and 3rd cycles (weeks 25-76) of-0.031. The HAQ reductions for the 2nd and 3rd cycles are taken from an observational follow up of patients in the GO-REVEAL trial.[4, 5] This HAQ improvement represents a belief that biologic therapies are associated with some continuing long term reversal of the progression of arthritis as well as having an initial benefit. The same rate of long term improvement in HAQ was assumed for all anti-TNF agents. From the 4th cycle (week 77) onwards, the HAQ score for responders was assumed to remain constant as long as patients remain on TNF-α inhibitor treatment.

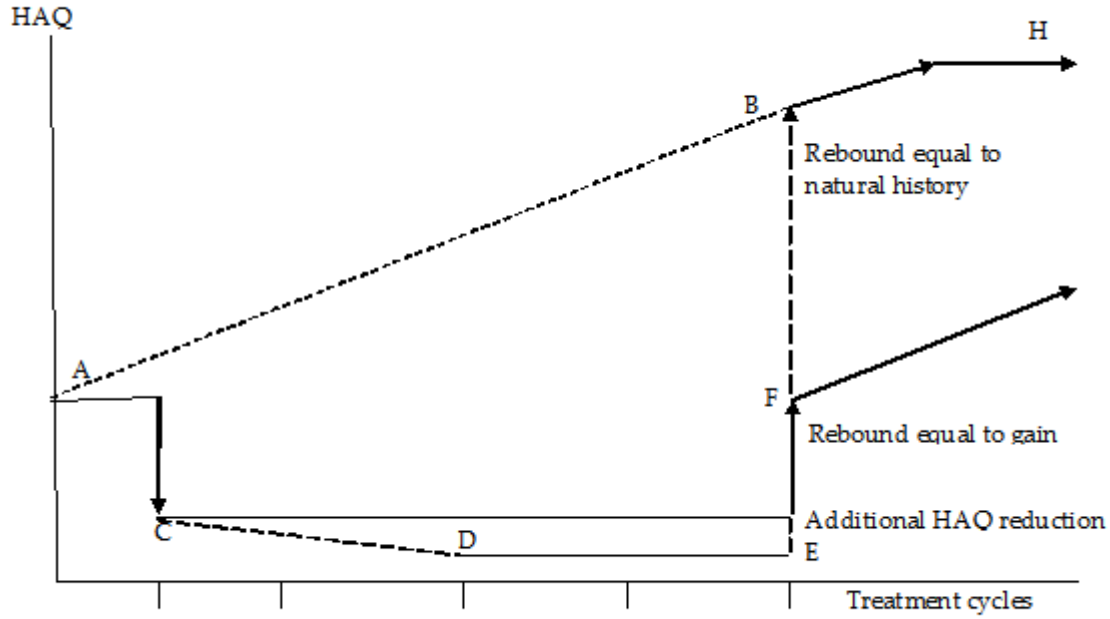
Patients who do not respond to PsARC at 12 weeks were assumed to have a benefit in terms of HAQ and PASI relative to baseline during the first cycle (0-12 weeks). The decision model made an adjustment for the so-called 'placebo effect'. This assumes that not all of the effectiveness of anti-TNF agents observed in RCTs would be generalisable to clinical practice, as patients are unlikely to incur the additional benefit associated with being observed as part of a randomised trial (the 'placebo effect'). To make the results more generalisable the placebo effect was adjusted for by subtracting the mean HAQ change in the placebo group (across PsARC responders and non-responders) from the HAQ change of patients on biologic therapy. A common placebo effect is applied across treatments in the base-case analysis, thus this will have a minimal impact upon the comparisons between TNF-α inhibitors. Differential placebo effects are applied in a sensitivity analysis.

In the base-case, 27% of the patient cohort was assumed to have no psoriasis and 73% had an average PASI of 9.9 at baseline. For patients with psoriasis, the decision model required an estimate of the mean absolute change in PASI after initiating biologic therapy. Absolute changes in PASI were inferred from the proportional changes reported in trials (that is, PASI 50, PASI 75 and PASI 90). The treatment benefit in terms of the PASI reduction was assumed only in the 1st cycle after which TNF-α inhibitor treatment was assumed to offer no additional PASI reduction. The MS stated that the golimumab RCT showed that PASI response was independent of PsARC response. The same lack of association between PASI and PSARC responses was assumed for all treatments. All patients with psoriasis in the model start with the same PASI score. PASI change was not assumed to be correlated with baseline score.

A sensitivity analysis was conducted assuming that HAQ progression for responders who are maintained on anti-TNF agents is the same rate as natural history (that is, without biologic therapy). No data was presented in the MS that estimated the change in HAQ following withdrawal from a

biologic drug (known as 'rebound'). Instead sensitivity analyses were carried out. Rebound was modelled under two scenarios: 1) rebound equal to the amount of HAQ gain (E to F) with natural history disease progression thereafter; and 2) rebound equal to the natural history disease progression as would have occurred from baseline with only palliative care (E to B) (see Figure 5.1).

Figure 5.1: HAQ reduction and rebound effect (MS: pg 116, Figure B11)



The arthritis component of patients on palliative care was assumed to progress (worsen) in line with natural history (0.0719 annually, as measured by HAQ), as estimated from the Leeds NESPAR study.[25] The distribution placed on this parameter for the probabilistic sensitivity analysis assumed that the rate of progression can only be non-negative (that is, continued worsening of arthritis in the absence of biologic therapy). The natural history of PASI was assumed constant over time, based on expert opinion (source for this is not stated). Following withdrawal of biologic therapy, patients were assumed to return to their baseline PASI score.

Adverse events were not included in the analysis. It was assumed that patients suffering from serious adverse events would not be counted as a PsARC treatment responder in the ITT analysis and would withdraw from biologic therapy. The MS stated that due to unavailability of data, the impact of serious adverse events leading to lasting complications or adverse events leading to temporary withdrawal from treatment and the associated disutility were not considered.

Annual withdrawals from treatment were taken from the previous York MTA[7] (16.5% per annum). The same withdrawal rate was applied to all strategies. After withdrawal, patients are assumed to go

onto palliative care in the base case model. Patients also were assumed to have an annual risk of death 1.6 to 1.7 times greater than the general population (Wong et al.[26]), but mortality was assumed to be independent of the therapy.

5.1.3. Health-related quality of life

In the base case, an algorithm estimating the utilities based on HAQ and PASI used in the previous MTA was used.[7] The regression estimates used in the base case are shown below in Table 5.2. A separate regression analysis using patient level data from the GO-REVEAL trial was used to predict 'utility' from HAQ and PASI scores. This was based on mapping of SF-36 data on to EQ-5D collected SF36 HRQOL data at each follow up. Two alternative methods to generate values for the utilities were explored: the Gray algorithm (selected as the base-case) and the Brazier algorithm (MS: pg 129-135). The Gray algorithm converts SF-36 to EQ-5D health states and then to utilities, whereas the Brazier algorithm estimates utilities directly from SF-36. Explanatory variables used in the model were: HAQ, PASI, HAQ squared and PASI squared. Interaction between PASI and HAQ was not explored. Following request by the ERG for clarification, this regression was revised to include an additional explanatory interaction variable, HAQ * PASI.

Table 5.2: Summary of quality of life values used in the base case manufacturers model (MS, pg 134, 209-210)

State	Regression estimate	SE
Intercept	0.897	0.006
HAQ	-0.298	0.006
PASI	-0.004	0.0003

5.1.4. Resources and costs

The costs, for each comparator for cycles 1, 2 and annual cycles thereafter, are shown below in Table 5.3. Resource use associated with treatment, administration and monitoring of infliximab, etanercept and adalimumab was taken from the previous York MTA model.[7] Resource use for golimumab was assumed to be the same as adalimumab. The MS stated that the systematic review for resource use data was updated but did not find any additional information (MS pg. 140). No details of this updated search were presented. In addition, four hours of staff nurse time were included for the initial cycle for golimumab, etanercept and adalimumab. No justification of including this additional cost was given.

The ERG understands that there is not yet a list price for golimumab, but the manufacturer assumed that the annual acquisition cost (including administration and monitoring) of golimumab (50 mg once

per calendar month) was comparable to adalimumab. The BNF [27] was used to cost other medications. Costs for infliximab were calculated assuming that vial optimization was allowed (3.5 vials).

The monitoring costs used in the manufacturer's model were those presented in Rodgers et al.[7] The administration costs, however, differ. Costs for outpatient and inpatient days are taken from the NHS reference costs but cover the costs for a finished consultant episode (FCE) (MS pg. 137-139, Table B27) and are, therefore, likely also to include any medications taken and procedures occurring during these visits.

Table 5.3: Summary of costs used in the manufacturers model (MS, pg 141, Table B28)

Items	Golimumab	Infliximab	Etanercept	Adalimumab
Treatment cost				
Cycle 1	£2,145.00	£5,363.37	£2,145.12	£2,145.00
Cycle 2	£2,145.00	£3,575.58	£2,145.12	£2,145.00
Annual cycle	£9,295.00	£11,620.64	£9,295.52	£9,295.00
Administration cost				
Cycle 1	£330.71	£372.00	£330.71	£330.71
Cycle 2	£91.50	£248.00	£91.50	£91.50
Annual cycle	£0.00	£806.00	£0.00	£0.00
Monitoring cost				
Cycle 1	£182.28	£90.78	£182.28	£182.28
Cycle 2	£72.30	£40.34	£72.30	£72.30
Annual cycle	£313.30	£174.81	£313.30	£313.30
Total				
Cycle 1	£2,658.00	£5,826.15	£2,658.12	£2,658.00
Cycle 2	£2,308.81	£3,863.92	£2,308.93	£2,308.81
Annual cycle	£9,608.30	£12,601.45	£9,608.82	£9,608.30

Ongoing costs as a function of HAQ were derived from the Kobelt, 2002 study.[24] Patients on biologic treatment only incur 85% of these costs (as they are assumed not to be using DMARDs), whilst those withdrawing from treatment incur 100% (assumed to revert to DMARDs). Ongoing costs as a function of PASI were estimated using the results of an independent data collection exercise. The information was collected using an internet based survey of 22 dermatologists (MS Appendix 16). The resource use estimates thus derived were the means of the resource use estimated by the individual respondent. The results from the survey estimated an additional cost of [redacted] including phototherapy costs and an additional [redacted].

5.1.5. Discounting

The manufacturer's model applied a discount rate of 3.5% per annum to expected costs and health effects (MS p.146, Table B30), in line with the NICE reference case.

5.1.6. Sensitivity analyses

A sub-group analysis was conducted based on the patients presumed to have predominantly rheumatic disease. For these patients only, the impact upon their rheumatic component was modelled, estimated using HAQ. No impact on the dermatological component of quality of life was assumed and, therefore, the PASI impact on utility was not modelled. The HAQ change was assumed identical for the subgroups with or without BSA $\geq$ 3% at baseline. This is suggested by the golimumab data, and is a model simplification. A subgroup analysis only including psoriatic arthritis patients with significant psoriasis was also modelled in the subgroup analysis.

Scenario analysis and probabilistic sensitivity analyses (PSA) were also undertaken by the manufacturer. A range of alternative scenarios was used to explore the implications of distinct model assumptions and of the use of alternative sources of data. These are summarised in Table 5.4 and are described in full in p.145-146, 150-151 of the MS. In addition a scenario specifying the HAQ rebound following withdrawal equal to natural history (as opposed to initial gain) was conducted.

Table 5.4: Model parameter values assumed in the base case and assumptions used in the PSA, in the MS

Variable	Base case	Sensitivity analysis	Rationale
Time horizon	40 years	5 years / 20 years	Shorter time horizon as limited long term data available
Discount rate	3.5%	0 - 6%	NICE reference case
Females	[redacted]	0 – 100%	Not given
Age	[redacted]	30 – 60 yrs	Not given
Mean weight	[redacted]	60 - 80 kg	Not given
Baseline HAQ score	[redacted]	± 50% change	Not given
Baseline PASI score	[redacted]	± 50% change	Not given
Placebo HAQ responses	Common	Individual from TNF-α inhibitor trials	Not given
Withdrawal rates	16.5%	11.14%	Taken from the previous assessment group model [17]
Psoriasis Costs	Included	Excluded	A proportion of patients do not have significant psoriasis
Phototherapy costs	Included	Excluded	Some psoriasis patients do not require phototherapy
QoL data	Rodgers et al.	Algorithm based on golimumab trial data	Not given
Golimumab annual acquisition cost	Equivalent to adalimumab	± 20%	Not given
HAQ change for responders	Continued up to 3 cycles	No HAQ benefit beyond the first cycle	Consistent with the previous NICE appraisals
HAQ change for non-responders	Trial based HAQ benefit in cycle 1	No HAQ benefit for non-responders	Not given
PASI change for non-responders	Trial based PASI benefit in cycle 1	No PASI benefit for non-responders	Not given
Natural history HAQ progression	0.0719 / year	0.1018 / year	Current evidence synthesis (Placebo HAQ change for non-responders)
PsA management cost of HAQ on TNF-α inhibitors	85% of costs for patients on palliative care	± 15%	Not given

5.1.7. Model validation

The MS did not report on any internal validation and debugging of the model. Instead it stated that the methods and results have been validated against those available in the literature, in particular the model by Woolacott et al.[17] and Rodgers et al.[7] The MS reported on the variations in methodology but stated that it is not anticipated that these will have a significant impact.

The clinical outcomes estimated by the model were compared with those from the selected trial results (MS pg. 148, Table B31). These were broadly similar for the majority of outcomes; however, there were some discernable differences, for example the PsARC response and number of patients still on treatment at the end of 2nd annual model cycle for adalimumab. These were not discussed further in the MS.

5.2. Critique of approach used

The manufacturers presented a cohort Markov model constructed in Excel with evidence synthesis undertaken in WinBUGS. This model has been checked and cross referenced with the results provided in the MS. The ERG confirms that the model results presented in the MS are equivalent to those generated by the deterministic and probabilistic version of the Excel model. The ERG can also confirm that the descriptions of the model provided in the MS are comparable to the execution of the model in Excel, with the exception of those clarifications and additional analyses described in Section 6.1.

The ERG has assessed the manufacturer's economic evaluation using the Philips et al.[28] checklist for quality assessing decision analytic models. This is shown in Appendix 5.1 and is used to assist the narrative critique in the following sections. In Table 5.5 the methods used in the manufacturer's model are also compared to those detailed in the NICE reference case.

Table 5.5: A consideration of the MS using a checklist based on NICE's reference case and other methodological recommendations, together with an indication of the inclusion of each of the elements in the MS and ERG's comments on whether the *de-novo* evaluation meets the requirements of NICE reference case

Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether de-novo evaluation meets requirements of NICE reference case
Comparator(s)	Therapies routinely used in the NHS, including technologies regarded as current best practice	Yes	Yes
Type of economic evaluation	Cost-effectiveness analysis	Yes	Yes
Perspective on costs	NHS and PSS	Yes	Yes
Perspective on outcomes	All health effects on individuals	Yes	Yes
Time horizon	Sufficient to capture differences in costs and outcomes	Yes	Yes
Synthesis of evidence on outcomes	Systematic review	Yes	Yes
Measure of health effects	QALYs	Yes	Yes
Source of data for measurement of HRQL	Reported directly by patients and/or carers	Yes	Yes
Source of preference data for valuation of changes in HRQL	Representative sample of the public	Yes	Yes
Discount rate	Annual rate of 3.5% on both costs and health effects	Yes	Yes
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes	Yes
Sensitivity analysis	Probabilistic sensitivity analysis	Yes	Yes

Despite its adherence to the reference case methodology, the ERG identified a number of shortcomings with the manufacturer's model. These are discussed in more detail in the sections below.

5.2.1. Interventions and comparators

All relevant comparators and interventions have been compared in the model. The comparator to TNF-α inhibitors was palliative care, which was defined as DMARDS. As the model uses a homogeneous cohort of patients considered representative of the groups of patients eligible for biologic therapies to treat PsA (that is, patients who have failed two or more conventional DMARDs) there is a possibility that palliative care would be the treatment option. However, palliative care may

consist of no therapy rather than further DMARDs. It is worth noting that the use of DMARDS to represent palliative care may have the potential to artificially inflate the cost-effectiveness of TNF-α inhibitors as, in practice, DMARDs are liable to be more effective than palliative care.

5.2.2. Structure of the model

In the base-case model, a decision is made to continue or withdraw from TNF-α inhibitors according to PsARC response at 12 weeks. In addition to the 12 week PsARC response decision rule, the model was constructed with the flexibility to allow a 24 week decision rule. Given that the license allows a higher dose of 100 mg per administration for patients over 100 kg if no response is achieved at 12 weeks[6] and a number of patients switched to a higher dose of golimumab in GO-REVEAL (Section 5.3.2) after failing to achieve a response, it may have been appropriate to have included this scenario in the sensitivity analysis. Further analysis on this issue was requested from the manufacturer but they stated that suitable data were not available to model this option.

The graphical depiction and explanation of the model were provided, however some clarifications were required in order to allow the ERG to fully comprehend the structure. These clarifications were provided and on the whole, the ERG felt that the model structure was reasonable. To assess the full impact of structural assumptions, the manufacturer's data were used as inputs in a model developed by the ERG to evaluate three anti-TNF agents (etanercept, adalimumab and infliximab) in Rodgers et al[7] (See Section 6).

5.2.3. Natural history

The baseline patients characteristics upon which the modelling is based are those observed in the GO-REVEAL trial (Section 5.3.2). These patents had active and progressive PsA and have responded inadequately to DMARDs. The baseline characteristics (age, patient weight, HAQ, PASI and proportion with psoriasis) are considered by the ERG to be appropriate for this population.

HAQ progression while not on biologic therapy (also called natural history progression) is estimated using the Leeds cohort study data.[25] The Leeds dataset is, however, small, including only 24 patients. In addition patients surveyed do not meet the requirements for this analysis in that many have not failed at least 2 previous DMARDs. It is also not clear if patients met the current guideline criteria for initiating anti-TNF agents for PsA (3 tender and 3 swollen joints).

5.2.4. Treatment effectiveness within the submission

Treatment effectiveness was derived through the use of a Bayesian indirect comparison (MTC). An outline of the evidence synthesis model is provided in Section 4.1.7. The base case model defined response to treatment by PsARC at 12 weeks, only those obtaining a PsARC response remained on treatment. This accords with NICE and BSR guidelines and is widely accepted as an appropriate outcome to assess response in psoriatic arthritis. The model predicted a change in HAQ conditional on PsARC response, which allowed a different HAQ outcome across all drugs for responders and non-responders. This approach implicitly incorporates a correlation between PsARC and HAQ outcomes, a correlation which was supported by the RCT evidence presented. A reduction (improvement) in HAQ for the first three cycles (12 weeks, 24 weeks and 52 weeks) was assumed, after which, HAQ for a responder was assumed to remain constant. The HAQ reductions for the 2nd and 3rd cycles appear to have been derived from the open label follow up period of patients in the GO-REVEAL trial; however, this was not explicit in the MS (pg 116) and clarification was sought from the manufacturer. For non-responders according to PsARC at 12 weeks, the HAQ change for non-responders is only applied for the first cycle after which it is assumed that the patient withdraws from biologic therapy and receives a natural history rate of progression (worsening) of HAQ (0.0719 per year taken from Rodgers et al.[7]).

PASI change from baseline at 12 weeks was estimated only for the subgroup with >3% body skin area and was independent of PsARC response. This assumption was supported by data from the GO-REVEAL trial, although it was applied across all TNF-α inhibitors. Given that this assumption was supported by limited data, some sensitivity analysis may have been warranted; although the ERG acknowledges the difficulty in undertaking such an analysis. Further to an ERG request for clarification, data to support the assumption of no correlation between PASI 75 and PsARC for the GO-REVEAL trial was provided by the manufacturer in the form of two-by-two tables.

The average absolute PASI change from baseline for each TNF-α inhibitor was used only for patients with clinically significant psoriatic component. The benefit, in terms of PASI reduction, was assumed only in the first cycle after which treatments were assumed to offer no additional PASI reduction. A number of additional assumptions were made relating to the PASI data including:

- All patients with BSA >3% were assumed to have identical PASI baseline values (equal to the mean PASI baseline score reported for this subgroup in the GO-REVEAL trial.

- The relationship between the proportions achieving different levels of PASI response (i.e. PASI 50, PASI 75 and PASI 90) in the GO-REVEAL trial was used to facilitate the calculation of average absolute change in PASI score.
- Absolute PASI change was not correlated with the PASI baseline score.

These assumptions may have been necessary to facilitate modelling; however, they are all based on one single trial and may have warranted further sensitivity analysis. TNF-α inhibitors are intended treat both joint disease and psoriasis. Clinical response at 3 months is measured using the PsARC for joints and PASI 75 for skin condition for these two aspects respectively. Whilst evidence from the GO-REVEAL trial shows no correlation between these responses, other data[7] suggests that they are not independent and thus further sensitivity analyses exploring this issue may have been appropriate. The model also assumed that psoriasis will not progress on or off treatment; that is, psoriasis will not on average worsen over time. This assumption was justified quoting clinical opinion, although this is not fully referenced in the MS.

The manufacturer's evidence synthesis estimated the treatment effect as the difference (or odds ratio) between the response rate in the intervention (TNF-α inhibitor) arm and the placebo arm. This considered what treatment effect is likely to be observed in general practice and assumed that the 'placebo effect', if it exists in the RCT, would apply equally to the intervention and the placebo. However, it is often unclear whether the 'placebo effect' would be observed in general practice if anti-TNF agents are given. In the MS model, the average HAQ gain in the placebo arm of the RCTs is subtracted from the HAQ gain for responders and non-responders on anti-TNF agents, estimated by the evidence synthesis model. However, palliative care was actually DMARDs (an active treatment) which may have led to an overestimation of the effectiveness of TNF-α inhibitors. This overestimation is likely to be small, and given that the same adjustment is applied to all TNF-α inhibitors, is unlikely to bias the comparison between TNF-α inhibitors but may affect the comparison with palliative care.

The MS model utilized UK life tables along with PsA specific mortality multipliers (Wong et al.[26]) to estimate mortality. The same mortality rate was assumed for all treatments and for no treatment (i.e. there was no differential impact of the alternative therapies on mortality). The ERG considers this assumption to be reasonable, although there may be a beneficial effect of TNF-α inhibitors on mortality; however, limited data to quantify this are not available

5.2.5. Withdrawals

In the MS patients are withdrawn from treatment if they are PsARC non-responders at 12 weeks, irrespective of PASI response. To estimate the probability of withdrawal whilst receiving TNF-α inhibitors, due to either loss of efficacy or adverse events, the model employed the same rates for TNF-α inhibitors as used in Rodgers et al.[7] (0.165 per year beyond the initial 12 week period). No withdrawals (from DMARDs) were assumed in the palliative care group. This assumption is considered appropriate by the ERG.

5.2.6. Health related quality of life

The MS model utilized the utility algorithm estimated by Rodgers et al.[7] in the base case analysis. It then explored two alternative methods to generate utilities: the Gray algorithm[29] and the Brazier algorithm.[30] The Gray algorithm converted SF-36 profiles to EQ-5D profiles and then EQ-5D profiles to utilities. The Brazier algorithm estimated utilities directly from SF-36. The Gray algorithm was used in the base-case analysis. The GO-REVEAL[4, 5] trial data were used in a multiple regression model using HAQ, PASI, HAQ squared and PASI squared, with no interaction terms, as explanatory variables. Again any interaction between HAQ and PASI was not explored.

There is some uncertainty regarding which of the regression models is appropriate to generate utilities. The ERG requested further analyses from the manufacturer using alternative functional forms (see Section 6.1).

5.2.7. Resource utilisation and costs

The MS states that it undertook an update of the systematic searches for resource use conducted by Rodgers et al.[7] However, no further details were provided. The MS reported that no additional resource use data was identified; therefore, the resource use information from the Rodgers et al. report[7] were used. The MS also included an additional 4 hours of staff nurse costs, apparently to cover training of patients to self-administer subcutaneous TNF-α inhibitors. The ERG considered that this may be unnecessary (that is, double-counting) and further justification for this assumption was requested.

The MS also used a survey of clinicians' opinions based on vignettes of 'typical cases' to estimate the costs associated with treating psoriasis. The MS did not give details of the statistical analyses used to

summarise these survey data or the unit costs used in their calculations. Further clarification was sought from the manufacturer.

All other costs used within the MS model were thought appropriate by the ERG.

5.2.8. Discounting

Discounting was appropriately conducted.

5.2.9. Subgroup analysis

Some subgroup analysis was undertaken on the impact of TNF-α inhibitors on patients with predominately rheumatic disease and a subgroup of patients with significant psoriasis. These analyses seemed appropriate given that PsA can have variable impact on both the joint and skin component of the disease.

5.2.10. Sensitivity analysis

The manufacturer undertook a detailed set of scenario analyses and probabilistic sensitivity analysis (PSA). However, the ERG considers that parameter uncertainty was not fully explored. This is because not all relevant parameters seem to have been considered uncertain in PSA (see Table 5.7). The ERG considers this to preclude a correct characterisation of uncertainty.

5.3. Results included in manufacturer's submission

It should be noted that all the results presented in this section (5.3) as taken from the MS, have the ICERs calculated incorrectly. Further details are given below.

The base case results of the model are presented in the manufacturer's submission p.150, Table B32. These are reproduced below in Table 5.6.

Table 5.6: Results from the base case

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) versus Palliation	ICER (£) vs TNF-α inhibitors
Palliation	£62,224	5.44				-
Adalimumab	£86,410	6.97	£24,186	1.53	£15,820	£15,820
Golimumab	£94,151	7.34	£7,740	0.37	£16,811	£20,901
Etanercept	£94,578	7.69	£428	0.35	£14,402	£1,232
Infliximab	£100,691	7.69	£6,112	0.00	£17,149	Dominated

These results suggest that golimumab is a cost-effective strategy compared to both palliative care and other TNF-α inhibitors at reasonable values for the threshold. However, the ICERs shown in Table 5.6 are incorrectly calculated. The comparisons with palliative care are not appropriate as strategies should be compared with the next best (in terms of QALYs) strategy excluding dominated and extendedly dominated strategies. The dominated and extendedly dominated strategies were also not excluded from the comparisons with other TNF-α inhibitors in the final column of Table 5.6. It is important to exclude extended dominated strategies in order to avoid inconsistent decision making. According to Table 5.6, the ICER of golimumab versus adalimumab is £20,901 per QALY. However the ICER of etanercept versus golimumab is only £1,232 per QALY. As etanercept is more effective (greater QALYs) than golimumab and has a lower cost-per-QALY, golimumab would never be chosen as long as etanercept is a viable alternative. Therefore golimumab should be excluded from the analysis and the ICERs recalculated. An appropriate incremental analysis is presented in Section 6.1.

A number of sensitivity and subgroup analyses were conducted by the MS. These are presented in the manufacturer's submission p.151-159, Tables B33- B36. These are also reproduced below in Table 5.7 – 5.10 below.

5.3.1. Univariate sensitivity analysis

The results of these univariate sensitivity analyses, presented by the manufacturer, are shown below in Table 5.7. It should be noted that as with the base case results presented in the MS, the ICERs are incorrectly calculated in this scenario analysis and only presented compared to palliative care. The ERG requested that this be revised (see Section 6.1).

Table 5.7: Sensitivity analysis conducted by the manufacturer, showing the incremental cost-effectiveness ratio (ICER) of golimumab versus palliative care

Variable	Base case	Parameter change	ICER vs. Palliative care
Time horizon	40 years	5 years 20 years	£41,799 £20,446
Discount rate	3.5%	0% costs & 0% outcomes 0% costs & 3.5% outcomes 3.5% costs & 0% outcomes	£12,396 £39,978 Dominant
Females	[redacted]	All males All females	£17,095 £16,367
Age	[redacted]	30 yrs 60 yrs	£15,478 £20,348
Baseline HAQ score	[redacted]	+ 50% change - 50% change	£18,802 £16,014
Baseline PASI score	[redacted]	+ 50% change - 50% change	£16,939 £16,807
Placebo HAQ responses	[redacted]	Individual from TNF-α inhibitor trials	£16,864
Withdrawal rates	[redacted]	11.14%	£17,311
Psoriasis Costs	Included	Excluded	£18,043
Phototherapy costs	Included	Excluded	£17,652
QoL data	Rodgers et al.	Algorithm based on previous NICE appraisal (Bravo Vergel 2007)	£19,218
Golimumab annual acquisition cost	Equivalent to adalimumab	+ 20% change - 20% change	£20,617 £13,004
HAQ change for responders	Continued up to 3 cycles	No HAQ benefit beyond the first cycle	£18,642
HAQ change for non-responders	Trial based HAQ benefit in cycle 1	No HAQ benefit for non-responders	£16,819
PASI change for non-responders	Trial based PASI benefit in cycle 1	No PASI benefit for non-responders	£16,839
Natural history HAQ progression	0.0719	0.1018	£14,825
PsA management cost on TNF-α inhibitors	85% of costs for patients on palliative care	+ 15% change - 15% change	£17,317 £16,305

5.3.2. Scenario analysis

A scenario analysis assuming a HAQ rebound equal to natural history following withdrawal from TNF-α inhibitors was performed by the manufacturer. This is a more pessimistic assumption than the base case which assumes that HAQ will rebound by the amount of the original gain. The results of this scenario, presented by the manufacturer, are shown below in Table 5.7. As with the base case results, the ICERs were incorrectly calculated by the MS in this scenario analysis. The correct ICERs are presented in Section 6.1.

Table 5.8: Results of the structural sensitivity analysis (rebound equal to natural history) reported in MS

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) versus Palliation
Palliation	£62,224	5.44			
Adalimumab	£87,533	6.20	£25,309	0.76	£33,514
Golimumab	£95,577	6.36	£8,044	0.16	£36,402
Etanercept	£96,028	6.69	£451	0.33	£27,090
Infliximab	£102,173	6.67	£6,145	-0.03	£32,693

5.3.3. Subgroup analysis

The results of the subgroup analysis, presented by the manufacturer, are shown below in Tables 5.8 and 5.9. As with the base case results, the ICERs were incorrectly calculated by the MS in this scenario analysis. The correct ICERs are presented in Section 6.1.

Table 5.9: Results of the subgroup analysis (rheumatic patients only)

	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) versus Palliation
Palliation	£40,275	5.85			
Adalimumab	£66,377	7.35	£26,102	1.50	£17,405
Golimumab	£74,542	7.71	£8,165	0.36	£18,378
Etanercept	£74,767	8.06	£225	0.35	£15,557
Infliximab	£81,990	8.04	£7,223	-0.03	£19,069

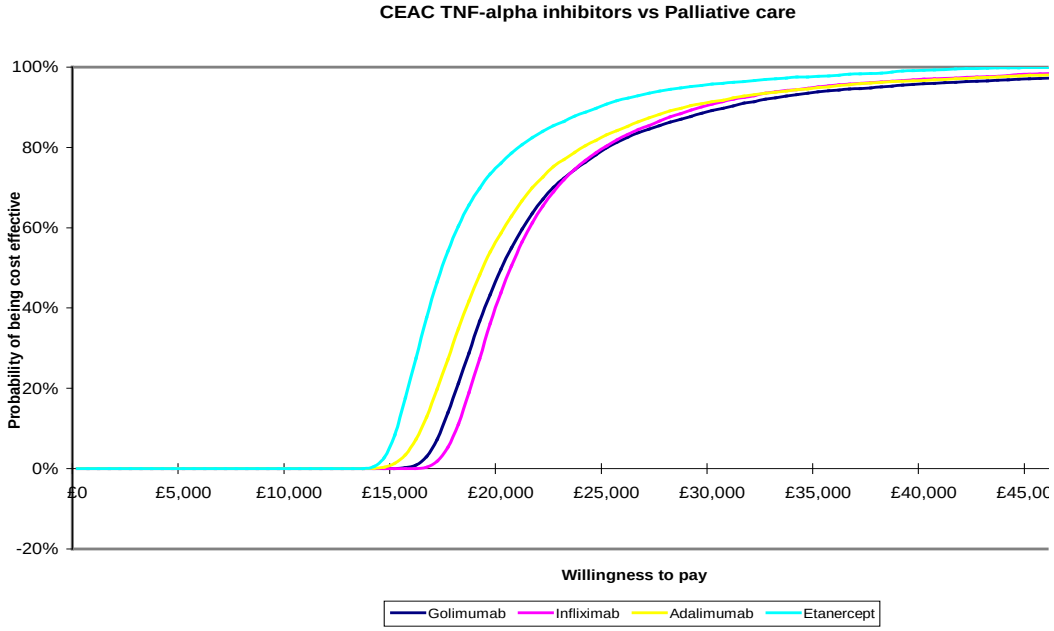
Table 5.10: Results of the subgroup analysis (rheumatic patients with significant psoriasis)

	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) versus Palliation
Palliation	£70,342	5.30			
Adalimumab	£93,820	6.83	£23,478	1.54	£15,249
Golimumab	£101,403	7.21	£7,583	0.37	£16,245
Etanercept	£101,906	7.55	£503	0.35	£13,982
Infliximab	£107,608	7.56	£5,702	0.01	£16,462

5.3.4. Probabilistic sensitivity analysis

A probabilistic sensitivity analysis was conducted by the manufacturers (MS: Section 6.7.8). This specifies parametric distributions for uncertain parameters within the model and samples from these distributions, using Monte Carlo simulation, to produce a distribution of costs and QALYs for each of the strategies. The results of the probabilistic sensitivity analysis were presented using a cost-effectiveness acceptability curve (CEAC), showing the probability that each strategy is the most cost-effective at various values for the cost-effectiveness threshold. The manufacturers CEAC is shown below in Figure 5.2.

Figure 5.2: CEAC of TNF-α inhibitors compared to palliative care



5.4. Comment on validity of results presented with reference to methodology used

Throughout the MS the incremental cost-effectiveness ratios were incorrectly calculated. The manufacturer failed to exclude extendedly dominated strategies. The base case results have been recalculated and are presented below (Table 5.11). As the table shows, etanercept compared to palliative care has an ICER of £14,379, with all other comparators dominated or extendedly dominated.

Table 5.11: MS base case results with ICERs correctly calculated

	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£)
Palliation	£62,224	5.44			
Adalimumab	£86,410	6.97	£24,186	1.53	ED
Golimumab	£94,151	7.34	£7,740	0.37	ED
Etanercept	£94,578	7.69	£428	0.35	£14,379
Infliximab	£100,691	7.69	£6,112	0.00	Dominated

Key: QALY quality adjusted life year; ICER incremental cost effectiveness ratio; ED extendedly dominated

The correct ICERs for the scenario analysis and subgroup analysis are all presented below in Table 5.12 and Table 5.13. The ERG requested that the ICER for the univariate sensitivity analysis be revised (see Section 6.1).

Table 5.12: Results of the subgroup analysis (rheumatic patients only) – correct calculation of ICERs

	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£)
Palliation	£40,275	5.85			-
Adalimumab	£66,377	7.35	£26,102	1.50	ED
Golimumab	£74,542	7.71	£8,165	0.36	ED
Etanercept	£74,767	8.06	£225	0.35	£15,607
Infliximab	£81,990	8.04	£7,223	-0.03	Dominated

Key: QALY quality adjusted life year; ICER incremental cost effectiveness ratio; ED extendedly dominated

Table 5.13: Results of the subgroup analysis (rheumatic patients with significant psoriasis) – correct calculation of ICERs

	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£)
Palliation	£70,342	5.30			-
Adalimumab	£93,820	6.83	£23,478	1.54	ED
Golimumab	£101,403	7.21	£7,583	0.37	ED
Etanercept	£101,906	7.55	£503	0.35	£14,028
Infliximab	£107,608	7.56	£5,702	0.01	£570,200

Key: QALY quality adjusted life year; ICER incremental cost effectiveness ratio; ED extendedly dominated

5.5. Summary of uncertainties and issues

A number of potential uncertainties are identified and described in Section 5.2 and summarised in Table 5.14. Several of these issues were subject to additional analyses by both the manufacturer, as part of their response to the ERG's points for clarification, and the ERG. The results of these additional analyses are presented in Section 6.

Table 5.14: Summary of uncertainties and issues identified in Section 5.2

Topic, uncertainty or issue	**Likely consequences for the results and conclusions**	**Additional analysis by manufacturer**	**Additional analysis by ERG**
Impact of structural assumptions (See Section 6).	Incorrect estimates of cost-effectiveness (ICER) and extent of uncertainty	-	Manufacturer's data were used as inputs in a model developed by the ERG
Evidence synthesis appropriateness (Section 5.2.2)	Incorrect estimates of PsARC response, HAQ gain and PASI change	Details of no correlation between PASI 75 and PsARC for the GO-REVEAL trial	Updating of ERGs evidence synthesis model to include data from GO-REVEAL Sensitivity analysis utilising ERGs evidence synthesis estimates in manufacturers model
Estimates of utilities	Incorrect estimates of QALYs	Re-estimating of utility algorithm	-
Resource use associated with administration of TNF-α inhibitors	Incorrect estimates of total costs	-	Sensitivity analysis excluding cost of staff nurse time in administration
Costs associated with psoriasis	Unclear calculation of estimates of mean costs per PASI point per year	Clarification on derivation of mean costs for PASI	Sensitivity analysis increasing and decreasing the costs associated with PASI

6. ADDITIONAL 'EXPLORATORY' OR OTHER WORK UNDERTAKEN BY THE ERG

6.1. Additional work undertaken by the manufacturer

A number of clarifications and additional analyses were requested from the manufacturer by the ERG. The justifications for these requests are discussed in Section 5.2 and the revised results are presented and discussed in the following section.

Two of the issues highlighted by the ERG led to a revised base-case analysis being presented by the manufacturer. These issues focused on (i) the QoL algorithms used (ii) vial sharing of infliximab between patients. The Gray algorithm used in sensitivity analysis for HRQOL included PASI- squared and HAQ squared terms. The ERG queried whether these coefficients were statistically significant and their impact on QOL. The ERG requested that these terms be excluded from the regression and the revised coefficients presented. In addition, the ERG requested a sensitivity analysis with the decision model using the revised Gray algorithm. The revised estimates are presented in Tables 6.1-6.4 below.

Table 6.1: Using the SF-36 data via Gray algorithm: combined data

Covariate	**Mean**	**Variance-Covariance matrix**			
		Intercept	**HAQ**	**PASI**	**HAQ x PASI**
Intercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HAQ	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PASI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HAQ x PASI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 6.2: Using the EQ-5D data: combined data

Covariate	**Mean**	**Variance-Covariance matrix**			
		Intercept	HAQ	PASI	HAQ x PASI
Intercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HAQ	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PASI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HAQ x PASI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Note: 7.779e-01 means 0.779; -3.259e-03 means -0.003259

Table 6.3: Using the SF-36 data via Gray algorithm: GO study only

Covariate	**Mean**	**Variance-Covariance matrix**			
		Intercept	HAQ	PASI	HAQ x PASI
Intercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HAQ	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PASI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HAQ x PASI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 6.4: Using the EQ-5D data: GO study only

Covariate	**Mean**	**Variance-Covariance matrix**			
		Intercept	HAQ	PASI	HAQ x PASI
Intercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HAQ	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PASI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HAQ x PASI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

The manufacturers incorporated these new estimates into a revised base-case analysis. In addition, in response to a request from the ERG, they removed vial sharing from the base-case. The revised results are presented in Table 6.5 below.

Table 6.5: MS revised base case results

Technologies	**Total costs (£)**	**Total QALYs**	**Incremental costs (£)**	**Incremental QALYs**	**ICER (£) versus Palliation (QALYs)**
Palliation	£62,224	6.61			
Adalimumab	£86,410	7.89	£24,186	1.28	£18,824
Golimumab	£94,151	8.21	£7,740	0.31	£19,993
Etanercept	£94,578	8.49	£428	0.29	£17,177
Infliximab	£106,620	8.49	£12,042	0.00	£23,578

In presenting these revised results the manufacturer has still incorrectly calculated the ICERs. They have failed again to account for extended dominance. The corrected ICERs are presented in Table 6.6 below.

Table 6.6: ERG correction of MS revised base case results

Technologies	**Total costs (£)**	**Total QALYs**	**Incremental costs (£)***	**Incremental QALYs***	**ICER (£) versus Palliation (QALYs)**	**ICER (£) incremental vs TNF-α inhibitors (QALYs)****
Palliation	£62,224	6.61				
Adalimumab	£86,410	7.89	£24,186	1.28	£18,824	E.D
Golimumab	£94,151	8.21	£7,740	0.31	£19,993	E.D
Etanercept	£94,578	8.49	£428	0.29	£17,177	£17,209
Infliximab	£106,620	8.49	£12,042	0.00	£23,578	Dominated

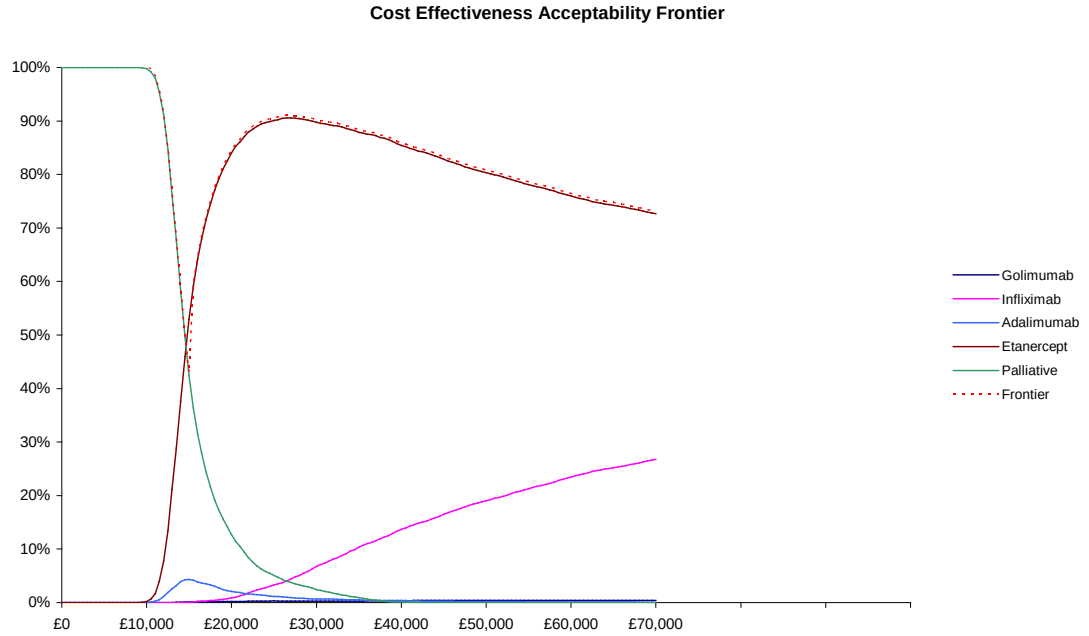
*The difference between the treatment and the lower ranked alternative, without considering extended dominance

**The difference between the treatment and the next best alternative, excluding extended dominated strategies

ED Extendedly dominated

Originally, the MS presented the cost effectiveness acceptability curves for each biologic relative to palliative care. The ERG felt that this was not an appropriate comparison (see Section 5.4) and requested a figure showing the probability that each strategy is the most cost effective compared with all the other strategies. The revised cost-effectiveness acceptability curve is presented below in Figure 6.1. The cost-effectiveness frontier is also presented in Figure 6.1, which shows the probability that the optimal strategy is cost effective at various values for the cost-effectiveness threshold. The probability that golimumab is cost-effective (according to the MS model) is almost zero at all values of the cost-effectiveness threshold, which is consistent with the result in the deterministic model that golimumab is extendedly dominated by etanercept.

Figure 6.1: Revised cost-effectiveness acceptability curve (TNF-α inhibitors compared to palliative care)



The ERG also asked the MS to carry out a sensitivity analysis that took account of NICE's decisions in the previous appraisal of anti-TNF agents for PsA. The issued guidance was based on the view that all the anti-TNF agents have similar effectiveness in terms of PASI, HAQ and PsARC response, and differed only in acquisition and administration cost. The MS conducted the additional analysis which assumes similar effectiveness in terms of PASI, HAQ and PsARC response for all four TNF-α inhibitors. The MS used identical values to those used in the previous MTA i.e. PsARC response of 0.713, HAQ change for responders to be -0.63 and HAQ change for non-responders to be -0.19. The MS model uses a different PASI calculation approach to the MTA and therefore we used the etanercept value for the absolute change from baseline in PASI score of -4.5278. This is in line with the previous approach where etanercept values were used to substitute for other TNF-α inhibitors. The results are presented in Table 6.7. The ICER for adalimumab, golimumab or etanercept versus palliative care is about £15,000 per QALY. Infliximab is dominated (i.e. more costly with no greater effectiveness than the other anti-TNF agents).

Table 6.7: MS model assuming equal clinical effectiveness for all anti-TNF agents

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) versus Palliation (QALYs)
Palliation	£62,224	6.61			
Adalimumab OR Golimumab	£92,877	8.59	£30,653	1.98	£15,494
Etanercept	£92,879	8.59	£2	0	£15,495
Infliximab	£104,401	8.59	£11,522	0	£21,319

6.2. Additional work undertaken by the ERG

6.2.1. Updating of previous ERG evidence synthesis model

The ERG updated the Rodgers et al.[7] Assessment Group evidence synthesis models to include golimumab through the use of the GO-REVEAL data. This has allowed some comparison of both the synthesis models used and the results obtained. A brief outline of the comparative modelling approaches is presented in Table 6.8.

Table 6.8: Summary of the alternative evidence synthesis approaches

	MS	**Rodgers et al.[7]**
Interventions	Etanercept, Infliximab, Adalimumab, Golimumab	Etanercept, Infliximab, Adalimumab, Golimumab
Studies used in the analysis	IMPACT, IMPACT 2, Mease 2000, Mease 2004, ADEPT, Genovese 2007, York HTA, GO-REVEAL.	IMPACT, IMPACT 2, Mease 2000, Mease 2004, ADEPT, Genovese 2007, GO-REVEAL.
Outcomes of interest		
PsARC	12 or 14 weeks data from all trials	12 or 14 weeks data from all trials
HAQ	HAQ at 12 and 24 weeks for Adalimumab/ 14 or 16 weeks for Infliximab/ 12 weeks for Etanercept and 14 weeks for Golimumab. (conditional on PsARC response).	HAQ at 12 or 14 weeks for all four anti-TNF agents (conditional on PsARC response at 12 weeks) Mease 2000 was not included due to the format of the results presented.
PASI 50/75/90	Estimated absolute PASI change from baseline. Incorporated data from week 24 for Adalimumab/ week 14 or 16 for Infliximab/ week 24 for Etanercept and week 14 for Golimumab	Estimated probability of achieving each of the three linked PASI outcomes. PASI 50/70/90 at 12 or 14 weeks (by biologic). Mease 2004 was not included due to a lack of 12 week data.
ACR 20/50/70	Not estimated.	ACR 20/50/70 at 12 or 14 weeks (by biologic)
Model	Two linked meta-analysis: estimating the change in HAQ from baseline conditional on PsARC response. Absolute change in PASI was modelled.	Fixed effect meta-analysis (PsARC, HAQ|PsARC, ordered logit model PASI/ACR)
Results Reported	Incremental HAQ change given PsARC response in treatment, Incremental HAQ change given PsARC non-response in treatment, Incremental HAQ change given PsARC response in placebo, Incremental HAQ change given PsARC non-response in placebo.	Probability of response in terms of PsARC, ACR and PASI. Changes in HAQ given PsARC response/non-response to treatment.

The results obtained by the Assessment Group are presented alongside the MS results in Table 6.10. In general the results are comparable with estimates varying by insignificant amounts and the rankings remaining generally the same. However, for the probability of achieving a PsARC response, the Assessment Group model ranked golimumab 2nd, [redacted]. For HAQ change in both responders and non-responders, the Assessment Group ranked infliximab 1st and etanercept 2nd, [redacted]. Table 6.9 shows the rankings of all anti-TNF agents across all outcomes for both the Assessment Group and MS models.

Table 6.9: Biologic comparative rankings from the Assessment Group & MS evidence synthesis

	Infliximab		**Etanercept**		**Adalimumab**		**Golimumab**	
	MS	**ERG**	**MS**	**ERG**	**MS**	**ERG**	**MS**	**ERG**
Probability of PsARC response	[redacted]	1ST	[redacted]	3RD	[redacted]	4TH	[redacted]	2ND
HAQ change responders	[redacted]	1ST	[redacted]	2ND	[redacted]	3RD	[redacted]	4TH
HAQ change non-responders	[redacted]	1ST	[redacted]	2ND	[redacted]	3RD	[redacted]	4TH
PASI change	[redacted]	1ST	[redacted]	4TH	[redacted]	2ND	[redacted]	3RD

Table 6.10: Comparative results from the Assessment Group & MS evidence synthesis

Outcomes	**Placebo**		**Infliximab**		**Etanercept**		**Adalimumab**		**Golimumab**	
	MS	ERG	MS	ERG	MS	ERG	MS	ERG	MS	ERG
PsARC response Mean (SD), 95% CrI	[redacted]	0.247 (0.036) [0.175, 0.318]	[redacted]	0.793 (0.057) [0.001, 0.799]	[redacted]	0.712 (0.070) [0.562, 0.832]	[redacted]	0.585 (0.070) [0.441, 0.716]	[redacted]	0.764 (0.065) [0.622, 0.871]
HAQ change from baseline, in PsARC responders Mean (SD), 95% CrI	[redacted]	-0.2663 (0.044) [-0.3555, -0.1816]	[redacted]	-0.659 (0.709) [-1.026, -0.286]	[redacted]	-0.635 (0.091) [-0.8144, -0.4563]	[redacted]	-0.4818 (0.065) [-0.6053, -0.3488]	[redacted]	-0.4404 (0.085) [-0.6088, -0.2756]
HAQ change from baseline, in PsARC non-responders Mean (SD), 95% CrI	[redacted]	0 [0,0]	[redacted]	-0.1981 (0.073) [-0.3382, 0.056]	[redacted]	-0.1949 (0.099) [-0.3917, 0.00023]	[redacted]	-0.136 (0.068) [-0.2684, 0.0017]	[redacted]	-0.0308 (0.088) [-0.2608, 0.1418]
PASI change from baseline, in patients ≥3% BSA psoriasis at baseline Mean (SD), 95% CrI	[redacted]	_	[redacted]	-7.2168	[redacted]	-2.5044	[redacted]	-5.17769	[redacted]	-4.486

6.2.2. Further exploration of the manufacturer's revised model by the ERG

In addition to updating of the Rodgers et al.[7] Assessment Group evidence synthesis (Section 6.2.1) and decision model (Section 6.2.3), the ERG has undertaken several sensitivity analysis using the manufacturer's model to address some of the other uncertainties identified in Section 5.2. These are:

1. Utilising the results from the ERG MTC.
2. Alternative estimates of cost of administration of drugs.
3. Alternative values for cost of PASI.
4. Alternative utility functions.
5. Alternative dosing strategy

6.2.2.1. Using the updated York Assessment Group MTC results

The ERG is aware of the limitations of the evidence base used to undertake the indirect treatment comparison (Section 5.2.3), and that the MS evidence synthesis has made a reasonable attempt to compare the effectiveness of the different therapies using the available data. Nevertheless, the ERG has noted a number of potential weaknesses in the MTC which are addressed in an updated Assessment Group MTC (Section 6.2.1). A sensitivity analysis utilising the results from the Assessment Group MTC within the manufacturers' model was therefore conducted. This incorporates the estimates of PsARC response, HAQ gain given response/no response and average change in PASI estimated by the Assessment Group MTC (see Section 6.2.1).

The results of this sensitivity analysis can be seen in Table 6.11 below. Using the updated Assessment Group evidence synthesis infliximab is no longer dominated by etanercept and has an ICER of £56,314 compared to etanercept. Adalimumab is still extendedly dominated, and golimumab is now dominated instead of extendedly dominated in the manufacturer's base case results (Table 6.5). The ICER for etanercept increases from £17,209, in the manufacturer's base case, to £20,614 using the Assessment Group updated evidence synthesis results.

Using the Assessment Group evidence synthesis produces fewer QALYs for all TNF-α inhibitor strategies. This is due to a combination of factors: the response rates are slightly lower for etanercept and golimumab; the HAQ gain for responders is lower for etanercept and adalimumab and the mean PASI change is lower for golimumab, infliximab and etanercept in the Assessment Group evidence synthesis compared to the manufacturer's estimates (see Section 6.2.1).

Table 6.11: Impact of using Assessment Group MTC within the manufacturers model

	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£)
Palliation	£62,224	6.61	-	-	-
Adalimumab	£88,181	7.83	£25,957	1.22	E.D
Golimumab 50 mg	£96,023	8.00	7,841	0.17	Dominated
Etanercept	£94,359	8.17	-£1,664	0.17	£20,614
Infliximab	£107,885	8.41	£13,526	0.24	£56,314

ED Extendedly dominated

6.2.2.2. Alternative estimates of cost of administration of drugs

The manufacturers' submission included a cost of staff nurse time for initial administration of etanercept, adalimumab and golimumab, in addition to the outpatient visit. This was justified on the basis that additional time would be required to train patients to self administer these drugs during the initial visit. The ERG, however, feels that the addition of this cost, whilst patients are already attending an outpatient visit, is likely to be double counting NHS staff time. The 4 hours of nurse staff time required for the first administration of etanercept, adalimumab and golimumab is, therefore, excluded in a sensitivity analysis. The results of this sensitivity analysis can be seen in Table 6.12 below. Excluding the additional hours of nurse staff time for etanercept, adalimumab and golimumab has little impact on the results, the ICER for etanercept compared to palliative care is broadly similar to the MS base case results at £17,019.

Table 6.12: Impact of excluding additional hours of staff nurse time for initial administration

	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£)
Palliation	£62,224	6.60	-	-	-
Adalimumab	£86,222	7.89	£23,998	1.28	ED
Golimumab	£93,962	8.20	£7,740	0.31	ED
Infliximab	£106,620	8.49	£12,657	0.29	Dominated
Etanercept	£94,390	8.49	-£12,229	0.0007	£17,019

ED Extendedly dominated

6.2.2.3. Alternative values for the cost of PASI

In the Rodgers et al.[7] Assessment Group model the cost assigned according to the degree of psoriasis, measured using PASI scores, proved to be a significant factor in determining the cost-effectiveness of TNF-α inhibitors - in particular, the comparison between etanercept and infliximab, where infliximab offers greater benefits in terms of PASI change.

The costs associated with PASI are implemented in the manufacturer's model by applying a cost per PASI point; estimated using results from a survey of dermatologists (see Section 5.1.5). An additional [redacted] was estimated including phototherapy costs and an [redacted] excluding these costs. Requests for further clarification on how these average costs were determined from the raw survey results were not fully satisfied, in particular in determining the unit costs associated with each of the resource use items relating to psoriasis (see Section 6.1). Thus it was not possible to examine the extent to which these costs were appropriate and valid. It was also not possible to use the costs associated with PASI directly from the Rodgers et al.[7] Assessment Group model as these were not specified according to costs per PASI point. As an alternative, a sensitivity analysis varying the manufacturer's PASI costs was conducted. This ranged the costs associated with PASI from half of that estimated by the manufacturer to double.

The results of this sensitivity analysis can be seen in Table 6.13 below. Decreasing or increasing the costs per PASI point does not change the relative ordering of strategies compared to the base case (Table 6.13). Adalimumab and golimumab are still extendedly dominated and infliximab is dominanted by etanercept. The ICER for etanercept increases from £1,491 in the manufacturer's base case, to £17,744 when halving PASI costs, and to £16,609 when doubling PASI costs.

Varying the costs associated with PASI does not change the estimates of total QALYs but does decrease costs for all strategies were PASI costs halve and increase costs for all strategies where PASI double. However, all other strategies remain either dominated or extendedly dominated and the comparison is between etanercept and palliative care.

Table 6.13: Impact of using York costs associated with PASI

	Total costs (£)	**Total QALYs**	**Incremental costs (£)**	**Incremental QALYs**	**ICER (£)**
Halving the cost per PASI point					
Palliation	£51,250	6.61	-	-	-
Adalimumab	£76,394	7.89	£25,144	1.28	ED
Golimumab 50 mg	£84,346	8.21	£7,952	0.31	ED
Infliximab	£97,270	8.49	£12,923	0.29	Dominated
Etanercept	£84,673	8.49	-£12,597	0.00	£17,744
Doubling the cost per PASI point					
Palliation	£73,199	6.61	-	-	-
Adalimumab	£96,427	7.89	£23,228	1.28	ED
Golimumab 50 mg	£103,955	8.21	£7,528	0.31	ED
Infliximab	£115,971	8.49	£12,016	0.29	Dominated
Etanercept	£104,484	8.49	-£11,486	0.00	£16,609

ED Extendedly dominated

6.2.2.4. Alternative utility function

The manufacturer's model utilised a revised version of the Rodgers et al.[7] Assessment Group utility function, including an additional explanatory variable, HAQ * PASI. There is some uncertainty regarding which of the regression models is appropriate to generate utilities. In addition to a request for further analyses from the manufacturer using alternative functional forms (see Section 6.1) the ERG has undertaken additional analyses using the original estimates of utility values from Rodgers et al.[7]

The results of this sensitivity analysis can be seen in Table 6.14 below. Utilizing the original Assessment Group utility function has little impact on the results, reducing the ICER slightly for etanercept compared to palliative care from £17,209 in the base case to £14,444.

Table 6.14: Impact of using York Assessment Group utility function

	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£)
Palliation	£62,224	5.45	-	-	-
Adalimumab	£86,410	6.98	£24,186	1.529	ED
Golimumab 50 mg	£94,151	7.35	£7,740	0.370	ED
Infliximab	£106,620	7.69	£12,469	0.344	Dominated
Etanercept	£94,578	7.69	-£12,042	0.003	£14,444

ED Extendedly dominated

6.2.2.5. Alternative dosing strategy for patients not achieving PsARC response at 3 months

The base-case MS model measures PsARC response rates for golimumab 50 mg after three months, and the model assumes that all patients who do not achieve PsARC at this time are withdrawn from biologic therapy to palliative care. However, the licence for golimumab allows the clinician to increase the dose to 100 mg for patients weighing more than 100 kg who fail to achieve an adequate response after 12-14 weeks of treatment. This might be reflected in the results of the open label phase of the GO-REVEAL trial. The study found that 28/146 (19%) of the patients in the 50 mg treatment arm titrated up to 100 mg after week 16, although the study does not give details of the weight or initial responses of these patients. If this strategy were repeated in clinical practice this would imply an increased cost for golimumab. The ERG does not have the data to conduct a full economic evaluation of this alternative strategy, but simply notes that patients who are maintained on an increased dose from 50 mg to 100 mg would cost an additional £2145 per 3 months. However, the clinical advisor to the ERG has stated that, in practice, few patients would be titrated up to a higher dose to achieve response because it would be less costly to switch the patient to an alternative biologic (a strategy also not modelled by the MS).

6.2.3. Alternative ERG model structure

The ERG was the Assessment Group for the recent Multiple Technology Appraisal of etanercept, infliximab and adalimumab for psoriatic arthritis. During that appraisal, the Assessment Group constructed a new decision model to compare the cost-effectiveness of these anti-TNF agents against each other and palliative care. In this section, the Assessment Group model is used to estimate the cost-effectiveness of etanercept, infliximab, adalimumab and golimumab. The aim of these analyses is to validate the MS model.

As described in Section 5.1, the structure of the Assessment Group model is broadly similar to the MS model. The main structural difference is in the way that PASI changes are estimated, though, as psoriasis does not contribute to major differences between the therapies in costs or HRQoL in either model in this patient group, this difference in model structure should not be of great importance. The Assessment Group model can only consider a homogeneous patient cohort, whereas the MS model can evaluate treatments for a mix of psoriatic arthritis patients with and without psoriasis. Therefore the Assessment Group model should be compared with the MS model for patients with significant psoriasis.

6.2.3.1. Methods of Assessment Group alternative model

Three analyses are undertaken. First, results are shown using the Assessment Group model when data predominately from the MS are used as inputs. This analysis aims to validate the structure of the MS model. In this validation analysis, it was not possible to use the MS estimates of the change in PASI for each treatment because the structure of the Assessment Group model does not permit these inputs in this format. The Assessment Group estimates of the PASI 50, 75 and 90 responses were used instead. Second, to validate the data used in the model, the estimates of PsARC and HAQ responses were replaced with those estimated by the updated Assessment Group evidence synthesis (Tables 6.10). Third, sensitivity analysis is carried out by varying the cost of treating psoriasis in the model. The calculation of this parameter was a source of uncertainty in the MS model and the ERG therefore investigated the impact of this parameter on the cost-effectiveness results generated.

6.2.3.2. Results of Assessment Group alternative model

Table 6.15 shows the results of using the Assessment Group model populated with the MS base-case data. The incremental cost-effectiveness ratio of etanercept versus palliative care is about £23,000 per QALY, indicating that etanercept is on the margin of being cost-effective, given these inputs to the model. The ICER of infliximab versus etanercept is over £280,000 per QALY. Adalimumab and golimumab are extendedly dominated by etanercept.

These results are can be compared with the results of the MS analysis for rheumatic patients with significant psoriasis (when the ICERs are correctly calculated in the MS) (Table 6.13). Both models find that adalimumab and golimumab are extendedly dominated, and both find that the ICER for infliximab versus etanercept is very high. The MS model finds that the ICER for etanercept versus palliative care is £14,000 per QALY. This is lower than the Assessment Group model estimate in Table 6.13 (£23,000 per QALY). This is because the Assessment Group meta-analysis and model estimates a lower improvement in PASI for etanercept than the MS meta-analysis and model.

Table 6.16 shows the results of the same model, but using ERG estimates of PsARC, HAQ and PASI response rates. Other data are the same as the analysis in Table 6.15. As in the previous analysis, adalimumab and golimumab are extendedly dominated. In Table 6.16, the ICER for infliximab falls to about £51,000 per QALY. This is because the ERG meta-analysis estimates a greater response in terms of PsARC and HAQ for infliximab than the MS meta-analysis.

Table 6.15: Results of Assessment Group model populated with MS data

Strategy	QALY	Lifetime Cost	Difference in QALYs	Difference in costs	Incremental QALY	Incremental cost	ICER
Palliative care	6.13	64094					
Adalimumab	7.13	89149	0.99	25055			E.Dominated*
Golimumab	7.38	97308	0.25	8159			E.Dominated*
Etanercept	7.64	98361	0.26	1053	1.51	34267	22761
Infliximab	7.68	108680	0.04	10319	0.04	10319	270174

*Extendedly dominated

Table 6.16: Results of Assessment Group model populated with ERG data on PsARC, HAQ and PASI response rates

Strategy	QALY	Lifetime Cost	Difference in QALYs	Difference in costs	Incremental QALY	Incremental cost	ICER
No treat		64094					
Adalimumab		90604	1.1	26510			
Golimumab		98422	0.28	7818			
Etanercept		96830	0.06	-1592	1.44	32736	22775
Infliximab		109792	0.25	12962	0.25	12962	50951

Table 6.17 shows the results of sensitivity analyses where the cost per PASI point per year is varied. The results of Tables 6.15 and 6.16 are based on a cost per PASI point per year of £167, assuming phototherapy is used to treat these patients, as in the MS model. This implies that reducing PASI from, say, 9.9 to 3.3 (a reduction of 6.6 points estimated for infliximab) would reduce the expected cost of treating psoriasis per year by £1,100. The sensitivity analyses in Table 6.17 varies this cost by

doubling the saving in psoriasis therapy to £2,200 per year and halving it to £550 per year. Neither scenario substantially changed the conclusions of the analysis.

Table 6.17: Sensitivity analyses varying the cost of treating psoriasis in the ERG model

Incremental cost-effectiveness ratios	Palliative care	Adalimumab	Golimumab	Etanercept	Infliximab
Base-case (Table 6.14)	-	Extendedly dominated	Extendedly dominated	£22761 per QALY	£270174 per QALY
Doubling cost of treating psoriasis	-	Extendedly dominated	Extendedly dominated	£22239 per QALY	£203686 per QALY
Halving the cost of treating psoriasis	-	Extendedly dominated	Extendedly dominated	£23036 per QALY	£305037 per QALY

6.2.3.3. Conclusions

Further analyses were conducted using the MS model and the ERG model previously developed by York Assessment Group during the recent appraisal of etanercept, infliximab and adalimumab. The MS model and the ERG alternative model have a broadly similar structure and data and give similar results, indicating golimumab is extendedly dominated by etanercept. Sensitivity analyses did not change these conclusions.

7. DISCUSSION

7.1. Summary of clinical effectiveness issues

The data from GO-REVEAL trial[4, 5] provides evidence to suggest that golimumab appears to be an efficacious treatment for active and progressive PsA patients who have not achieved an adequate response with DMARDs or NSAIDs. The effect sizes of point estimates of joint and skin disease response and functional status were moderate to large, implying that these treatment effects could be clinically significant. However, the analyses for efficacy outcome were limited to only one RCT (GO-REVEAL[4, 5]) with limited sample size. In particular, few patients provided data on the psoriasis response to golimumab treatment.

There was a lack of robustness for the data analyses in the GO-REVEAL trial.[4, 5]Based on the data further provided by the manufacturer, it appears that the analyses of PASI 50 and PASI 90 at 14 weeks and all the PASI outcomes at 24 weeks were not performed on the basis of intention-to-treat analysis, thereby compromising the internal validity of these skin disease outcomes. Furthermore, whilst the analyses at 24 weeks involved all the intention-to-treat data from the randomisation, it appears that such analyses failed to adjust for the treatment contamination due to patients' crossing-over at week 16. Therefore, this may have threatened the internal validity of results for all the efficacy and safety outcomes at 24 weeks.

The radiographic outcomes in the GO-REVEAL trial[4, 5]were evaluated over the short follow-up period of 24 weeks, which were often considered inadequate to assess radiographic changes in response to the treatment. There was a lack of long-term efficacy data of radiographic assessments. Given the fact that the treatment effect on the joint disease is more accurately reflected by the more objective radiographic measure, radiographic long-term data could be valuable to provide more generalisable estimates of the treatment effect in responding to golimumab therapy.

In terms of the results of manufacturer's MTC analyses, the credible intervals of most outcomes for all four anti-TNF agents overlapped between each other. In particular, there were substantial uncertainties for these estimates of PASI change from baseline due to a small sample size of patients evaluable for psoriasis. Furthermore, no comparisons of anti-TNF agents in the MTC were performed for the treatment duration beyond 14 weeks, since only 12-14 week data from the included trials were

used to establish the relative efficacy between these anti-TNF agents. Additionally, there were also no comparisons for disease progression (radiographic data) in the MTC analyses.

There was a concern of the generalisability of treatment effects observed in these trial participants in MTC to those treated in routine clinical practice. Despite most patients in the included trials of MTC being under licensed conditions, the majority of patients had previously received at least one DMARD and no trial specified the failure to respond to at least two DMARDs (patients whom the current BSR guidelines and NICE guidance for etanercept, infliximab and adalimumab consider eligible for the biologic treatment) as a recruitment criterion. Thus, it remains unclear that the beneficial effects observed in these trial participants were similar in those treated in routine clinical practice.

The ERG further considered the evidence for the safety evaluation of golimumab to be inadequate. The evidence for the safety evaluation of golimumab was exclusively based on 14 and 24 week data from a single RCT with PsA patients (GO-REVEAL[4, 5]). The manufacturer failed to provide longer term data or to consider adverse event data of golimumab from controlled studies in other conditions such as rheumatoid arthritis and ankylosing spondylitis. Whilst the adverse effects profile of golimumab appears similar to other anti-TNF agents, the MS did not include a detailed comparison of the adverse effects profiles of the 4 anti-TNF agents. The longer-term safety profile of golimumab remained uncertain. Given these limitations and uncertainties, the manufacturer's conclusion that golimumab is a safe treatment option similar to other anti-TNF agents may be premature and may not be reliable.

7.2. Summary of cost effectiveness issues

The MS decision model took account of all the important elements of the decision problem, in terms of the rules for the continuation of biologic therapy, the natural history of arthritis and psoriasis in these patients, the treatment effects, the relationship between psoriasis, arthritis and HRQoL and the costs.

The MS report incorrectly calculated the ICERs (incremental costs-per-QALY) of each biologic relative to the next best alternative. When the ICERs are correctly calculated using the MS model, the ICER of etanercept is about £17,000 per QALY versus palliative care. Other anti-TNF agents are extendedly dominated or dominated by etanercept. Sensitivity analyses did not change these conclusions.

A key area of uncertainty is whether the anti-TNF agents should be considered equally clinically effective, that is, to treat them as a class. This was the position adopted in the guidance issued by NICE for the previous appraisal of etanercept, adalimumab and infliximab for psoriatic arthritis.[2] If all anti-TNF agents are considered equally effective (in terms of PsARC, HAQ and PASI responses), then etanercept, adalimumab and golimumab have very nearly equal costs and equal QALYs and all have an ICER of about £15,000 per QALY versus palliative care. Infliximab has higher acquisition costs and is dominated by other biologic strategies if they are considered equally effective.

The licence for golimumab indicates that patients over 100 kg in weight who fail to respond to golimumab 50 mg at 3 months can be trialled on a higher dose of 100 mg. A full economic analysis of this option could not be undertaken because of lack of data. The ERG notes that, if patients are titrated and maintained on a higher dose, the additional acquisition costs will be around £2,145 per 3 months. However, the clinical adviser to the ERG suggests that, in practice, this scenario is unlikely because of the additional cost and eligible patients are more likely to be tried on an alternative biologic.

A remaining source of uncertainty is the annual cost of treating psoriasis. Although the MS conducted a survey of dermatologists and presented the raw data from the survey, there was no detail of the statistical method used to calculate the mean costs from the raw data and, therefore, the ERG could not validate the calculations. However, the ERG conducted some sensitivity analysis on the PASI cost using the ERG model. Doubling or halving the cost per PASI point of £167 did not materially affect the results of the ERG model, indicating that this is not a key parameter for the decision, at least in patients who do not have severe psoriasis.

7.3. Implications for research

- Long-term observational studies of golimumab with large sample sizes of patients with PsA are required to demonstrate that beneficial effects for joint and skin disease and improvement of function are maintained.
- Further monitoring of the safety profile of golimumab (e.g. through the BSR Biologics Register) is required. Future studies should also establish whether long-term patterns of adverse events of golimumab are similar to other anti-TNF agents.

8. References

1. Mease P, Ganguly R, Wanke L, Yu E, Singh A. How much improvement in functional status is considered important by patients with active psoriatic arthritis: applying the outcome measures in rheumatoid arthritis clinical trials (OMERACT) group guidelines. *Ann Rheum Dis* 2004;63 Suppl 1:391.

2. National Institute for Health and Clinical Excellence. *Final appraisal determination. Etanercept, infliximab and adalimumab for the treatment of psoriatic arthritis (review of technology appraisal guidance 104 and 125).* London: NICE; 2010 May. Available from: http://www.nice.org.uk/nicemedia/live/11966/49113/49113.pdf

3. Kyle S, Chandler D, Griffiths CEM, Helliwell P, Lewis J, McInnes I, et al. Guideline for anti-TNF-a therapy in psoriatic arthritis. *Rheumatology (Oxford).* 2005;44:390-7.

4. Centocor Ltd. *A multicenter, randomized, double-blind, placebo-controlled trial of golimumab, a fully human anti-TNFα monoclonal antibody, administered subcutaneously in subjects with active psoriatic arthritis*. Malvern, PA: Centocor; 2007 16 Oct.

5. Kavanaugh A, McInnes I, Mease P, Krueger GG, Gladman D, Gomez-Reino J, et al. Golimumab, a new human tumor necrosis factor antibody, administered every four weeks as a subcutaneous injection in psoriatic arthritis twenty-four–week efficacy and safety results of a randomized, placebo-controlled study. *Arthritis Rheum* 2009;60:976-86.

6. *Summary of Product Characteristics- Simponi 50 mg solution for injection in pre-filled pen (Centocor B.V.).* 2010. [cited 2010 Jul 7]. Available from: http://www.ema.europa.eu/humandocs/PDFs/EPAR/simponi/emea-combined-h992en.pdf.

7. Rodgers M, Epstein D, Bojke L, Yang H, Craig D, Fonseca T, et al. *Etanercept, infliximab and adalimumab for the treatment of psoriatic arthritis: a systematic review and economic evaluation*. London: National Institute for Clinical Excellence; 2009. Available from: http://www.nice.org.uk/nicemedia/live/11966/46713/46713.pdf

8. Mease PJ, Goffe BS, Metz J, VanderStoep A, Finck B, Burge DJ. Etanercept in the treatment of psoriatic arthritis and psoriasis: a randomised trial. *Lancet* 2000;356:385-90.

9. Mease PJ, Kivitz AJ, Burch FX, Siegel EL, Cohen SB, Ory P, et al. Etanercept treatment of psoriatic arthritis: safety, efficacy, and effect on disease progression. *Arthritis Rheum* 2004;50:2264-72.

10. Antoni CE, Kavanaugh A, Kirkham B, Tutuncu Z, Burmester GR, Schneider U, 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;52:1227-36.

11. Antoni C, Krueger GG, de Vlam K, Birbara C, Beutler A, Guzzo C, et al. Infliximab improves signs and symptoms of psoriatic arthritis: results of the IMPACT 2 trial. *Ann Rheum Dis* 2005;64:1150-7.

12. Mease PJ, Gladman DD, Ritchlin CT, Ruderman EM, Steinfeld SD, Choy EH, 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;52:3279-89.

13. Genovese MC, Mease PJ, Thomson GT, Kivitz AJ, Perdok RJ, Weinberg MA, et al. Safety and efficacy of adalimumab in treatment of patients with psoriatic arthritis who had failed disease modifying antirheumatic drug therapy. *J Rheumatol* 2007;34:1040-50.

14. Sterry W, Ortonne J-P, Kirkham B, Brocq O, Robertson D, Pedersen RD, et al. Comparison of two etanercept regimens for treatment of psoriasis and psoriatic arthritis: PRESTA randomised double blind multicentre trial. *BMJ* 2010;340:c147

15. Smith CH, Anstey AV, Barker JN, Burden AD, Chalmers RJ, Chandler DA, et al. British Association of Dermatologists' guidelines for biologic interventions for psoriasis 2009. *Br J Dermatol* 2009;161:987-1019.

16. Sutton A, Ades AE, Cooper N, Abrams K. Use of indirect and mixed treatment comparisons for technology assessment. *Pharmacoeconomics* 2008;29:753-67.

17. Woolacott N, Bravo Vergel Y, Hawkins N, Kainth A, Khadjesari Z, Misso K, et al. Etanercept and infliximab for the treatment of psoriatic arthritis: a systematic review and economic evaluation. *Health Technol Assess* 2006;10(31):1-239.

18. Roberts C, Torgerson DJ. Understanding controlled trials: baseline imbalance in randomised controlled trials. *BMJ* 1999;319:195.

19. Dumville JC, Torgerson DJ, Hewitt CE. Reporting attrition in randomised controlled trials. *BMJ* 2006;332:969-71.

20. Song F, Altman DG, Glenny A-M, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ* 2003;326:472

21. Schering-Plough Ltd. *Remicade in the treatment of psoriatic arthritis in the United Kingdom: a submission to the National Institute for Clinical Excellence [industry submission]*. Kenilworth, NJ: Schering-Plough Ltd; 2004.

22. Abbott. *Adalimumab (HUMIRA): Multiple technology appraisal of adalimumab, etanercept and infliximab for psoriatic arthritis. National Institute for Health and Clinical Excellence (NICE) Health Technology Appraisal*: Abbott Laboratories Limited; 2009.

23. Bansback NJ, Ara R, Barkham N, Brennan A, Fraser AD, Conway P, et al. Estimating the cost and health status consequences of treatment with TNF antagonists in patients with psoriatic arthritis. *Rheumatology* 2006;45:1029-38.

24. Kobelt G, Jonsson L, Lindgren P, Young A, K. E. Modelling the progression of rheumatoid arthritis: a two-country model to estimate costs to estimate costs and consequences of rheumatoid arthritis. *Arthritis Rheum* 2002;46:2310-19.

25. Scott DL, Pugner K, Kaarela K, Doyle DV, Woolf A, Holmes J, et al. The links between joint damage and disability in rheumatoid arthritis. *Rheumatology (Oxford).* 2000;39:122-32.

26. Wong K, Gladman DD, Husted J, Long JA, Farewell VT. Mortality studies in psoriatic arthritis: results from a single outpatient clinic. 1. Causes and risk of death *Arthritis Rheum* 1997;40:1868-72.

27. Joint Formulary Committee. *British National Formulary*. London: Pharmaceutical Press; 2009.

28. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004;8(36):1-158

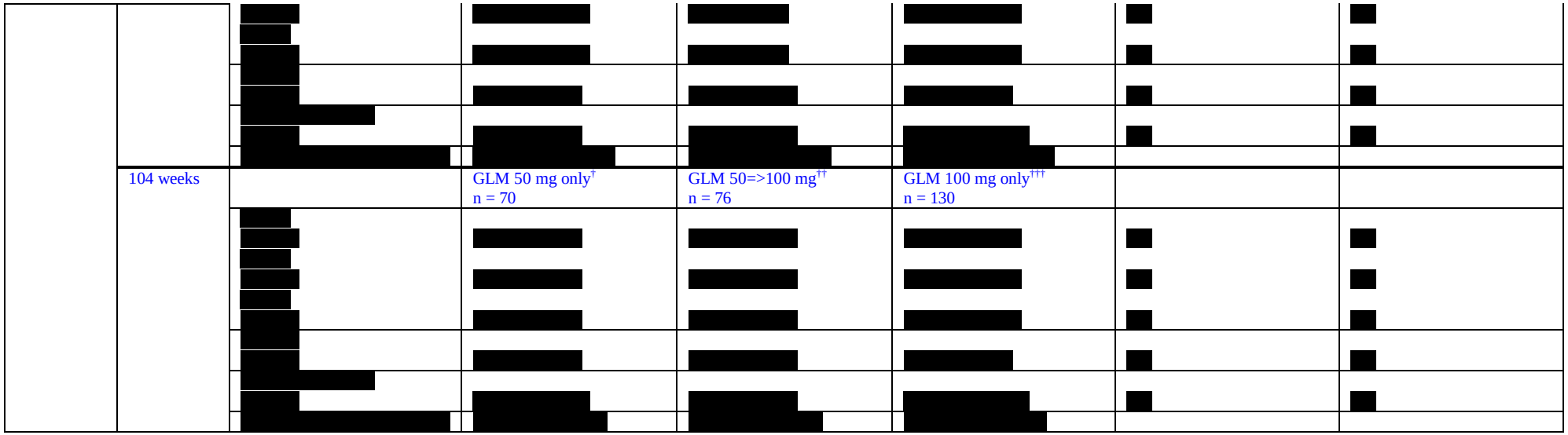
29. Gray AM, Rivero-Arias O, Clarke PM. Estimating the association between SF-12 responses and EQ-5D utility values by response mapping. *Med Decis Making* 2006;26:18-29.

30. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. *Med Care* 2004;42:851-9.

9. Appendices

Appendix 1: Study results for golimumab RCT of GO-REVEAL trial (The manufacturer's Point clarification A3)

Trial	Duration	Outcomes	Placebo	Golimumab 50 mg	Golimumab 100 mg	RR or mean difference (95% CI)	
						50 mg	100 mg
GO REVEAL[4, 5]	14 weeks	PsARC	24/113 (21.2%)	107/146 (73.3%)	105/146 (71.9%)	3.451 (2.49 - 4.87)	3.386 (2.43 - 4.80)
		ACR 20 All pts +MTX -MTX	 10/113 (8.8%) 8/55 (14.5%) 2/58 (3.4%)	 74/146 (50.7%) 38/71 (53.5%) 36/75 (48.0%)	 66/146 (45.2%) 32/71 (45.1%) 34/75 (45.3%)	 5.727 (3.24 - 10.56) 3.680 (1.98 – 7.25) 13.920 (4.13 – 51.64)	 5.108 (2.86 - 9.48) 3.099 (1.63 – 6.22) 13.147 (3.88 – 48.88)
		ACR 50 All pts	 2/113 (1.8%)	 44/146 (30.1%)	 41/146 (28.1%)	 17.027 (4.81 – 63.32)	 15.866 (4.47 – 59.11)
		ACR 70 All pts	 1/113 (0.9%)	 18/146 (12.3%)	 25/146 (17.1%)	 13.932 (2.46 – 81.82)	 19.349 (3.48 – 112.44)
		HAQ change from baseline (mean (SD))	N/A	N/A	N/A		
		PASI 50 All pts	 7/73 (9.6%)	 63/106 (59.4%)	 83/107 (77.6%)	 6.198 (3.22 – 12.7)	 8.089 (4.38 – 16.04)
		PASI 75 All pts	 2/79 (2.5%)	 44/109 (40.4%)	 63/108 (58.3%)	 15.945 (4.62 – 59.11)	 23.042 (6.85 – 84.59)
		PASI 90 All pts	 0/73 (0.0%)	 22/106 (20.8%)	 26/107 (24.3%)	 Inf (4.21 – Inf)	 Inf (4.95 – Inf)
	24 weeks[‡]		n = 113	n = 146	n = 146		
		PsARC	33/113 (29.2%)	102/146 (69.9%)	124/146 (84.9%)	2.392 (1.81 – 3.20)	2.908 (2.28 – 3.68)
		ACR 20 All pts	 14/113 (12.4%)	 76/146 (52.1%)	 89/146 (61.0%)	 4.202 (2.60 – 7.03)	 4.920 (3.09 – 8.13)
		ACR 50 All pts	 4/113 (3.5%)	 47/146 (32.2%)	 55/146 (37.7%)	 9.094 (3.62 – 23.94)	 10.642 (4.27 – 27.85)
		ACR 70 All pts	 1/113 (0.9%)	 27/146 (18.5%)	 31/146 (21.2%)	 20.897 (3.77 – 121.19)	 23.993 (4.35 – 138.68)
		HAQ change from baseline (mean (SD))	- 0.01 ± 0.49	0.33 ± 0.55 p < 0.001	0.39 ± 0.50 p < 0.001		
		PASI 50 All pts	 6/73 (8.2%)	 77/102 (75.5%)	 87/106 (82.1%)	 9.185 (4.69 – 19.45)	 9.986 (5.21 – 20.76)
		PASI 75 All pts	 1/73 (1.4%)	 57/102 (55.9%)	 70/106 (66.0%)	 40.794 (7.86 – 232.88)	 48.208 (9.44 – 274.39)
		PASI 90 All pts	 0/73 (0.0%)	 33/102 (32.4%)	 34/106 (32.1%)	 Inf (6.65 – Inf)	 Inf (6.59 – Inf)
	52 weeks		GLM 50 mg only[†] n = 102	GLM 50=>100 mg[††] n = 26	GLM 100 mg only[†††] n = 115		
		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]



‡At wk24 all pts randomised to the respective treatment arm are included; †Includes patients randomised to golimumab 50 mg and did not change dose; ††Includes patients on placebo at baseline who entered early escape or crossed over to golimumab 50 mg and later dose escalated to 100 mg and patients randomised to golimumab 50 mg who entered early escape or dose escalated to golimumab 100 mg; †††Includes patients randomised to golimumab 100 mg and did not change dose.

Appendix 2: Results of mixed treatment comparison in the MS and revised analyses in point clarifications

Outcomes	Placebo		Infliximab		Etanercept	
	MS base case	MS (revised inputs for Mease2005)	MS base case	MS (revised inputs for Mease2005	MS base case	MS (revised inputs for M
PsARC response Mean (SD), 95% CrI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HAQ change from baseline, in PsARC responders Mean (SD), 95% CrI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HAQ change from baseline, in PsARC non-responders Mean (SD), 95% CrI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

PASI change from baseline, in patients ≥3% BSA psoriasis at baseline	**Placebo**			**Infliximab**			**Etanercept**		
Outcomes	MS base case	MS (revised inputs for Mease2004)		MS base case	MS (revised inputs for Mease2005		MS base case	MS (revised inputs for Mease2005	
Mean (SD), 95% CrI									
PsARC response									
Mean (SD), 95%									

In the base case MS, the Mease 2005 data incorporated was HAQ at 24 weeks. Further to a request the manufacturer revised the analysis using Mease 12 week data. The results presented demonstrate that this revision has a marginal effect on the estimates, suggesting that the use of 24 week outcomes did not different the results in a significant manner.

CrI						
HAQ change from baseline, in PsARC responders Mean (SD), 95% CrI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HAQ change from baseline, in PsARC non-responders Mean (SD), 95% CrI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PASI change from baseline, in patients ≥3% BSA psoriasis at baseline	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Mean (SD), 95% CrI							
Outcomes	**Placebo**			**Infliximab**			**Etanercept**

In the base case MS, the Mease 2004 data incorporated was PASI at 24 weeks. Further to a request the manufacturer revised the analysis removing the 24 week data. The lack of availability of 12 week data meant that the Mease 2004 study was not included in the scenario. The results presented demonstrate that this revision has an effect on the estimates, although it is not clear the impact that the difference observed would have on the results of the decision model.

	MS base case	MS (revised for GO-REVEAL 2009)	MS base case	MS (revised for GO-REVEAL 2009)	MS base case	MS (revised for GO-R
PsARC response Mean (SD), 95% CrI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HAQ change from baseline, in PsARC responders Mean (SD), 95% CrI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HAQ change from baseline, in PsARC non-responders Mean (SD), 95% CrI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

PASI change from baseline, in patients ≥3% BSA psoriasis at baseline Mean (SD), 95% CrI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

In clarifications the manufacturer also supplied what they refer to as 'revised for GO-REVEAL 2009' results. It is not clear what this reanalysis included and, as demonstrated in the table above, there is no real difference between the revised results and the base case.

Appendix 3: Additional adverse effects data from the GO-REVEAL trial submitted in clarification letter (Summary of serious adverse effects, serious infection, tuberculosis, and adverse effects leading to discontinuation)

Adverse event	**Placebo**	**Placebo => GLM 50 mg**	**GLM 50 mg**	**GLM 50 mg => GLM 100 mg**	**GLM 100 mg**	**Combined 50 mg & 100 mg**	**All GLM**
Week 16	n = 113		n = 146		n = 146	n = 292	
Patients with ≥ 1 AE	63/113 (55.8%)		85/146 (58.2%)		82/146 (56.2%)	167/292 (57.2%)	
			1.044 (0.85-1.30)		1.007 (0.81-1.26)	1.026 (0.86-1.26)	
+MTX	27/54 (50.0%)		36/71 (50.7%)		39/69 (56.5%)	75/140 (53.6%)	
			1.521 (1.04-2.22)		1.696 (1.18-2.43)	1.607 (1.16-2.28)	
-MTX	36/59 (61.0%)		49/75 (65.3%)		43/77 (55.8%)	92/152 (60.5%)	
			1.071 (0.83-1.39)		0.915 (0.70-1.22)	0.992 (0.80-1.29)	
Patients with ≥ 1 serious AE	6/113 (5.3%)		3/146 (2.1%)		2/146 (1.4%)	5/292 (1.7%)	
			0.387 (0.11-1.39)		0.258 (0.06-1.10)	0.322 (0.11-0.98)	
+MTX	1/54 (1.9%)		2/71 (2.8%)		1/69 (1.4%)	3/140 (2.1%)	
			1.521 (0.20-11.56)		0.771 (0.08-7.36)	1.133 (0.17-7.88)	
-MTX	5/59 (8.5%)		1/75 (1.3%)		1/77 (1.3%)	2/152 (1.3%)	
			0.157 (0.02-0.98)		0.153 (0.02-0.96)	0.155 (0.04-0.68)	
Patients with ≥ 1 serious infections	3/113 (2.7%)		1/146 (0.7%)		0/146 (0.0%)	1/292 (0.3%)	
			0.258 (0.04-1.78)		0.000 (0.00-0.98)	0.129 (0.02-0.89)	
+MTX	0/54 (0.0%)		1/71 (1.4%)		0/69 (0.0%)	1/140 (0.7%)	
			Inf (0.2 – Inf)		-	Inf (0.10 – Inf)	
-MTX	3/59 (5.1%)		0/75 (0.0%)		0/77 (0.0%)	0/142 (0.0%)	
			0.000 (0.00-1.04)		0.000 (0.00-1.01)	0.000 (0.00-0.55)	
AE leading to discontinuation	4/113 (3.5%)		2/146 (1.4%)		2/146 (1.4%)	4/292 (1.4%)	
			0.387 (0.08-1.78)		0.387 (0.08-1.78)	0.387 (0.11-1.40)	
+MTX	0/54 (0.0%)		2/71 (2.8%)		0/69 (0.0%)	2/140 (1.4%)	
			-		-	-	
-MTX	4/59 (6.8%)		0/75 (0.0%)		2/77 (2.6%)	2/152 (1.3%)	
			0.000 (0.00-0.73)		0.383 (0.08-1.74)	0.194 (0.04-0.89)	
Subjects with ≥ 1 injection site reactions	3/113 (2.7%)		5/146 (3.4%)		5/146 (3.4%)	10/292 (3.4%)	
			1.290 (0.35-4.84)		1.290 (0.35-4.84)	1.290 (0.39-4.32)	

Subjects with tuberculosis	0/113 (0.0%)		0/146 (0.0%)		0/146 (0.0%)	0/292 (0.0%)	
Adverse event	**Placebo**	**Placebo => GLM 50 mg**	**GLM 50 mg**	**GLM 50 mg => GLM 100 mg**	**GLM 100 mg**	**Combined 50 mg & 100 mg**	**All GLM**
Week 24	n = 113	n = 51	n = 146	n = 28	n = 146	n = 292	n = 343
Patients with ≥ 1 AE	67/113 (59.3%)	26/51 (51.0%)	99/146 (67.8%)	4/28 (14.3%)	95/146 (65.1%)	196/292 (67.1%)	222/343 (64.7%)
			1.144 (0.95-1.38)		1.097 (0.91-1.33)	1.132 (0.96-1.36)	1.092 (0.93-1.31)
+MTX	30/54 (55.6%)	10/25 (40.0%)	45/71 (63.4%)	1/14 (7.1%)	44/69 (63.8%)	89/140 (63.6%)	99/165 (60.0%)
			1.141 (0.86-1.54)		1.148 (0.86-1.55)	1.144 (0.90-1.53)	1.08 (0.85-1.44)
-MTX	37/59 (62.7%)	16/26 (61.5%)	54/75 (72.0%)	3/14 (21.4%)	51/77 (66.2%)	107/152 (70.4%)	123/178 (69.1%)
			1.148 (0.91-1.46)		1.056 (0.83-1.36)	1.123 (0.92-1.42)	1.102 (0.91-1.40)
Patients with ≥ 1 AE of severe intensity	12/113 (10.6%)	1/51 (2.0%)	8/146 (5.5%)	0/28 (0.0%)	8/146 (5.5%)	16/292 (5.5%)	17/343 (5.0%)
			0.516 (0.22-1.19)		0.516 (0.22-1.19)	0.516 (0.26-1.05)	0.467 (0.23-0.94)
Patients with ≥ 1 serious AE	7/113 (6.2%)	0/51 (0.0%)	3/146 (2.1%)	0/28 (0.0%)	4/146 (2.7%)	7/292 (2.4%)	7/343 (2.0%)
			0.332 (0.09-1.15)		0.442 (0.14-1.38)	0.387 (0.14-1.04)	0.329 (0.12-0.89)
+MTX	1/54 (1.9%)	0/25 (0.0%)	2/71 (2.8%)	0/14 (0.0%)	1/69 (1.4%)	3/140 (2.1%)	3/165 (1.8%)
			1.521 (0.20-11.56)		0.783 (0.08-7.46)	1.157 (0.17-8.05)	0.982 (0.14-6.83)
-MTX	6/59 (10.2%)	0/26 (0.0%)	1/75 (1.3%)	0/14 (0.0%)	3/77 (3.9%)	4/152 (2.6%)	4/178 (2.2%)
			0.131 (0.02-0.80)		0.383 (0.11-1.35)	0.259 (0.08-0.83)	0.221 (0.07-0.71)
Patients with ≥ serious infections	4/113 (3.5%)	0/51 (0.0%)	1/146 (0.7%)	0/28 (0.0%)	1/146 (0.7%)	2/292 (0.7%)	2/343 (0.6%)
			0.193 (0.03-1.27)		0.193 (0.03-1.27)	0.193 (0.04-0.89)	0.165 (0.04-0.76)
+MTX	0/54 (0.0%)	0/25 (0.0%)	1/71 (1.4%)	0/14 (0.0%)	0/69 (0.0%)	1/140 (0.7%)	1/165 (0.6%)
			Inf (0.2-Inf)		-	Inf (0.10-Inf)	Inf (0.09-Inf)
-MTX	4/59 (6.8%)	0/26 (0.0%)	0/75 (0.0%)	0/14 (0.0%)	1/77 (1.3%)	1/152 (0.7%)	1/178 (0.6%)
			0.000 (0.00-0.73)		0.192 (0.03-1.24)	0.097 (0.02-0.63)	0.083 (0.01-0.54)
AE leading to discontinuation	5/113 (4.4%)	0/51 (0.0%)	2/146 (1.4%)	0/28 (0.0%)	6/146 (4.1%)	8/292 (2.7%)	8/343 (2.3%)
			0.310 (0.07-1.36)		0.929 (0.31-2.82)	0.619 (0.22-1.78)	0.527 (0.19-1.51)
+MTX	0/54 (0.0%)	0/25 (0.0%)	2/71 (2.8%)	0/14 (0.0%)	2/69 (2.9%)	4/140 (2.9%)	4/165 (2.4%)
			-		-	-	-
-MTX	5/59 (8.5%)	0/26 (0.0%)	0/75 (0.0%)	0/14 (0.0%)	4/77 (5.2%)	4/152 (2.6%)	4/178 (2.2%)
			0.000 (0.00-0.58)		0.613 (0.18-2.05)	0.311 (0.09-1.04)	0.265 (0.08-0.89)
Subjects with ≥ 1 injection site	3/113 (2.7%)	0/51 (0.0%)	7/146 (4.8%)	0/28 (0.0%)	7/146 (4.8%)	14/292 (4.8%)	14/343 (4.1%)

reactions			1.806 (0.52-6.35)		1.806 (0.52-6.35)	1.806 (0.57-1.83)	1.537 (0.49-4.96)
Subjects with tuberculosis	0/113 (0.0%) -		0/146 (0.0%) -		0/146 (0.0%) -	0/292 (0.0%) -	

Appendix 5: Philips et al.[28] checklist

Question(s)	*Response* (Y, N, or NS)	*Comments*
Is there a clear statement of the decision problem?	Y	
Is the objective of the evaluation and model specified and consistent with the stated decision problem?	Y	
Is the primary decision-maker specified?	Y	
Is the perspective of the model stated clearly?	Y	
Are the model inputs consistent with the stated perspective?	Y	
Has the scope of the model been stated and justified?	Y	
Are the outcomes of the model consistent with the perspective, scope and overall objective of the model?	Y	
Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	Y	
Are the sources of data used to develop the structure of the model specified?	Y	Some sources unclear in MS. Further clarification s sought from manufacturers
Are the causal relationships described by the model structure justified appropriately?	Y	
Are the structural assumptions transparent and justified?	N	Issue of 100 mg dose for golimumab is unclear in the MS
Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Y	
Is there a clear definition of the options under evaluation?	Y	
Have all feasible and practical options been evaluated?	Y	

Is there justification for the exclusion of feasible options?	Y	
Is the chosen model type appropriate given the decision problem and specified causal relationships within the model?	Y	
Is the time horizon of the model sufficient to reflect all important differences between options?	Y	
Are the time horizon of the model, the duration of treatment and the duration of treatment effect described and justified?	Y	
Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions?	Y	
Is the cycle length defined and justified in terms of the natural history of disease?	Y	
Are the data identification methods transparent and appropriate given the objectives of the model?	Y	
Where choices have been made between data sources, are these justified appropriately?	Y	Alternatives presented as sensitivity analysis
Has particular attention been paid to identifying data for the important parameters in the model?	N	Incomplete details of methods used to search for resource use data
Has the quality of the data been assessed appropriately?	N	Quality of data derived from expert opinion not discussed
Where expert opinion has been used, are the methods described and justified?	N	No discussion of how expert opinions on the cost of treating psoriasis were synthesised
Is the data modelling methodology based on justifiable statistical and epidemiological techniques?	Y	
Is the choice of baseline data described and justified?	Y	
Are transition probabilities calculated appropriately?	Y	
Has a half-cycle correction been applied to both cost and outcome?	N	

If not, has this omission been justified?		
If relative treatment effects have been derived from trial data, have they been synthesised using appropriate techniques?	Y	
Have the methods and assumptions used to extrapolate short-term results to final outcomes been documented and justified?	N/A	
Have alternative extrapolation assumptions been explored through sensitivity analysis?	N/A	
Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?	Y	
Have alternative assumptions regarding the continuing effect of treatment been explored through sensitivity analysis?	Y	
Are the costs incorporated into the model justified?	N	No justification for including the costs of staff nurse time for the administration of etanercept, adalimumab and golimumab. Additional clarification sought.
Has the source for all costs been described?	N	
Have discount rates been described and justified given the target decision-maker?	Y	
Are the utilities incorporated into the model appropriate?	N	Unclear how the function form for the utility algorithm was derived. Further clarification sought.
Is the source for the utility weights referenced?	Y	
Are the methods of derivation for the utility weights justified?	Y	
Have all data incorporated into the model been described and referenced in sufficient detail?	N	Little detail on the estimates of HAQ gain for cycles 2 and 3 in the model. Little detail on the costs of psoriasis as estimated by expert opinion.
Has the use of mutually inconsistent data been justified (i.e. are assumptions and choices appropriate)?	Y	

Is the process of data incorporation transparent?	Y	
If data have been incorporated as distributions, has the choice of distribution for each parameter been described and justified?	Y	
If data have been incorporated as distributions, is it clear that second order uncertainty is reflected?	Y	
Have the four principal types of uncertainty been addressed?	Y	
If not, has the omission of particular forms of uncertainty been justified?	N/A	
Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?	Y	
Is there evidence that structural uncertainties have been addressed via sensitivity analysis?	Y	
Has heterogeneity been dealt with by running the model separately for different subgroups?	Y	
Are the methods of assessment of parameter uncertainty appropriate?	Y	
If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	N	Unclear where the ranges for the costs associated with HAQ and PASI were obtained from.
Is there evidence that the mathematical logic of the model has been tested thoroughly before use?	N	
Are any counterintuitive results from the model explained and justified?	N	
If the model has been calibrated against independent data, have any differences been explained and justified?	N	Model results compared against trial results utilised in the evidence synthesis
Have the results of the model been compared with those of previous models and any differences in results explained?	N	