Golimumab for the Treatment of Rheumatoid Arthritis after Failure of Previous Disease-Modifying Antirheumatic Drugs

Submission to National Institute of Health and Clinical Excellence

Single technology appraisal (STA)

02 July 2010

Manufacturer

Schering-Plough Ltd (part of MSD)

Contents

List of tables

List of Figures

EXECUTIVE SUMMARY

Introduction

With nearly 400,000 patients affected by Rheumatoid Arthritis (RA) within England and Wales, early initiation of therapy with tumour necrosis factor α (TNFα) inhibitors is a critical component in the pharmacologic management of this chronic, debilitating condition. Clinical evidence suggests that RA patients respond variably to TNFα inhibitors with up to 70% responding within 6 months. RA patients will benefit from proven treatment options which address the challenges throughout the treatment pathway in methotrexate-experienced and TNFα inhibitor-experienced patient populations.

MSD has submitted clinical and cost-effectiveness evidence for the first monthly TNFα inhibitor, golimumab (Simponi®) to inform the appraisal of golimumab for the treatment of RA after failure of previous disease-modifying antirheumatic drugs (DMARDs).

Golimumab (GOL) is the first TNFα inhibitor with randomised, controlled evidence to support a significant clinical response and well tolerated safety profile in DMARD-experienced as well as TNFα inhibitor experienced patients. GOL has been found to have comparable efficacy and safety to the existing biologics. With comparable acquisition costs across the biologic class, the evidence suggests GOL is a cost effective treatment alternative within the NICE willingness to accept threshold. The incremental cost effectiveness ratio (ICER) of GOL compared to standard care is similar to ICERs of other biologics which are currently recommended for treatment by NICE.

Background

RA is the most common inflammatory arthritis in England and Wales associated with severe disability, premature mortality and considerable economic implications: total costs of RA in the UK are estimated to exceed £1 billion per annum.

DMARD experienced
In the treatment of RA after failure of two DMARDs, the National Institute for Health and Clinical Excellence (NICE) has recommended the use of the TNFα inhibitors certolizumab pegol (TA186), adalimumab, etanercept and infliximab (TA130).

TNFα inhibitor experienced
Final guidance for the appraisal of biologic DMARDs after the failure on one TNFα inhibitor remains outstanding, with treatment options limited to only rituximab within this patient population (TA126).

Golimumab RCT evidence strongly supports a robust clinical and safety profile which improves the signs and symptoms of RA, slowing progression of joint damage and improving physical function in both patient populations.

Technology

The key features of GOL are presented in Table 1.

Table 1. Golimumab key features

Approved Name	Golimumab

Brand Name	Simponi®
Marketing Status	European Commission granted marketing authorisation valid throughout the European Union on 1 October 2009.
Pharmacological Action	Human immunoglobulin G1κ (IgG1κ) monoclonal antibody produced by murine hybridoma cell line with recombinant DNA technology. Binds with high affinity and specificity to both soluble and transmembrane forms of TNFα, neutralizing the biological activity of TNFα.
Formulation	One 0.5 ml pre-filled pen/syringe contains 50 mg GOL (injected subcutaneously)
Dosing Frequency	50 mg given once a month, on the same date each month. No loading dose.
Average Length of a Course of Treatment	In accordance with NICE Guidelines, treatment should continue as long as an adequate response (an improvement of ≥1.2 in disease activity score (DAS)) is maintained. Available data suggest that clinical response is usually achieved within 12 to 14 weeks of treatment with GOL (after 3-4 doses). Continued therapy should be reconsidered in patients who show no evidence of therapeutic benefit within this time period.
Average Interval between courses of Treatments	Continuous treatment for at least 12 weeks and then until no response or loss of response
Indications	●*Rheumatoid arthritis (RA*): GOL, in combination with methotrexate (MTX), is indicated for the treatment of moderate to severe, active rheumatoid arthritis in adult patients when the response to DMARD therapy including MTX has been inadequate. ●*Psoriatic arthritis (PsA):* GOL, alone or in combination with MTX, is indicated for the treatment of active and progressive PsA in adult patients when the response to previous DMARD therapy has been inadequate. ●*Ankylosing spondylitis (AS):* GOL is indicated for the treatment of severe, active AS in adult patients who have responded inadequately to conventional therapy.
Annual Acquisition Cost	£9,294.96
Comparators' Annual Acquisition Costs	Adalimumab (£9,295.00) Certolizumab: (£8,794.50)[1] Etanercept: (£9,295.52) Infliximab: (£8,626.11)[2] Rituximab: (£8,335.62)[3]

[1] *Average over 5 year time horizon including year 1 Patient Access Scheme*
[2] *Average over 5 year time horizon including year 1 loading dose*
[3] *Average over 5 year time horizon based on 6 month dosing frequency*

There are no head-to-head trials of any anti-TNF-alpha agents for the treatment of RA. Randomised trials have shown adalimumab, etanercept, infliximab & golimumab to be effective in the treatment of RA. It is therefore important to consider the advantages that golimumab provides over existing treatment options.

Issues with injection site reactions and ease of administration are common concerns for RA patients. Golimumab's profile addresses these areas of concern:

- Golimumab is delivered in a L-histidine buffer (compared to citrate-buffered solution of other TNFα inhibitors) and has low injection volume of 0.5ml thus leading to low incidence of injection site reactions.

- Golimumab's monthly dosing (12 injections per year) reduces injection frequency in comparison to either once weekly or twice weekly dosing with etanercept (52 or 104 injections per year) or adalimumab (26 injections per year).

Clinical Effectiveness of Golimumab

The safety and efficacy of GOL has been robustly assessed with two large, randomised, controlled trials:

- GO-FORWARD (MTX experienced)
- GO-AFTER (TNFα inhibitor-experienced)

DMARD experienced: GO-FORWARD
Treatment of patients with active RA despite methotrexate therapy with GOL 50mg significantly reduced the signs and symptoms of RA and improved physical function.

All primary endpoints were achieved within GO-FORWARD; GOL 50mg demonstrated significant benefit in achieving ACR20 response at week 14 (55.1% vs 33.1%; p=0.001) and median change from baseline in HAQ-DI at week 24 (-0.38 vs -0.13%; p<0.001) compared to placebo. A systematic review of published literature and clinical trials informed meta-analyses which found golimumab to be statistically superior to placebo for the following efficacy endpoints:

- ACR 20 responders: 14, 24 and 52 weeks
- ACR 50 responders: 14, 24 and 52 weeks

The meta-analyses informed a mixed treatment comparison between the TNFα inhibitors (adalimumab, certolizumab pegol, etanercept, infliximab and golimumab) and found no statistical differences at all time points for ACR 20 and ACR 50 responders. A significant amount of heterogeneity was found as trials spanned over a large date range reinforcing the uncertainty within the point estimates of the mixed treatment comparison. Baseline characteristics of recruited RCT patients differed (i.e, disease severity, prior treatment experience) in line with shifts in treatment pathways within the UK as patients with more severe RA are being treated earlier and more aggressively. The TNFα inhibitors are comparable across the main rheumatic treatment outcomes; it is therefore appropriate to view these products as a class with golimumab as a novel addition, in line with previous RA appraisals (TAG130, TAG186 & FAD for ongoing appraisal of sequential use of biologics in RA).

Long term data from GO-FORWARD found GOL 50mg to be clinically effective over year 1 with ACR20, 50 and 70 response rates of 64.0%, 43.8% and 24.7%, respectively. Patients with active RA despite MTX therapy continued to benefit from treatment with GOL 50mg; 90.6% of those patients achieving ACR20 response at week 24 maintained the response at week 52. 61.4% showed DAS28 remission (≤2.6) at week 52 with 36.8% achieving sustained DAS28 remission.

GOL 50mg is generally well-tolerated in combination with methotrexate with serious adverse events (5.6% vs 2.3%) and serious infections (2.2% vs 0.8%) comparable to placebo through week 16. GOL 50mg has a low incidence of injection-site reactions (ISR) and the majority of those ISRs were mild or moderate, with no serious ISRs. Meta-analyses found no statistical

difference between GOL and placebo for the following assessed safety parameters: serious adverse events, serious infections and ISRs. Indirect comparison among the biologics found no statistically significant differences for serious adverse events or serious infections. Golimumab was statistically superior to etanercept with regard to ISRs in the indirect comparison.

TNFα inhibitor experienced: GO-AFTER

GO-AFTER is the first prospective, randomised, double-blind, placebo-controlled phase III trial which investigates the sequential use of a TNFα inhibitor in RA patients. In GO-AFTER, of the patients randomised in the placebo group (n=155) and the GOL 50mg group (n=153), significantly more patients achieved the primary endpoint of ACR20 at week 14 with GOL 50mg than placebo (35.3% vs 18.1%; p<0.001).

Among the subgroup of patients who discontinued one or more prior TNFα inhibitors due to a lack of efficacy, a greater proportion of patients achieved an ACR20 response with GOL 50mg than placebo (35.7% vs 17.7%; p=0.006). Significantly more patients achieved the secondary endpoints of ACR 50 and 70 at week 14 and ACR 20, 50 and 70 at week 24 with GOL 50mg than placebo. At week 24, significantly more GOL 50mg patients had a clinically important reduction in HAQ-DI than in the placebo group (50.0% vs 34.1%; p=0.0044).

Meta-analyses found golimumab and rituximab to be statistically superior to placebo. A mixed treatment comparison found no statistically significant differences between the two biologics.

Cost Effectiveness of Golimumab

Systematic literature reviews and meta-analyses informed a Markov model which found GOL 50mg to be a cost-effective treatment option for patients who demonstrate an inadequate response to DMARDs or a TNFα inhibitor.

DMARD experienced

In a DMARD-experienced patient population, a treatment strategy including golimumab was associated with 5.827 QALYS at a total lifetime cost of £67,747. In comparison, a non-biologic treatment strategy comprising of DMARD therapy, was associated with 4.569 QALYs at a total lifetime cost of £35,870. The incremental cost-effectiveness ratio (ICER) for GOL 50mg compared to non-biologic therapy was estimated at £25,346/QALY (deterministic). Table 2 presents the base case results for all comparators within a DMARD experienced RA patient population.

Table 2: Base-case cost effectiveness results (DMARD experienced RA patient population)

	Golimumab	Methotrexate	Adalimumab	Certolizumab pegol	Etanercept	Infliximab
Total costs	£67,747	£35,869	£66,875	£73,571	£74,208	£69,899
Difference in tl costs	-	£31,878	£872	-£5,824	-£6,461	-£2,152
QALYs	5.827	4.569	5.792	5.768	6.133	5.651
QALY difference	-	1.258	0.035	0.059	-0.306	0.176
ICER	-	£25,346	£25,097	Dominated	*£21,099**	Dominated

* *Based on non-significant point estimates, etanercept was found to be more costly and more effective*

The results indicate golimumab to be a cost effective treatment alternative compared to standard care. Based on non significant efficacy point estimates, golimumab ranges from being cost effective compared to adalimumab to dominating both certolizumab pegol and infliximab. The incremental analysis has been displayed in the Table 3 below.

Table 3: Incremental cost effectiveness results (DMARD experienced RA patient population)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) versus Baseline (Methotrexate)	Incremental analysis
Methotrexate	£35,869	4.569	-	-	-	-
Adalimumab	£66,875	5.792	£31,006	1.223	£25,353	£25,353
Golimumab	£67,747	5.827	£872	0.035	£25,346	£24,914
Infliximab	£69,899	5.651	£2,152	-0.176	£31,464	Dominated
Certolizumab	£73,571	5.768	£3,672	0.117	£31,444	£31,385
Etanercept	£74,208	6.133	£637	0.365	£24,514	£1,745

Table 3 presents ICERs versus standard care which are within the range of accepted values in previous RA NICE appraisals. Whilst the above incremental results have differences in absolute terms, the mixed treatment comparison which inputted into the economic evaluation was found to have no statistically significant differences between the TNFα inhibitors. The figures in the table should thus be viewed a providing a range of values for the class of drugs as a whole rather than providing evidence of difference for cost-effectiveness.

TNFα inhibitor experienced

In a TNFα inhibitor-experienced patient population, a treatment strategy including golimumab was associated with 3.712 QALYS at a total lifetime cost of £50,175. In comparison, a non-biologic treatment strategy comprising of DMARD therapy, was associated with 3.129 QALYs at a total lifetime cost of £33,673. The incremental cost-effectiveness ratio (ICER) for GOL 50mg compared to non-biologic therapy was estimated at £28,826/QALY (deterministic) in a TNFα inhibitor-experienced patient population.

The results of the base case analysis are displayed in Table 4 below.

Table 4. Base-case cost effectiveness results (TNFα inhibitor experienced RA patient population)

	Golimumab	Methotrexate	Rituximab
Total costs	£50,175	£33,673	£50,206
Difference in total costs	-	£16,502	-£31
QALYs	3.712	3.129	3.523
QALY difference	-	0.583	0.189
ICER	-	£28,826	Dominated

The results indicate golimumab to be a cost effective treatment alternative compared to standard care and rituximab. The ICER for golimumab compared to standard care was comparable to ICERs of other subcutaneous TNFα inhibitors already recommended by NICE in RA. The ICERs in relation to rituximab found golimumab to be less costly and more effective in TNFα inhibitor experienced RA patients.

The incremental analysis has been displayed in the Table 5 below.

Table 5: Incremental cost effectiveness results (TNFα inhibitor experienced RA patient population)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) versus Baseline (Methotrexate)	Incremental analysis
Methotrexate	£33,673	3.129	-	-	-	-
Golimumab	£50,175	3.712	£16,502	0.583	£28,286	£28,286
Rituximab	£50,206	3.523	£31	-0.189	£41,935	Dominated

The limited evidence available for TNFα inhibitor experienced patients added to significant uncertainty within the indirect comparison and the resulting incremental analysis. Golimumab and rituximab are comparable to each other across rheumatic treatment outcomes and thus it is appropriate to view these two biologics as a class within the TNFα inhibitor experienced RA patient population.

Budget Impact

The budget implications of introducing GOL into the UK market as a treatment for DMARD experienced and TNFα inhibitor experienced RA populations were estimated with a 5-year time horizon model. The assessment found that the introduction of GOL would increase the budget by £101,482 in 2011, based on a [redacted] GOL market share, increasing to £1,308,426 in year 2014, based on a [redacted] GOL market share.

Conclusion

Robust clinical and safety evidence in the form of two, large RCTs support the conclusion that patients would substantially benefit from the introduction of golimumab for the treatment of moderate to severe RA. Mixed treatment comparisons found golimumab to be superior to standard care and no statistically significant differences in efficacy or safety to the other biologics.

Golimumab would offer DMARD experienced patients the first monthly option with minimal budgetary implications to the NHS. In addition to being a clinically effective alternative within a DMARD experienced patient population, golimumab is the first TNFα inhibitor with clinical evidence from a RCT in a TNFα inhibitor-experienced population. This clinical evidence supports comparable efficacy to the only management option available for those 30% of patients who do not respond adequately to a first TNFα inhibitor. TNFα inhibitor experienced patients would benefit from an auto-injector which would increase treatment options beyond the bi-annual intravenous infusions of rituximab.

In both a DMARD- and TNFα inhibitor-experienced patient population, golimumab was found to have minimal incremental costs to achieve the clinical benefits and thus demonstrated cost-effectiveness.

Golimumab is not only a clinically- and cost-effective treatment in first line treatment of rheumatoid arthritis but has also shown to provide further benefit to English and Welsh patients when used sequentially.

Section A – Decision problem

1 Description of technology under assessment

1.1 Give the brand name, approved name and, when appropriate, therapeutic class. For devices, provide details of any different versions of the same device.

Golimumab (GOL; Simponi®) is a tumour necrosis factor α (TNFα) inhibitor. It is available in either a 0.5 ml pre-filled pen (autoinjector) or pre-filled syringe containing 50 mg of golimumab.

1.2 What is the principal mechanism of action of the technology?

GOL is a human immunoglobulin G1κ (IgG1κ) monoclonal antibody produced by murine hybridoma cell line with recombinant DNA technology. GOL binds with high affinity and specificity to both soluble and transmembrane forms of TNFα, thereby neutralizing the biological activity of TNFα.

1.3 Does the technology have a UK marketing authorisation/CE marking for the indications detailed in this submission? If so, give the date on which authorisation was received. If not, state current UK regulatory status, with relevant dates (for example, date of application and/or expected approval dates).

Positive opinion received from the Committee for Medicinal Products for Human Use on 25 June 2009. The European Commission granted a marketing authorisation valid throughout the European Union for GOL on 1 October 2009.

1.4 Describe the main issues discussed by the regulatory organisation (preferably by referring to the [draft] assessment report [for example, the EPAR]). If appropriate, state any special conditions attached to the marketing authorisation (for example, exceptional circumstances/conditions to the licence).

None.

1.5 What are the (anticipated) indication(s) in the UK?

The approved indications for GOL are as follows:

- Rheumatoid arthritis (RA): GOL, in combination with methotrexate (MTX), is indicated for the treatment of moderate to severe, active rheumatoid arthritis in adult patients when the response to disease-modifying anti-rheumatic drug (DMARD) therapy including MTX has been inadequate.

- Psoriatic arthritis (PsA): GOL, alone or in combination with MTX, is indicated for the treatment of active and progressive PsA in adult patients when the response to previous DMARD therapy has been inadequate.
- Ankylosing spondylitis (AS): GOL is indicated for the treatment of severe, active AS in adult patients who have responded inadequately to conventional therapy.

1.6 Please provide details of all completed and ongoing studies from which additional evidence is likely to be available in the next 12 months for the indication being appraised.

The evidence from completed studies has been included within the clinical effectiveness section. Open label extensions of two and three year efficacy and safety data for C0524T06 (Centocor, Inc. C0524T06. Accessed on: 11 May 2010. Available at: http://clinicaltrials.gov/ct2/results?term=golimumab) and C0524T11 (Centocor, Inc. C0524T11. Accessed on: 11 May 2010. Available at: http://clinicaltrials.gov/ct2/results?term=golimumab) are likely to be published in abstract form by quarter 4, 2010.

1.7 If the technology has not been launched, please supply the anticipated date of availability in the UK.

[redacted]

1.8 Does the technology have regulatory approval outside the UK? If so, please provide details.

Table 6 presents details for GOL regulatory approval received outside of the UK for the treatment of RA, AS and PsA.

Table 6. Golimumab international regulatory approval

Country	Approval Date
Canada	07 April 2009
United States	24 April 2009
European Union	01 October 2009
Australia	06 November 2009
New Zealand	17 December 2009
Croatia	29 March 2010

1.9 Is the technology subject to any other form of health technology assessment in the UK? If so, what is the timescale for completion?

GOL will be submitted for consideration to the Scottish Medicines Consortium in [redacted]

1.10 For pharmaceuticals, please complete the table below. If the unit cost of the pharmaceutical is not yet known, provide details of the anticipated unit cost, including the range of possible unit costs.

Table 7. Unit costs of technology being appraised

Pharmaceutical formulation	One 0.5 ml pre-filled pen/syringe contains 50 mg GOL
Acquisition cost (excluding VAT)	Anticipated to be similar to adalimumab
Method of administration	Injected subcutaneously
Doses	50 mg
Dosing frequency	50 mg given once a month, on the same date each month
Average length of a course of treatment	In accordance with NICE Guidelines, treatment should continue as long as an adequate response (an improvement of ≥1.2 in disease activity score (DAS)) is maintained (NICE. CG79: Rheumatoid arthritis: the management of rheumatoid arthritis in adults. 2009) Available data suggest that clinical response is usually achieved within 12 to 14 weeks of treatment with GOL (after 3-4 doses). Continued therapy should be reconsidered in patients who show no evidence of therapeutic benefit within this time period (MSD. Simponi® 50mg solution for injection. Summary of Product Characteristics. 2010.)
Average cost of a course of treatment	£774.58
Anticipated average interval between courses of treatments	Continuous treatment for at least 12 weeks and then until no response or loss of response
Anticipated number of repeat courses of treatments	Continuous treatment for at least 12 weeks and then until no response or loss of response
Dose adjustments	In patients weighing more than 100 kg who do not achieve an adequate clinical response after 3 or 4 doses, increasing the dose of golimumab to 100 mg once a month may be considered. Continued therapy should be reconsidered in patients who show no evidence of therapeutic benefit after receiving 3 to 4 additional doses of 100 mg.

1.11 For devices, please provide the list price and average selling price. If the unit cost of the device is not yet known, provide details of the anticipated unit cost, including the range of possible unit costs.

Not applicable.

1.12 Are there additional tests or investigations needed for selection, or particular administration requirements for this technology?

Tests, investigations, precautions and monitoring for GOL are consistent with those for all other TNFα inhibitors.

Treatment with GOL is contraindicated in patients with active tuberculosis (TB), other severe infections, moderate or severe heart failure (NYHA class III/IV), or hypersensitivity to the active substance or to any of the following excipients (MSD. Simponi® 50mg solution for injection. Summary of Product Characteristics. 2010.)

- Sorbitol(E420)
- L-histidine
- L-histidine monohydrochloride monohydrate
- Polysorbate 80
- Water for injections

1.13 Is there a need for monitoring of patients over and above usual clinical practice for this technology?

No.

1.14 What other therapies, if any, are likely to be administered at the same time as the intervention as part of a course of treatment?

- Methotrexate,
- Disease Modifying Antirheumatic Drugs/Systemic Immunosuppressive therapy,
- Corticosteroid therapy,
- Nonsteroidal anti-inflammatory drugs (NSAIDs), or
- Other analgesics.

2 Context

2.1 Please provide a brief overview of the disease or condition for which the technology is being used. Include details of the underlying course of the disease.

Rheumatoid arthritis (RA) is a chronic inflammatory disease characterized by progressive joint damage and disability. Swelling of the synovial lining of joints leads to tenderness, pain, impaired joint function and, ultimately, joint erosion. Approximately 400,000 people in the UK have RA (NICE CG79 2009, Symmons et al 2002) with nearly 10,000 new cases being diagnosed each year (NICE CG79 2009, Symmons et al 1994).

RA is associated with significant morbidity and mortality. The ability to perform daily activities is impaired by joint damage, pain and fatigue (Luqmani et al, 2009). Many patients experience severe disability and 2–5 years after the onset of RA about a third of patients are unable to work (Barrett et al 2000, Young et al 2002). RA is associated with a reduced life expectancy (Goodson et al 2005, Radovits et al 2010), and the risk of death from cardiovascular causes is 60% higher in patients with RA than in the general population (Meune et al 2009).

2.2 How many patients are assumed to be eligible? How is this figure derived?

The total numbers of rheumatoid arthritis patients eligible to receive biologic treatment was estimated as 71,062 patients in 2010 as presented in Table 10.

Table 8 and Table 9 present estimates for the number of existing and newly diagnosed rheumatoid arthritis patients in 2010, respectively.

Population estimates for England and Wales (≥ 15 years old) were extracted from the Government Actuary Department 2008 national population projections (Government Actuary's Department. Projections database. Accessed on 12 April 2010. Available at: www.gad.gov.uk/demographyData/Population/index.aspx). Prevalence and incidence were extracted from published sources as 0.81% (Symmons et al. 2002) and 0.015% (Garcia Rodriguez et al. 2009), respectively. Percentage of patients diagnosed (60%); treated (75%) and biologic eligible (42%) were derived from data on file (MSD, 2010).

Table 8. Existing rheumatoid arthritis patients

PREVALENCE BASED	Percentage	2010
Eng/Wales Pop ≥ 15yrs		45,574,176
Prevalence	0.81%	369,151
Diagnosed	60%	221,490
Treated	75%	166,118
Biologic Eligible	42%	69,770

Table 9. Newly diagnosed rheumatoid arthritis patients

INCIDENCE BASED	Percentage	2010
Eng/Wales Pop ≥ 15yrs		45,574,176
Incidence	0.015%	6,836
Diagnosed	60%	4,102
Treated	75%	3,076
Biologic Eligible	42%	1,292

Table 10. Total RA patients in England & Wales eligible for biologic treatment

TOTAL		2010
Prevalence		69,770
Incidence		1,292
Total		**71,062**

2.3 Please give details of any relevant NICE guidance or protocols for the condition for which the technology is being used. Specify whether any specific subgroups were addressed.

NICE has published clinical guidelines and several technology appraisals to review the treatment of rheumatoid arthritis. These include: Clinical Guideline (CG) 79 (NICE, 2009), Technology Appraisal (TA) 126 (NICE. TA126: Rituximab for the treatment of rheumatoid arthritis. 2007), TA130 (NICE. TA130. Adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis. 2007), TA141 (NICE. TA141: Abatacept for the treatment of rheumatoid arthritis. 2008), TA186 (NICE. TA186: Certolizumab pegol for the treatment of rheumatoid arthritis. 2010).

The British Society for Rheumatology (BSR) has updated and published two guidelines on the eligibility criteria for the first biological therapy (Deighton et al, 2010) and the management of RA (after first two years) (Luqmani et al 2009).

The European League Against Rheumatism (EULAR) recently published recommendations for the management of RA with synthetic and biological DMARDs (Smolen et al 2010).

2.4 Please present the clinical pathway of care that depicts the context of the proposed use of the technology. Explain how the new technology may change the existing pathway. If a relevant NICE clinical guideline has been published, the response to this question should be consistent with the guideline and any differences should be explained.

There is no cure for RA; therefore, timely diagnosis and treatment strategies that aim to limit permanent joint damage and disability are crucial. Relief of symptoms, particularly pain, is also a very important goal for patients. NSAIDs and glucocorticoids are recommended for rapid control of symptoms. For reducing disease progression, conventional DMARDs are the established first-line therapy. Conventional DMARDs include methotrexate (MTX), sulphasalazine, hydroxycholoroquinine, leflunomide and gold injections. Of these, MTX is considered the gold standard for RA therapy.

Current NICE guidelines for the management of active RA recommend MTX plus another conventional DMARD and short-term treatment with glucocorticoids as initial therapy. Although conventional DMARD therapy is effective for some patients, for many the response achieved is inadequate. The development of biologic DMARDs has had an important impact on RA management, providing second-line therapy for such patients.

In patients with active disease despite treatment with conventional DMARDs, the addition of a biologic DMARD to the treatment strategy improves symptom control, physical function and quality of life, and reduces radiological progression. In the UK, four biologic DMARDs that block TNFα, namely infliximab, etanercept, adalimumab and certolizumab pegol, are approved for the treatment of RA following the failure of conventional DMARDs. Rituximab, a B-cell-targeted therapy, is approved as a third-line agent for patients with an inadequate response or intolerance to at least one TNFα inhibitor. Tocilizumab, an anti-interleukin-6 agent, and abatacept, a T-cell co-stimulation modulator, are licensed in Europe for the treatment of RA and are undergoing appraisal by NICE.

2.5 Please describe any issues relating to current clinical practice, including any variations or uncertainty about best practice.

Current clinical practice includes multiple options after treatment on a first TNFα inhibitor. These include switching to:

- a different TNFα inhibitor,
- rituximab,
- abatacept or
- tocilizumab.

The treatment of RA after the failure of a TNFα inhibitor with adalimumab, etanercept, infliximab, rituximab and abatacept is currently undergoing a Multiple Technology Appraisal (MTA) by NICE.

Technology Appraisal (TA) 141 does not recommend abatacept for the treatment of RA (NICE, 2008).

Guidance for tocilizumab is not available as the biologic is undergoing the Single Technology Appraisal (STA) NICE process.

In line with the regulatory approval received for GOL, this submission considers the use of GOL in TNFα inhibitor experienced patients. Comparators within this analysis include those treatments which are currently licensed and approved by NICE for use within this patient population. Rituximab is the only biologic currently licensed and NICE approved for the treatment of RA in patients who have had an inadequate response to or intolerance of other DMARDs, including treatment with at least TNFα inhibitor (NICE. TA126: Rituximab for the treatment of rheumatoid arthritis. 2007).

2.6 Please identify the main comparator(s) and justify their selection.

Standard therapy: Management of RA is aimed at limiting permanent joint damage and disability. NSAIDs and glucocorticoids remain an important initial intervention but current practice is aimed at early diagnosis and early use of potential DMARDs to reduce disease progression. Conventional DMARDs include methotrexate (MTX), sulphasalazine, hydroxycholoroquinine, leflunomide and gold injections. Of these, MTX is considered the gold standard for RA therapy.

Biologic DMARD therapy: Patients failing standard care are likely to be offered TNFα inhibitor therapy. Etanercept, infliximab, adalimumab and certolizumab pegol are currently in use for management of active RA in the UK. Patients who achieve an inadequate response on TNFα inhibitors may be offered treatment with rituximab. All five agents are likely to be used in the current practice depending on the patient and physician choice and are therefore deemed to be appropriate comparators.

2.7 Please list therapies that may be prescribed to manage adverse reactions associated with the technology being appraised.

No significant adverse reactions of these treatments are known.

2.8 Does the technology require additional infrastructure to be put in place?

No.

3 Equity and equality

3.1 Identification of equity and equalities issues

3.1.1 Please specify any issues relating to equity or equalities in NICE guidance, or protocols for the condition for which the technology is being used.

None.

3.1.2 Are there any equity or equalities issues anticipated for the appraisal of this technology (consider issues relating to current legislation and any issues identified in the scope for the appraisal)?

No.

3.1.3 How have the clinical and cost-effectiveness analyses addressed these issues?

Not applicable.

4 Statement of the decision problem

Table 11 Decision problem statement

	Final scope issued by NICE	Decision problem addressed in the submission	Rationale if different from the scope
Population	Adults with RA who have had an inadequate response to DMARDs	Adults with moderate to severe, active RA who have had an inadequate response to DMARDs, including methotrexate (MTX)	Patient population defined further as per the GOL draft SPC (MSD,2010). NICE Guideline 79 includes MTX as a first-line treatment (NICE,2009).
Intervention	GOL in combination with methotrexate	Same as in final scope.	N/A
Comparator(s)	Management strategies involving DMARDs without golimumab, including treatment with: • conventional DMARDs (for example, sulfasalazine, leflunomide) • biological agents (including adalimumab, etanercept, infliximab, rituximab, tocilizumab*, certolizumab pegol, abatacept*). *Subject to ongoing appraisal	Management strategies involving DMARDs without golimumab, including treatment with: • conventional DMARDs (for example, sulfasalazine, leflunomide) • biological agents (including adalimumab, etanercept, infliximab, rituximab, certolizumab pegol).	Tocilizumab nor abatacept received positive NICE guidance by the invitation to submit date of 28 April 2010.
Outcomes	The outcome measures to be considered include: • disease activity • physical function • joint damage • pain • mortality • fatigue • radiological progression • extra-articular manifestations of disease • adverse effects of treatment • health related quality of life	The outcome measures addressed include: • disease activity • physical function • joint damage • pain • mortality • fatigue • radiological progression • adverse effects of treatment • health related quality of life	Extra-articular manifestations of disease not routinely reported in RCT
Economic	The reference case stipulates that the cost	Cost effectiveness of treatments expressed in terms of	N/A

analysis	effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective.	incremental cost per quality-adjusted life year. Time horizon considered is lifetime of the patient. Costs are considered from an NHS and Personal Social Services perspective.	
Subgroups to be considered	If evidence allows, the appraisal will consider subgroups of people defined by the baseline severity of their RA. If the evidence allows, the appraisal will consider the costs of joint replacement therapy and hospital admissions. Guidance will only be issued in accordance with the marketing authorisation.	Subgroups include: Biologic experienced patients who discontinued treatment due to lack of efficacy The submission considers cost of hospital admissions. Submission in line with the current marketing authorisation.	Joint replacement data is not available from the pivotal trials.
Special considerations, including issues related to equity or equality	NIL	NIL	N/A

Section B – Clinical and cost effectiveness

5 Clinical evidence

5.1 Identification of studies

5.1.1 Describe the strategies used to retrieve relevant clinical data, both from the published literature and from unpublished data that may be held by the manufacturer or sponsor. The methods used should be justified with reference to the decision problem. Sufficient detail should be provided to enable the methods to be reproduced, and the rationale for any inclusion and exclusion criteria used should be provided. Exact details of the search strategy used should be provided in section 9.2, appendix 2.

To identify relevant studies, searches were performed on 23 March 2010 in the following databases:

- MEDLINE and MEDLINE In process and other non-indexed citations (not restricted by date) via PubMed
- EMBASE (not restricted by date) via Embase.com
- Cochrane Library Central Trials Register (not restricted by date) via Wiley Interscience

Articles retrieved from the database searches were initially screened by title and those that were not relevant to the objective were eliminated. The next stage of the selection process involved screening abstracts; any articles that failed to meet the inclusion criteria were eliminated. For studies that remained after the initial screens, full-text articles were obtained and a further selection was conducted. Included and excluded studies were compared with those in other meta-analyses of biologic DMARDs to confirm that all relevant RCTs had been captured (Blumenauer *et al.*, 2003; Blumenauer *et al.*, 2002; Maxwell and Singh, 2009; Navarro-Sarabia *et al.*, 2005; Roche, 2009; UCB Pharma, 2009). For some of the individual analyses, particular studies could not be included (e.g. because a particular outcome was not reported); such cases are detailed in the relevant results sections, with the reasons for exclusion.

5.2 Study selection

5.2.1 Describe the inclusion and exclusion selection criteria, language restrictions and the study selection process. A justification should be provided to ensure that the rationale is transparent. A suggested format is provided below.

Studies in the systematic review were included or excluded according to the eligibility criteria described in
Table 12.

Table 12. Eligibility criteria used in search strategy

Inclusion criteria	
Populations	1. Adult patients (≥ 18 years) with active RA despite treatment with at least one conventional DMARD for ≥ 3 months; no previous use of anti-TNF-α agents or other biologic agents. 2. Adult patients (≥ 18 years) with active RA despite treatment with at least one anti-TNF-α agent.
Interventions	• Abatacept, adalimumab, certolizumab, etanercept, golimumab, infliximab, rituximab or tocilizumab compared with any other agent including placebo.
Study design	• Double-blind, randomized controlled trials.
Outcomes	• Any of the following outcomes of interest: • measures of treatment efficacy: ACR responses, mean DAS or DAS28, number of patients achieving low DAS (< 3.2) or DAS remission (< 2.6), HAQ-DI. • measures of safety and tolerability: adverse events, treatment discontinuations.
Report characteristics	• Articles for which the full text was available in English. • No publication date restrictions were imposed.
Exclusion criteria	
Populations	• Conventional DMARD-naïve patients. • Mixed populations of both DMARD experienced and anti-TNF-α-experienced patients (> 10% from each group), unless analysed separately
Study design	• Studies with no appropriate comparisons between biologic agents and other active comparators or placebo (e.g. open-label extensions and observational studies). • Studies in which the drug of interest is not administered at the EMA-approved dose or details of dosing are not given. If a study includes more than one treatment arm of the intervention of interest, one of them must be at the approved dose.
Report characteristics	• Reviews, systematic reviews and meta-analyses.

Justification for the above inclusion / exclusion criteria is as follows:

An appropriate trial treatment period with a conventional DMARD is typically considered to be at least 24 weeks (National Institute for Health and Clinical Excellence, 2009a). In line with this guidance, only studies in which all patients had previously received conventional DMARDs were included in this review.

A preliminary search indicated that several large RCTs of biologic DMARDs did not specify the number of previous DMARDs administered or included some patients who had previously received only one DMARD. In addition, the mean duration of previous DMARD therapy was not always specified, and many studies had DMARD treatment for ≥ 3 months (rather than ≥ 24 weeks) as an inclusion criterion. The eligibility criteria of the present analysis allowed for the inclusion of such studies.

To ensure the retrieval of high-quality data, only double-blind RCTs were included in the present analysis. The outcomes selected correspond to the most commonly reported measures of efficacy and safety in RA trials and are in line with outcomes reported in previous NICE submissions from manufacturers of biologic DMARDs. Studies were excluded if an English language full text version of the article was not available.

5.2.2 A flow diagram of the numbers of studies included and excluded at each stage should be provided using a validated statement for reporting systematic reviews and meta-analyses such as the QUOROM statement flow diagram (www.consort-statement.org/?o=1065). The total number of studies in the statement should equal the total number of studies listed in section 5.2.4.

See below in section 5.2.3.

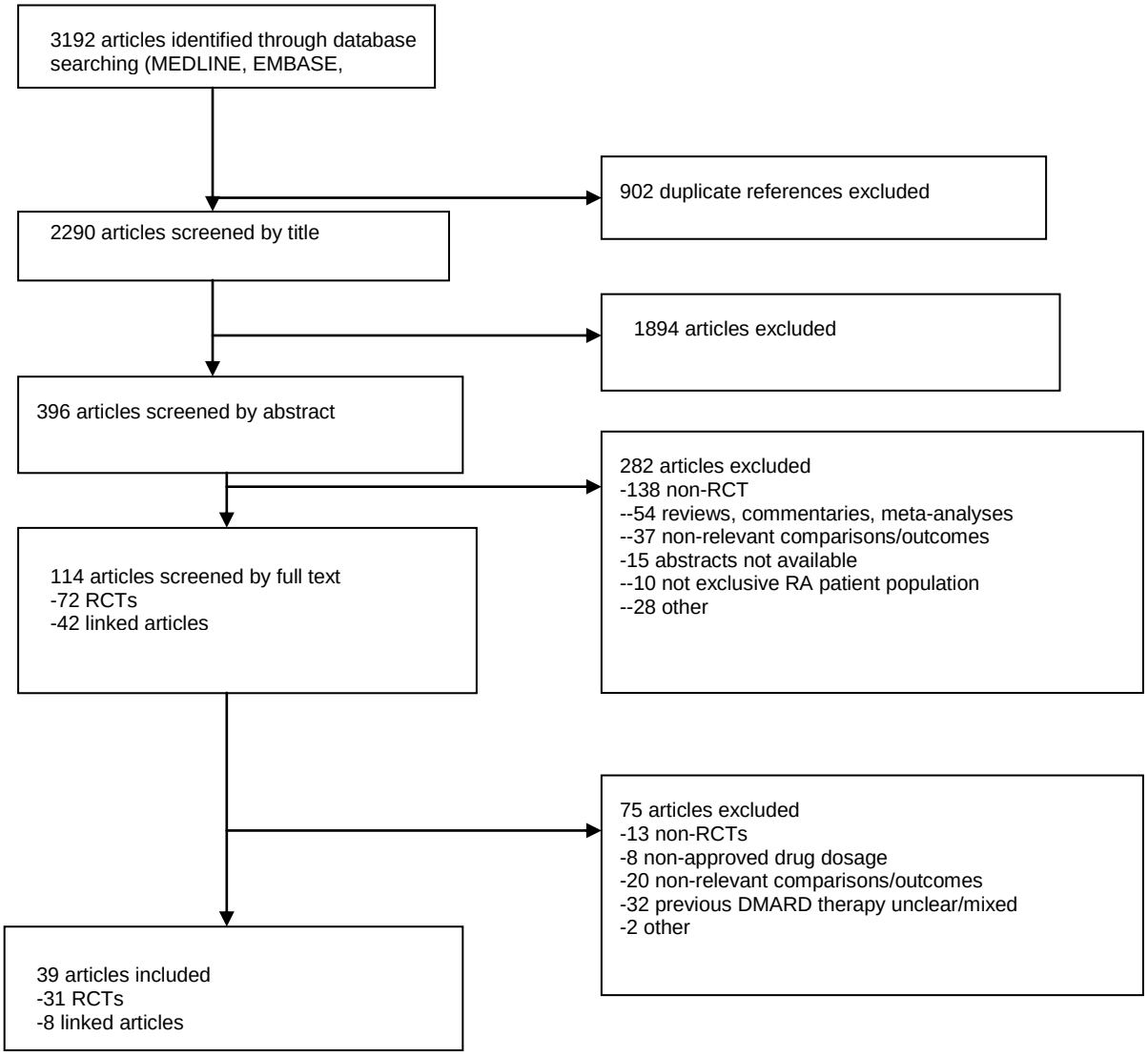
5.2.3 When data from a single RCT have been drawn from more than one source (for example, a poster and a published report) and/or when trials are linked (for example, an open-label extension to an RCT), this should be made clear.

Data presented in this report is based on Clinical Study Reports (CSRs) and published papers. The search strategy identified 2290 articles (after duplicates had been excluded). Based on screening of titles, 1894 articles were excluded. Review of the abstracts for the remaining 396 articles resulted in the exclusion of a further 282 articles, the majority because the studies were not RCTs (A total of 39 articles were selected for inclusion in the meta-analysis; the full list is presented in Table 13. These 39 references report results from a total of 31 individual RCTs; 8 articles were based on other included RCTs.

Figure 1). The full text versions of 114 articles were reviewed to determine eligibility for inclusion. A further 75 articles were excluded: 13 studies were not RCTs; 8 did not use the study drug at the EMEA-approved dose; 20 did not report relevant outcomes or comparisons; and in 32 studies, previous DMARD therapy was unclear or mixed.

A total of 39 articles were selected for inclusion in the meta-analysis; the full list is presented in Table 13. These 39 references report results from a total of 31 individual RCTs; 8 articles were based on other included RCTs.

Figure 1. Flow diagram of study selection process



Complete list of relevant RCTs

5.2.4 Provide details of all RCTs that compare the intervention with other therapies (including placebo) in the relevant patient group.

Table 13. List of relevant RCTs

Reference	*Trial name (acronym or primary reference)*	*Study drug*	*Population*
(Kremer *et al.*, 2003)	(Kremer *et al.*, 2003), phase 2b	abatacept	cDMARD-experienced
(Kremer *et al.*, 2005)	(Kremer *et al.*, 2003), phase 2b	abatacept	cDMARD-experienced
(Kremer *et al.*, 2006)	AIM	abatacept	cDMARD-experienced
(Russell *et al.*, 2007)	AIM	abatacept	cDMARD-experienced
(Weinblatt *et al.*, 2006)	ASSURE	abatacept	cDMARD- and anti-TNF-experienced
(Genovese *et al.*, 2005)	ATTAIN	abatacept	anti-TNF-experienced
(Kim *et al.*, 2007)	(Kim *et al.*, 2007), phase 3	adalimumab	cDMARD-experienced
(van de Putte *et al.*, 2004)	(van de Putte *et al.*, 2004), phase 3	adalimumab	cDMARD-experienced
(Weinblatt *et al.*, 2003)	ARMADA	adalimumab	cDMARD-experienced
(Miyasaka, 2008)	CHANGE	adalimumab	cDMARD-experienced
(Keystone *et al.*, 2004)	DE019	adalimumab	cDMARD-experienced
(Furst *et al.*, 2003)	STAR	adalimumab	cDMARD-experienced
(Chen *et al.*, 2009)	(Chen *et al.*, 2009)	adalimumab	cDMARD-experienced
(Keystone *et al.*, 2008b)	RAPID 1	certolizumab	cDMARD-experienced
(Strand *et al.*, 2009)	RAPID 1	certolizumab	cDMARD-experienced
(Smolen *et al.*, 2009b)	RAPID 2	certolizumab	cDMARD-experienced
(Klareskog *et al.*, 2004)	TEMPO	etanercept	cDMARD-experienced
(van der Heijde *et al.*, 2006b)	TEMPO	etanercept	cDMARD-experienced
(Keystone *et al.*, 2009a)	TEMPO and ERA	etanercept	cDMARD-experienced
(Combe *et al.*, 2006)	(Combe *et al.*, 2006)	etanercept	cDMARD-experienced
(Moreland *et al.*, 1999)	(Moreland *et al.*, 1999)	etanercept	cDMARD-experienced
(Weinblatt *et al.*, 1999)	(Weinblatt *et al.*, 1999)	etanercept	cDMARD-experienced
(Smolen *et al.*, 2009a)	GO-AFTER	golimumab	anti-TNF-experienced

Reference	*Trial name (acronym or primary reference)*	*Study drug*	*Population*
(Keystone *et al.*, 2009b)	GO-FORWARD	golimumab	cDMARD-experienced
(Kay *et al.*, 2008)	(Kay *et al.*, 2008)	golimumab	cDMARD-experienced
(Schiff *et al.*, 2008)	ATTEST	infliximab, abatacept	cDMARD-experienced
(Lipsky *et al.*, 2000)	ATTRACT	infliximab	cDMARD-experienced
(Maini *et al.*, 1999)	ATTRACT	infliximab	cDMARD-experienced
(Westhovens *et al.*, 2006b)	START	infliximab	cDMARD-experienced
(Abe *et al.*, 2006)	(Abe *et al.*, 2006)	infliximab	cDMARD-experienced
(Maini *et al.*, 1998)	(Maini *et al.*, 1998)	infliximab	cDMARD-experienced
(Edwards *et al.*, 2004)	(Edwards *et al.*, 2004), phase 2	rituximab	cDMARD-experienced
(Strand *et al.*, 2006)	(Edwards *et al.*, 2004), phase 2	rituximab	cDMARD-experienced
(Cohen *et al.*, 2006)	REFLEX	rituximab	anti-TNF-experienced
(Keystone *et al.*, 2008a)	REFLEX	rituximab	anti-TNF-experienced
(Smolen *et al.*, 2008)	OPTION	tocilizumab	cDMARD-experienced
(Emery *et al.*, 2008)	RADIATE	tocilizumab	anti-TNF-experienced
(Nishimoto *et al.*, 2009)	SATORI	tocilizumab	cDMARD-experienced
(Genovese *et al.*, 2008)	TOWARD	tocilizumab	cDMARD-experienced

A list of excluded studies and reasons for exclusion are presented in Section 9.2.6 in Table 168

5.2.5 Please highlight which of the RCTs identified above compares the intervention directly with the appropriate comparator(s) with reference to the decision problem. If there are none, please state this.

All studies compare intervention with placebo; No head to head studies were available.

5.2.6 When studies identified above have been excluded from further discussion, a justification should be provided to ensure that the rationale for doing so is transparent. For example, when studies have been identified but there is no access to the level of trial data required, this should be indicated.

No studies have been excluded.

List of relevant non-RCTs

5.2.7 Please provide details of any non-RCTs (for example experimental and observational data) that are considered relevant to the decision problem and a justification for their inclusion.

No non-RCT evidence was included in the clinical section.

5.3 Summary of methodology of relevant RCTs

5.3.1 As a minimum, the summary should include information on the RCT(s) under the subheadings listed in this section.

Please see section 5.3.2.

Methods

5.3.2 Describe the RCT(s) design (for example, duration, degree and method of blinding, and randomisation) and interventions. Include details of length of follow-up and timing of assessments. The following tables provide a suggested format for when there is more than one RCT.

DMARD experienced population

Of the 32 RCTs identified by the selection process, 28 were identified for inclusion in the meta-analysis for the DMARD experienced population: 2 golimumab studies, 3 abatacept studies, 7 adalimumab studies, 3 certolizumab pegol studies, 4 etanercept studies, 5 infliximab studies (one of which [Schiff *et al.* 2008] included an abatacept

treatment arm), 1 rituximab study and 3 tocilizumab studies. For the following studies, data were sought from more than 1 reference:

- Kremer *et al.* (2003) (abatacept): 2 references (Kremer *et al.*, 2005; Kremer *et al.*, 2003)
- ATTRACT (infliximab): 2 references (Lipsky *et al.*, 2000; Maini *et al.*, 1999)
- Edwards *et al.* (2004) (rituximab): 2 references (Edwards *et al.*, 2004; Strand *et al.*, 2006)
- AIM (abatacept): 2 references (Kremer *et al.*, 2006; Russell *et al.*, 2007)
- RAPID 1 (certolizumab): 2 references (Keystone *et al.*, 2008b; Strand *et al.*, 2009)
- TEMPO (etanercept): 3 references (Keystone *et al.*, 2009a; Klareskog *et al.*, 2004; van der Heijde *et al.*, 2006b)

TNFα experienced population

Of the 32 RCTs identified by the selection process, five RCTs were identified for inclusion in the meta-analysis for the TNFα experienced population: 1 golimumab study, 2 abatacept studies, 1 rituximab study and 1 tocilizumab study. For the REFLEX trial (rituximab), data were sought from 2 references (Cohen *et al.*, 2006; Keystone *et al.*, 2008a).

The ASSURE trial involved both conventional DMARD experienced patients and TNFα experienced patients, as separate subgroups. Data from this study were therefore included in the analyses for both populations.

Table 14. Comparative summary of methodology of the RCTs (DMARD experienced population)

Trial	*Trial design (blinding, phase, location, sponsor)*	*Number of patients*	*Interventions compared (number of patients)*	*Dosing and administration*	*Duration of follow-up (weeks)*	*Concomitant medication*	*Primary outcomes*
(Kremer *et al.*, 2003)	Randomized, double-blind, placebo-controlled phase 2b study, location not reported, sponsored by Bristol-Myers Squibb	339[1]	Placebo + MTX (119) Abatacept + MTX (115)	Abatacept (10 mg/kg) or placebo i.v. days 1, 15, 30 and monthly thereafter MTX (10–30 mg) every week	52	Stable low-dose corticosteroids (≤10 mg per day) and NSAIDs were permitted; cDMARDs other than MTX discontinued prior to study	ACR20 response at 24 weeks
AIM	Randomized, double-blind, placebo-controlled phase 3 study, multinational, sponsored by Bristol-Myers Squibb	652	Placebo + MTX (219) Abatacept + MTX (433)	Abatacept (10 mg/kg) or placebo i.v. days 1, 15, 29 and every 28 days thereafter MTX (≥ 10 mg) every week	52	Stable doses of NSAIDs and corticosteroid (≤ 10 mg prednisone/day) permitted if stabilized for 25 days before randomization; cDMARDs other than MTX discontinued prior to study	ACR20 response at 24 weeks; proportion of patients achieving ≥ 0.3 improvement in HAQ-DI at 1 year; radiographic progression of joint erosions (Genant-modified Sharp score) at 1 year
ASSURE	Randomized, double-blind, placebo-controlled study, USA and Canada, sponsored by Bristol-Myers Squibb	1274	Placebo + cDMARDs (417) Abatacept + cDMARDs (855)	Abatacept (10 mg/kg) or placebo i.v. days 1, 15, 29 and every 28 days thereafter Background DMARD doses not reported	52	All patients continued to receive their background cDMARD therapies (MTX, [hydroxy]chloroquine, sulfasalazine, leflunomide, gold, azathioprine); oral corticosteroids (≤ 10 mg/day) and stable doses of NSAIDs were permitted	Incidence of adverse events at every study visit (day 1, 15, 29 and every 4 weeks thereafter)
(Kim *et al.*, 2007)	Randomized, double-blind, placebo-controlled, phase 3 study, Korea, sponsored by Abbott Laboratories	128	Placebo + MTX (63) Adalimumab + MTX (65)	Adalimumab (40 mg) or placebo s.c. every other week MTX (no dosing guidelines)	24	MTX required; other concomitant medications not specified; cDMARDs other than MTX discontinued prior to study	ACR20 response at week 24

Trial	*Trial design (blinding, phase, location, sponsor)*	*Number of patients*	*Interventions compared (number of patients)*	*Dosing and administration*	*Duration of follow-up (weeks)*	*Concomitant medication*	*Primary outcomes*
(van de Putte *et al.*, 2004)	Randomized, double-blind, placebo-controlled, phase 3 study, multinational, sponsored by Abbott Laboratories	544[1]	Placebo (110) Adalimumab (113)	Adalimumab (40 mg) s.c. every other week, placebo s.c. every week (placebo given to adalimumab group in alternate weeks)	26	Analgesics such as propoxyphene, codeine, or aspirin were permitted, but not within 12 hours of study visits; all cDMARDs discontinued prior to study	ACR20 response at week 26
ARMADA	Randomized, double-blind, placebo-controlled study, USA and Canada, sponsored by Abbott Laboratories	271[1]	Placebo + MTX (62) Adalimumab + MTX (67)	Adalimumab (40 mg) or placebo s.c. every other week MTX (12.5–25 mg) every week	24	Salicylates, NSAIDs and corticosteroids (≤ 10 mg/day oral prednisone or equivalent), non-opioid analgesics permitted; cDMARDs other than MTX discontinued prior to study	ACR20 response
CHANGE	Randomized, double-blind, placebo-controlled phase 2/3 study, Japan, sponsored by Abbott Japan	352[1]	Placebo (87) Adalimumab (91)	Adalimumab (40 mg) or placebo s.c. every other week	24	MTX and other cDMARDs were not allowed; use of other concomitant medication not specified	ACR20 response at week 24
DE019	Randomized, double-blind, placebo-controlled study, USA and Canada, sponsored by Abbott Laboratories	619[1]	Placebo + MTX (200) Adalimumab + MTX (207)	Adalimumab (40 mg) s.c. every other week, placebo s.c. every week (placebo given to adalimumab group in alternate weeks) MTX (12.5–25 mg) every week	52	MTX, corticosteroids (≤ 10 mg/day), NSAIDs permitted and kept constant throughout study; cDMARDs other than MTX discontinued prior to study	ACR20 response at week 24

Trial	*Trial design (blinding, phase, location, sponsor)*	*Number of patients*	*Interventions compared (number of patients)*	*Dosing and administration*	*Duration of follow-up (weeks)*	*Concomitant medication*	*Primary outcomes*
STAR	Randomized, double-blind, placebo-controlled study, USA and Canada, sponsored by Abbott Laboratories	636	Placebo + cDMARDs (318) Adalimumab + DMARDs (318)	Adalimumab (40 mg) or placebo s.c. every other week Background DMARD doses not reported	24	cDMARDs, low dose corticosteroids (≤ 10 mg/day prednisone), NSAIDs or analgesics were permitted if at stable dose for ≥ 28 days	Incidence of adverse events, physical examination findings, standard laboratory test results
(Chen *et al.*, 2009)	Randomized, double-blind, placebo-controlled study, Taiwan, sponsor not provided	47	Placebo + MTX (12) Adalimumab + MTX (35)	Adalimumab (40 mg) or placebo s.c. every other week MTX (10–15 mg) every week	12	Concomitant use of NSAIDs, oral corticosteroids, MTX and aspirin permitted as long as dose was maintained; other cDMARDs discontinued prior to study	ACR20 response at week 12
RAPID 1	Randomized, double-blind, placebo-controlled phase 3 study, multinational, sponsored by UCB Inc	982[1]	Placebo + MTX (199) Certolizumab 200 mg + MTX (393)	Certolizumab (400 mg) or placebo s.c. weeks 0, 2, 4, 200 mg or placebo every 2 weeks thereafter MTX (≥ 10 mg) every week	52	NSAIDs, oral corticosteroids (≤ 10 mg prednisone/day) were allowed if dosage stable for ≥ 4 weeks prior to study; cDMARDs other than MTX discontinued prior to study	ACR20 response at week 24, mean change from baseline in modified total Sharp score at week 52
RAPID 2	Randomized, double-blind, placebo-controlled phase 3 study, multinational, sponsored by UCB Inc	619[1]	Placebo + MTX (127) Certolizumab 200 mg + MTX (246)	Certolizumab (400 mg) or placebo s.c. weeks 0, 2, 4, 200 mg or placebo every 2 weeks thereafter MTX (≥ 10 mg) every week	24	Oral corticosteroids (≤ 10 mg prednisone/day) and NSAIDs permitted if doses stable within 28 and 14 days of baseline, respectively	ACR20 response at week 24
TEMPO	Randomized, double-blind, placebo-controlled study, multinational, sponsored by Wyeth Research	682	Placebo + MTX (228) Etanercept + placebo (223) Etanercept + MTX (231)	Etanercept (25 mg) or placebo s.c. twice weekly MTX (7.5–20 mg) every week or placebo	52	Not specified, assumed NSAIDs and corticosteroids as these were listed in patient baseline characteristics table; cDMARDs including MTX discontinued prior to study	Numeric index of ACR response area under the curve over first 24 weeks

Trial	*Trial design (blinding, phase, location, sponsor)*	*Number of patients*	*Interventions compared (number of patients)*	*Dosing and administration*	*Duration of follow-up (weeks)*	*Concomitant medication*	*Primary outcomes*
(Combe *et al.*, 2006)	Randomized, double-blind, placebo-controlled study, multinational, sponsored by Wyeth Research	254	Placebo + sulfasalazine (50) Etanercept + placebo (103) Etanercept + sulfasalazine (101)	Etanercept (25 mg) or placebo s.c. twice weekly Sulfasalazine (2–3 g/day) or placebo	24	NSAIDs, oral corticosteroids (≤ 10 mg/day), simple analgesics and aspirin (≤ 300 mg) were permitted; cDMARDs other than sulfasalazine discontinued prior to study	ACR20 response at week 24
(Moreland *et al.*, 1999)	Randomized, double-blind, placebo-controlled study, USA and Canada, sponsored by Immunex	234[1]	Placebo (80) Etanercept (78)	Etanercept (25 mg) or placebo s.c. twice weekly	26	NSAIDs (not exceeding manufacturer dose) and corticosteroids (≤ 10 mg/day prednisone) were permitted if doses stable; all cDMARDs discontinued prior to study	ACR20 response and ACR50 response at 3 and 24 weeks
(Weinblatt *et al.*, 1999)	Randomized, double-blind, placebo-controlled study, USA, sponsored by Immunex	89	Placebo + MTX (30) Etanercept + MTX (59)	Etanercept (25 mg) or placebo s.c. twice weekly MTX (15–25 mg) every week	24	NSAIDs and prednisone (≤ 10 mg/kg) were permitted if dose stable for ≥ 4 weeks; cDMARDs other than MTX discontinued prior to study	ACR20 response at 24 weeks, incidence of adverse events
GO-FORWARD	Randomized, double-blind, placebo controlled phase 3 study, multinational, sponsored by Centocor	444	Placebo + MTX (133) Golimumab + MTX (89)	Golimumab (50 mg) or placebo s.c. every 4 weeks MTX (≥ 15 mg) every week	24	NSAIDs, analgesics, oral prednisone ≤ 10 mg/day allowed if stable for ≥ 2 weeks; other cDMARDs discontinued prior to study	ACR20 response at 24 weeks

Trial	*Trial design (blinding, phase, location, sponsor)*	*Number of patients*	*Interventions compared (number of patients)*	*Dosing and administration*	*Duration of follow-up (weeks)*	*Concomitant medication*	*Primary outcomes*
(Kay *et al.*, 2008)	Randomized, double-blind, placebo-controlled study, multinational, sponsored by Centocor	172[1]	Placebo + MTX (35) Golimumab (every 4 weeks) + MTX (35)	Golimumab (50 mg) or placebo s.c. every 2 weeks (golimumab every 4 weeks with placebo on alternate weeks) MTX (≥ 10 mg) every week	52	NSAIDs, oral prednisone ≤ 10 mg/day allowed if stable for ≥ 4 weeks; other cDMARDs discontinued prior to study	ACR20 response at week 16
ATTEST	Randomized, double-blind, placebo- and active (infliximab)-controlled phase 3 study, multinational, sponsored by Bristol-Myers Squibb	431	Abatacept + MTX (156) Placebo + MTX (110) Infliximab + MTX (165)	Abatacept (10 mg/kg) or placebo i.v. day 1, 15, 29, and every 28 days thereafter Infliximab (3 mg/kg) or placebo i.v. day 1, 14, 43, 85, and every 56 days thereafter MTX (≥ 15 mg) every week	52	Permitted between days 1–197: oral corticosteroids (≤ 10 mg/day, stable for ≥ 25 of 28 days prior to study), NSAIDs; cDMARDs other than MTX discontinued prior to study; dose adjustments were permitted after day 198 as well as add-on therapy of other DMARDs	Change from baseline in DAS28 at 24 weeks (abatacept vs placebo)
ATTRACT	Randomized, double-blind, placebo-controlled phase 3 study, multinational, sponsored by Centocor	428[1]	Placebo + MTX (88) Infliximab (every 4 weeks) + MTX (86) Infliximab (every 8 weeks) + MTX (86)	Infliximab (3 mg/kg) or placebo i.v. week 0, 2, 6, and every 4 weeks thereafter (note: some infliximab patients received drug every 8 weeks, placebo every other infusion) MTX (≥ 12.5 mg) every week	54	NSAIDs and oral corticosteroids (≤ 10 mg/kg) were permitted if dose stable for ≥ 4 weeks; other cDMARDs discontinued prior to study	ACR20 response at week 30

Trial	*Trial design (blinding, phase, location, sponsor)*	*Number of patients*	*Interventions compared (number of patients)*	*Dosing and administration*	*Duration of follow-up (weeks)*	*Concomitant medication*	*Primary outcomes*
START	Randomized, double-blind, placebo-controlled study, multinational, sponsored by Centocor	1082[1]	Placebo + MTX (361) Infliximab + MTX (360)	Infliximab (3 mg/kg) or placebo i.v. week 0, 2, 6, 14, and every 8 weeks thereafter MTX (≤ 25 mg) every week	54	NSAIDs, oral corticosteroids, other cDMARDs permitted if dosing stable for ≥ 4 weeks	Proportion of patients experiencing a serious infection within first 22 weeks
(Abe *et al.*, 2006)	Randomized, placebo-controlled, double-blind study, Japan, sponsor not reported	147[1]	Placebo + MTX (47) Infliximab + MTX (49)	Infliximab (3 mg/kg) or placebo i.v. week 0, 2, 6, and every 8 weeks thereafter MTX (≥ 6 mg) every week	14	NSAIDs, folic acid, corticosteroids (≤ 10 mg/day) permitted if dose stable for ≥ 4 weeks; other cDMARDs discontinued prior to study	ACR20 response at week 14
(Maini *et al.*, 1998)	Randomized, double-blind, placebo-controlled phase 2 trial, multinational, sponsored by Centocor	101[1]	Placebo + MTX (14) Infliximab + placebo (14) Infliximab + MTX (15)	Infliximab (3 mg/kg) or placebo i.v. week 0, 2, 6, 10, 14 MTX (7.5–15 mg) every week	26	cDMARDs other than MTX were discontinued prior to the study, patients were permitted oral corticosteroids (≤ 7.5 mg/day) if dosage stable for ≥ 4 weeks, NSAIDs permitted	Total time (weeks) that patients exhibited response to treatment (defined by Paulus 20% index)
(Edwards *et al.*, 2004)	Randomized, double-blind, placebo-controlled phase 2 study, multinational, sponsored by Roche	161	Placebo + MTX (40) Rituximab + placebo (40) Rituximab + cyclophosphosphamide + placebo (41) Rituximab + MTX + placebo (40)	Rituximab (1000 mg/day) or placebo i.v. day 1, 15 Cyclophosphamide (750 mg) day 3, 17 MTX (≥ 10 mg) every week	48	NASIDs at stable doses or corticosteroids (≤ 12.5 mg/day) were allowed; other cDMARDs discontinued prior to study	ACR50 response at week 24

Trial	*Trial design (blinding, phase, location, sponsor)*	*Number of patients*	*Interventions compared (number of patients)*	*Dosing and administration*	*Duration of follow-up (weeks)*	*Concomitant medication*	*Primary outcomes*
OPTION	Randomized, double-blind, placebo-controlled phase 3 study, multinational, sponsored by F Hoffmann-La Roche and Chugai Pharmaceutical	622[1]	Placebo + MTX (204) Tocilizumab + MTX (205)	Tocilizumab (8 mg/kg) or placebo i.v. every 4 weeks MTX (10–25 mg) every week	32	Oral glucocorticoids (≤ 10 mg/day) and NSAIDs were permitted if doses were stable for ≥ 6 weeks before study; cDMARDs other than MTX discontinued prior to study	ACR20 response at 24 weeks
SATORI	Randomized, double-blind, placebo-controlled study, Japan, sponsored by Chugai Pharmaceutical	125	Placebo + MTX (64) Tocilizumab + MTX (61)	Tocilizumab (8 mg/kg) or placebo i.v. every 4 weeks MTX (8 mg) every week	24	Oral corticosteroids (≤ 10 mg/day stable for minimum 2 weeks), intra-articular corticosteroids, hyaluronate preparations, NSAIDs were allowed; other cDMARDs discontinued prior to study	ACR20 response at week 24
TOWARD	Randomized, double-blind, placebo-controlled phase 3 study, multinational, sponsored by F Hoffmann-La Roche	1216	Placebo + DMARDs (413) Tocilizumab + DMARDs (803)	Tocilizumab (8 mg/kg) or placebo i.v. every 4 weeks Background DMARD doses not reported	24	Oral glucocorticoids (≤ 10 mg/day) and NSAIDs were permitted if doses were stable for ≥ 6 weeks before study	ACR20 response at week 24

[1]Intent-to-treat population provided for all treatment arms, including those not analysed in systematic review.

Table 15. Comparative summary of methodology of the RCTs (TNFα experienced population)

Trial	*Trial design (blinding, phase, location, sponsor)*	*Number of patients*	*Interventions compared (number of patients)*	*Dosing and administration*	*Duration of follow-up (weeks)*	*Concomitant medication*	*Primary outcomes*
ASSURE	Randomized, double-blind, placebo-controlled study, USA and Canada, sponsored by Bristol-Myers Squibb	167	Placebo + DMARDs (64) Abatacept + DMARDs (103)	Abatacept (10 mg/kg) or placebo i.v. days 1, 15, 29, and every 28 days thereafter Background DMARD doses not reported	52	All patients continued to receive their background DMARD therapies (MTX, [hydroxy]chloroquine, sulfasalazine, leflunomide, gold, azathioprine); oral corticosteroids (≤ 10 mg/day) and stable doses of NSAIDs were permitted	Incidence of adverse events at every study visit (day 1, 15, 29 and every 4 weeks thereafter)
ATTAIN	Randomized, double-blind, placebo-controlled phase 3 study, multinational, sponsored by Bristol-Myers Squibb	389	Placebo + DMARDs (133) Abatacept + DMARDs (256)	Abatacept (10 mg/kg) or placebo i.v. day 1, 15, 29, and every 28 days thereafter Background DMARD doses not reported	24	Oral corticosteroids permitted (≤ 10 mg/day) if dose stable for ≥ 28 days; 38–41% were taking concomitant anti-TNF therapies; approx 70% were taking NSAIDs; biologics discontinued prior to study	Proportion of patients achieving ACR20 and proportion of patients with an improvement of ≥ 0.3 from baseline in HAQ-DI, at 24 weeks
GO-AFTER	Randomized, double-blind, placebo-controlled phase 3 study, multinational, sponsored by Centocor and Schering-Plough	461[1]	Placebo (155) Golimumab (153) (note: background DMARDs optional)	Golimumab (50 mg) or placebo s.c. every 4 weeks MTX (≥ 10 mg) every week Other DMARD doses not specified	24	DMARDs (MTX, sulfasalazine and hydroxychloroquine) permitted, but not required, if patients tolerated dose for ≥ 12 weeks and dose stable for ≥ 4 weeks prior to study; prednisone ≤ 10 mg/day or NSAIDs also allowed if dose stable for ≥ 2 weeks prior to study	Proportion of patients achieving ACR20 at week 14
REFLEX	Randomized, double-blind, placebo-controlled phase 3 study, multinational, sponsored by Hoffmann-La Roche, Biogen Idec, Genentech	499	Placebo + MTX (201) Rituximab + MTX (298)	Rituximab (2 x 1000 mg) or placebo i.v. days 1, 15 MTX (10–25 mg) every week	24 (note: overall REFLEX study 2 years)	Glucocorticoids (≤ 10 mg/day) and NSAIDs were permitted if dosage stable for ≥ 4 weeks and ≥ 2 weeks, respectively; all patients were treated with i.v. methylprednisone (100 mg day 1, 15), oral prednisone (60 mg days 2–7, 30 mg days 8–14); biologic DMARDs discontinued prior to study	Proportion of patients achieving ACR20 at week 24

Trial	*Trial design (blinding, phase, location, sponsor)*	*Number of patients*	*Interventions compared (number of patients)*	*Dosing and administration*	*Duration of follow-up (weeks)*	*Concomitant medication*	*Primary outcomes*
RADIATE	Randomized, double-blind, placebo-controlled phase 3 study	489[1]	Placebo + MTX (158) Tocilizumab + MTX (158)	Tocilizumab (8 mg/kg) or placebo i.v. every 4 weeks MTX (10–25 mg/week)	24	All patients received stable MTX and folate, other DMARDs discontinued prior to study; patients were allowed to continue stable oral corticosteroids (≤ 10 mg/day) and/or NSAIDS	Proportion of patients achieving ACR20 at week 24

[1]Intent-to-treat population provided for all treatment arms, including those not analyzed in systematic review.

Participants

5.3.3 Provide details of the eligibility criteria (inclusion and exclusion) for the trial. The following table provides a suggested format for the eligibility criteria for when there is more than one RCT. Highlight any differences between the trials.

Inclusion and exclusion criteria can be found in Appendix 8.14, Table 177.

5.3.4 Describe the patient characteristics at baseline. Highlight any differences between study groups. The following table provides a suggested format for the presentation of baseline patient characteristics for when there is more than one RCT.

Patient baseline characteristics were similar among the studies in the analysis of the conventional DMARD experienced population (Table 16) and were also similar among the studies in the analysis of the TNFα experienced population (Table 17).

Table 16. Characteristics of participants in the RCTs across randomised groups (DMARD experienced population)

Trial	*Interventions compared (number of patients)*	*Gender (% female)*	*Age (yrs)*	*Disease duration (yrs)*	*Prior DMARD treatment (names, mean no., mean duration)*	*Positive rheumatoid factor (%)*	*Dose MTX (mg/ week)*	*Concomitant NSAID therapy (%)*	*Concomitant glucocorticoid therapy (%)*	*Mean DAS28*	*CRP (mg /dL)*	*HAQ-DI*	*Tender joint count*	*Swollen joint count*	*Patient assessment of pain (VAS)*
(Kremer *et al.*, 2003)	placebo + MTX (119)	66	54.7 (23–80)[1]	8.9 ± 8.3	MTX, leflunomide, infliximab (< 3% patients), mean no. not specified, mean duration 2.9 ± 3.5 years	90	15.8 ± 4.1	–	–	5.5 ± 0.87	3.2 ± 3.2	1.0 ± 0.6	29.2 ± 13.0	21.8 ± 8.8	6.52 ± 2.21
	abatacept + MTX (115)	75	55.8 (17–83) [1]	9.7 ± 9.8	MTX, lefluno-mide, infliximab (< 3% patients), mean no. not specified, mean duration 2.5 ± 2.7 years	99	15.0 ± 4.4	–	–	5.5 ± 0.63	2.9 ± 2.8	1.0 ± 0.5	30.8 ± 12.2	21.3 ± 8.4	6.21 ± 2.14
AIM	placebo + MTX (219)	81.7	50.4 ± 12.4	8.9 ± 7.1	MTX and other DMARDs, mean no. not specified, mean duration not specified (MTX ≥ 3 months)	78.5	15.7 ± 3.5	82.6	68.5	6.4	2.8 ± 2.5	1.7 ± 0.6 (n = 219)	32.3 ± 13.6	22.1 ± 8.8	6.59 ± 2.06
	abatacept + MTX (433)	77.8	51.5 ± 12.9	8.5 ± 7.3	as above	81.8	16.1 ± 3.6	85.5	72.1	6.4	3.3 ± 3.1	1.7 ± 0.7 (n = 431)	31.0 ± 13.2	21.4 ± 8.8	6.33 ± 2.11
ASSURE	placebo + MTX (417)	83.7	52.0 ± 12.1	9.5 ± 9.1	MTX, (hydroxy)chloroquine, sulfasalazine, leflunomide, gold, azathioprine, mean no. not specified, mean duration not specified (≥ 3 months)	–	–	–	73.7	–	2.1 ± 2.6	1.5 ± 0.7	–	–	6.13 ± 2.08
	adalimumab + MTX (855)	83.1	52.2 ± 11.8	9.5 ± 8.7	as above	–	–	–	71.6	–	1.9 ± 2.4	1.5 ± 0.6	–	–	6.11 ± 2.04

Trial	*Interventions compared (number of patients)*	*Gender (% female)*	*Age (yrs)*	*Disease duration (yrs)*	*Prior DMARD treatment (names, mean no., mean duration)*	*Positive rheumatoid factor (%)*	*Dose MTX (mg/ week)*	*Concomitant NSAID therapy (%)*	*Concomitant glucocorticoid therapy (%)*	*Mean DAS28*	*CRP (mg /dL)*	*HAQ-DI*	*Tender joint count*	*Swollen joint count*	*Patient assessment of pain (VAS)*
(Kim *et al.*, 2007)	placebo + MTX (63)	85.7	49.8 ± 10.5	6.9 ± 4.5 (n = 36)	MTX, sulfasalazine, hydroxychloroquine and cyclosporin, mean no. not specified, mean duration not specified (MTX ≥ 24 weeks)	82.5	16.3 ± 3.4	–	–	–	2.7 ± 2.6	1.3 ± 0.6	20.3 ± 8.6	12.8 ± 5.8	5.94 ± 1.86
	adalimumab + MTX (65)	95.4	48.5 ± 10.2	6.8 ± 4.2 (n = 37)	as above	76.9	16.6 ± 3.3	–	–	–	2.2 ± 2.2	1.4 ± 0.6	19.2 ± 9.2	12.2 ± 5.6	5.76 ± 1.82
(van de Putte *et al.*, 2004)	placebo (110)	77.3	53.5 ± 13.2	11.6 ± 9.3	MTX and other unspecified DMARDs, mean no. 3.6 ± 1.8 , mean duration not specified	81.8	NA	83.6	67.3	7.09 ± 0.87	5.70 ± 4.90	1.88 ± 0.64	35.5 ± 14.2	19.8 ± 9.3	7.02 ± 1.81
	adalimumab (113)	79.6	52.7 ± 13.3	10.6 ± 6.9	MTX and other unspecified DMARDs, mean no. 3.8 ± 1.8 , mean duration not specified	79.6	NA	82.3	68.1	7.07 ± 0.86	5.26 ± 3.74	1.83 ± 0.59	33.7 ± 15.9	20.5 ± 10.6	7.01 ± 1.99
ARMADA	placebo + MTX (62)	82.3	56.0 ± 10.8	11.1 ± 8.0	MTX, other DMARDs included hydroxychloroquine, mean no. 3.0, mean duration not specified	–	16.5 ± 5.0	–	58.1	–	3.1 ± 3.9	1.64 ± 0.63	28.7 ± 15.2	16.9 ± 9.5	5.72 ± 2.10
	adalimumab + MTX (67)	74.6	57.2 ± 11.4	12.2 ± 11.1	MTX, other DMARDs included hydroxychloroquine, mean no. 2.9, mean duration not specified	–	16.4 ± 4.1	–	– for individual treatment arms	–	2.1 ± 1.8	1.55 ± 0.61	28.0 ± 12.7	17.3 ± 8.6	5.30 ± 2.20
CHANGE	placebo (87)	77.0	53.4 ± 12.8	8.4 ± 8.2	All patients had ≥ 1 previous DMARD, common DMARDs included MTX, sulfasalazine, bucillamine, mean no. not specified, mean duration not specified	89.0	NA	–	–	–	5.86 ± 3.30	1.39 ± 0.75	23.7 ± 8.8	19.3 ± 7.0	6.27 ± 2.28
	adalimumab (91)	79.1	56.9 ± 10.3	9.9 ± 7.9	as above	86.2	NA	–	–	–	6.48 ± 4.45	1.64 ± 0.70	24.4 ± 10.7	19.1 ± 7.3	6.81 ± 2.10

Trial	*Interventions compared (number of patients)*	*Gender (% female)*	*Age (yrs)*	*Disease duration (yrs)*	*Prior DMARD treatment (names, mean no., mean duration)*	*Positive rheumatoid factor (%)*	*Dose MTX (mg/ week)*	*Concomitant NSAID therapy (%)*	*Concomitant glucocorticoid therapy (%)*	*Mean DAS28*	*CRP (mg /dL)*	*HAQ-DI*	*Tender joint count*	*Swollen joint count*	*Patient assessment of pain (VAS)*
DE019	adalimumab + MTX (207)	76.3	56.1 ± 13.5	11.0 ± 9.2	MTX, details of other DMARDs –, mean no 2.4, mean duration not specified (MTX ≥ 24 weeks)	81.6	16.7 ± 4.5	–	–	–	1.8 ± 2.3	1.8 ± 2.3	27.3 ± 12.7	19.3 ± 9.8	5.59 ± 2.04
	placebo + MTX (200)	73.0	56.1 ± 12.0	10.9 ± 8.8	as above	89.5	16.7 ± 4.1	–	–	–	1.8 ± 2.1	1.8 ± 2.1	28.1 ± 13.8	19.0 ± 9.5	5.63 ± 2.29
STAR	adalimumab + DMARDs (318)	79.6	55.0 ± 12.8	9.3 ± 8.8	Permitted DMARDs included MTX, gold (oral and parenteral), corticosteroids, (hydroxy)chloroquine, lefluonimide, sulfasalazine, mean no. 1.1, mean duration –	63.4	–	62.3	50.9	–	1.5 ± 2.0	1.37 ± 0.62	27.3 ± 13.0	20.9 ± 11.0	5.51 ± 2.25
	placebo + DMARDs (318)	79.2	55.8 ± 12.4	11.5 ± 9.7	Permitted DMARDs included MTX, gold (oral and parenteral), corticosteroids, (hydroxy)chloroquine, lefluonimide, sulfasalazine, mean no. 1.2, mean duration –	62.3	–	63.8	54.4	–	1.5 ± 1.9	1.43 ± 0.60	27.6 ± 13.8	21.3 ± 13.8	5.56 ± 2.25
(Chen *et al.*, 2009)	adalimumab + MTX (35)	74.3	53.0 (29.0, 75.0)[2]	6.2 (0.3, 19.2) [2]	MTX, sulfasalazine, hydroxychloroquine, cyclosporin, mean no. not specified, mean duration not specfied (MTX ≥ 4 weeks)	85.7	–	–	–	6.41 ± 0.33	–	–	–	–	–
	placebo + MTX (12)	91.7	53.0 (35.0, 73.0) [2]	8.3 (1.3, 15.6) [2]	as above	91.7	–	–	–	6.54 ± 0.42	–	–	–	–	–
FAST-4WARD	placebo (109)	89.0	54.9 ± 11.6	10.4 ± 9.6	MTX, other DMARDs not specified, mean no. 2 ± 1.25, mean duration not specified (MTX ≥ 24 weeks)	100	NA	–	58.7	6.3 ± 0.9	11.3 (8.6, 14.9)[3]	1.6 ± 0.65	28.3 ± 12.5	19.9 ± 9.3	5.48 ± 2.08

Trial	*Interventions compared (number of patients)*	*Gender (% female)*	*Age (yrs)*	*Disease duration (yrs)*	*Prior DMARD treatment (names, mean no., mean duration)*	*Positive rheumatoid factor (%)*	*Dose MTX (mg/ week)*	*Concomitant NSAID therapy (%)*	*Concomitant glucocorticoid therapy (%)*	*Mean DAS28*	*CRP (mg /dL)*	*HAQ-DI*	*Tender joint count*	*Swollen joint count*	*Patient assessment of pain (VAS)*
	certolizumab (111)	78.4	52.7 ± 12.7	8.7 ± 8.2	MTX, other DMARDs not specified, mean no. 2 ± 1.19, mean duration not specified (MTX ≥ 24 weeks)	100	NA	–	55.9	6.3 ± 1.1	11.6 (9.1, 14.9) [3]	1.4 ± 0.63	29.6 ± 13.7	21.2 ± 10.1	5.82 ± 2.19
RAPID 1	placebo + MTX (199)	83.9	52.2 ± 11.2	6.2 ± 4.4	MTX, other DMARDs not listed in detail, mean no. 1.4 ± 1.4 (excluding MTX), mean duration not specified (MTX ≥ 24 weeks)	82.8	13.4 ± 4.2	–	–	7.0 ± 0.9	1.60 (0.20, 16.2) [4]	1.7 ± 0.6	29.8 ± 13.0	21.2 ± 9.7	6.36 ± 1.99
	certolizumab 200 mg + MTX (393)	82.4	51.4 ± 11.6	6.1 ± 4.2	MTX, other DMARDs not listed in detail, mean no. 1.3 ± 1.3 (excluding MTX), mean duration not specified (MTX ≥ 24 weeks)	79.6	13.6 ± 4.3	–	–	6.9 ± 0.8	1.60 (0.10, 23.4) [4]	1.7 ± 0.6	30.8 ± 12.4	21.7 ± 9.9	6.21 ± 2.00
	certolizumab 400 mg + MTX (390)	83.6	52.4 ± 11.7	6.2 ± 4.4	MTX, other DMARDs not listed in detail, mean no. 1.3 ± 1.3 (excluding MTX), mean duration not specified (MTX ≥ 24 weeks)	83.6	13.6 ± 4.0	–	–	6.9 ± 0.8	1.40 (0.20, 27.3) [4]	1.7 ± 0.6	31.3± 13.3	21.5 ± 9.8	6.38 ± 1.72
RAPID 2	placebo + MTX (127)	84.3	51.5 ± 11.8	5.6 ± 3.9	MTX, other DMARDs not specified, mean no. 1.2 ± 1.2 (excluding MTX), mean duration not specified (MTX ≥ 24 weeks)	78.2	12.2 ± 3.3	–	59.8	6.83 ± 0.87	1.35 (18.58) [5]	1.6 ± 0.6	30.4 ± 13.4	21.9 ± 9.7	5.99 ± 2.22
	certolizumab 200 mg (246)	83.7	52.2 ± 11.1	6.1 ± 4.1	MTX, other DMARDs not specified, mean no. 1.2 ± 1.3 (excluding MTX), mean duration not specified (MTX ≥ 24 weeks)	77.5	12.5 ± 3.6	–	55.3	6.85 ± 0.84	1.42 (19.08) [5]	1.6 ± 0.6	30.1 ± 14.5	20.5 ± 9.6	6.18 ± 1.93
	certolizumab 200 mg (246)	78.0	51.9 ± 11.8	6.5 ± 4.3	MTX, other DMARDs –, mean no. 1.3 ± 1.2 (excluding MTX), mean duration not specified (MTX ≥ 24 weeks)	75.5	12.6 ± 3.7	–	61.8	6.80 ± 0.79	1.31 (16.99) [5]	1.6 ± 0.6	30.0 ± 13.9	21.0 ± 10.2	6.05 ± 2.00

Trial	*Interventions compared (number of patients)*	*Gender (% female)*	*Age (yrs)*	*Disease duration (yrs)*	*Prior DMARD treatment (names, mean no., mean duration)*	*Positive rheumatoid factor (%)*	*Dose MTX (mg/ week)*	*Concomitant NSAID therapy (%)*	*Concomitant glucocorticoid therapy (%)*	*Mean DAS28*	*CRP (mg /dL)*	*HAQ-DI*	*Tender joint count*	*Swollen joint count*	*Patient assessment of pain (VAS)*
TEMPO	placebo + MTX (228)	79	53.0 ± 12.8	6.8 ± 5.5	MTX, other DMARDs not specified, mean no 2.3 ± 1.6, mean duration not specified	71	17.2	86	64	5.5 ± 1.2	2.55 ± 2.82	1.7 ± 0.7 (n = 227)	33.1 ± 13.4	22.6 ± 10.7	–
	etanercept + placebo (223)	77	53.2 ± 13.8	6.3 ± 5.1	MTX, other DMARDs not specified, mean no 2.3 ± 1.4, mean duration not specified	75	NA	88	57	5.7 ± 1.1	3.24 ± 3.77	1.7 ± 0.7	35.0 ± 14.5	23.0 ± 10.7	–
	etanercept + MTX (231)	74	52.5 ± 12.4	6.8 ± 5.4	MTX, other DMARDs not specified, mean no 2.3 ± 1.4, mean duration not specified	76	16.9	88	62	5.5 ± 1.2	2.99 ± 3.26	1.8 ± 0.6 (n = 230)	34.2 ± 14.8	22.1 ± 11.3	–
(Combe *et al.*, 2006)	placebo + sulfasalazine (50)	82.0	53.3 ± 12.8	5.6 ± 4.4	sulfasalazine, other DMARDs not specified, mean no. not specified, mean duration not specified (sulfasalazine ≥ 4 months)	–	(2.1 ± 0.4, sulfasalazine, g/day)	–	40.0	5.0 ± 1.1	1.155[6]	1.6 ± 0.5	31.3 ± 14.0	18.65 ± 11.1	5.88 ± 2.00
	etanercept + placebo (103)	78.6	51.3 ± 13.5	7.1 ± 5.2	as above	–	(2.1 ± 0.4, sulfasalazine, g/day)	–	59.2	5.1 ± 1.1	1.43[6]	1.7 ± 0.6	29.7 ± 14.7	19.1 ± 10.1	6.26 ± 2.17
	etanercept + sulfasalazine (101)	80.2	50.6 ± 12.3	6.5 ± 5.1	as above	–	(2.1± 0.5, (sulfasalazine, g/day)	–	44.6	5.2 ± 1.2	1.16[6]	1.6 ± 0.6	31.3 ± 14.1	19.4 ± 10.4	5.85 ± 2.07
(Moreland *et al.*, 1999)	placebo (80)	76	51	12	MTX, hydroxychloroquinine, gold, sulfasalazine, azathioprine, penacillamine, mean no 3.0, mean duration not specified	79	NA	84	58	–	4.1	1.7	35	25	6.5

Trial	*Interventions compared (number of patients)*	*Gender (% female)*	*Age (yrs)*	*Disease duration (yrs)*	*Prior DMARD treatment (names, mean no., mean duration)*	*Positive rheumatoid factor (%)*	*Dose MTX (mg/ week)*	*Concomitant NSAID therapy (%)*	*Concomitant glucocorticoid therapy (%)*	*Mean DAS28*	*CRP (mg /dL)*	*HAQ-DI*	*Tender joint count*	*Swollen joint count*	*Patient assessment of pain (VAS)*
	etanercept (78)	74	53	11	MTX, hydroxychloroquinine, gold, sulfasalazine, azathioprine, penacillamine, mean no 3.3, mean duration not specified	79	NA	67	81	–	4.7	1.6	33	25	6.7
(Weinblatt *et al.*, 1999)	placebo + MTX (30)	73	53	13	MTX, other DMARDs not specified, mean no. 2.8, mean duration 35 months	90	–	80	53	–	2.6[6]	1.5[6]	28[6]	17[6]	5.6[6]
	etanercept + MTX (59)	90	48	13	MTX, other DMARDs not specified, mean no. 2.7, mean duration 58 months	84	–	75	70	–	2.2[6]	1.5[6]	28[6]	20[6]	5.6[6]
GO-FORWARD	placebo + MTX (133)	82.0	52.0 (42.0, 58.0)	6.5 (3.1, 11.9)	MTX, other DMARDs not specified, mean no. not specified, mean duration not specified (MTX ≥ 3 months)	81.2	15.0 (15.0, 20.0)[7]	–	65.4	6.111 (5.260, 6.574)	0.8 (0.3, 2.0)[7]	1.250 (0.750, 1.750)	21.0 (14.0, 34.0)	12.0 (8.0, 19.0)[7]	5.7 (3.6, 7.5)[7]
	golimumab + MTX (89)	80.9	52.0 (43.0, 57.0)	4.5 (2.1, 9.7)	as above	86.5	15.0 (15.0, 20.0)[7]	–	75.3	6.105 (5.366, 6.940)[7]	1.0 (0.4, 2.8)[7]	1.375 (1.000, 1.875)[7]	26.0 (16.0, 39.0)[7]	13.0 (8.0, 22.0)[7]	6.1 (4.7, 7.7)[7]
(Kay *et al.*, 2008)	placebo (35)	74.3	52.0 (46.0, 66.0)	5.6 (1.4, 10.9)	MTX, other DMARDs not specified, mean no. not specified, mean duration not specified (MTX ≥ 3 months)	–	–	–	–	6.3 (5.7, 7.0)[7]	2.0 (1.3, 3.4)[7]	1.3 (0.9, 1.9)[7]	22 (16, 38)[7]	13 (10, 18)[7]	7.0 (5.1, 7.9)[7]
	golimumab every 4 weeks + MTX (35)	85.7	57.0 (50.0, 64.0)	8.2 (4.1, 14.3)	as above	–	–	–	–	6.4 (5.6, 7.3)[7]	2.1 (1.2, 3.4)[7]	1.7 (1.4, 2.0)[7]	28 (18, 40)[7]	14 (10, 21)[7]	7.0 (6.3, 8.6)[7]
ATTEST	abatacept + MTX (156)	83.3	49.0 ± 12.5	7.9 ± 8.5	MTX, other DMARDs not specified, mean no. not specified, mean duration not specified (MTX ≥ 3 months)	87.2	16.5 ± 3.7	85.3	75.6	6.9 ± 1.0	3.1 ± 2.7	1.8 ± 0.6	31.6 ± 13.9	21.3 ± 8.6	–
	placebo + MTX (110)	87.3	49.4 ± 11.5	8.4 ± 8.6	as above	77.3	16.6 ± 3.7	84.5	70	6.8 ± 1.0	2.7 ± 2.6	1.8 ± 0.7	30.3 ± 11.7	20.1± 7.0	–

Trial	*Interventions compared (number of patients)*	*Gender (% female)*	*Age (yrs)*	*Disease duration (yrs)*	*Prior DMARD treatment (names, mean no., mean duration)*	*Positive rheumatoid factor (%)*	*Dose MTX (mg/ week)*	*Concomitant NSAID therapy (%)*	*Concomitant glucocorticoid therapy (%)*	*Mean DAS28*	*CRP (mg /dL)*	*HAQ-DI*	*Tender joint count*	*Swollen joint count*	*Patient assessment of pain (VAS)*
	infliximab + MTX (165)	82.4	49.1 ± 12.0	7.3 ± 6.2	as above	84.8	16.3 ± 3.6	86.1	71.5	6.8 ± 0.9	3.3 ± 3.2	1.7 ± 0.7	31.7 ± 14.5	20.3 ± 8.0	–
ATTRACT	infliximab every 4 weeks + MTX (86)	77	52 ± 13	9 ± 8	MTX, other DMARDs not specified, mean no. 2.6 ± 1.5 (excluding MTX), mean duration not specified (MTX ≥ 3 months)	80	16 ± 4	76	54	–	3.5 ± 4.2	1.7 ± 0.6	31 ± 15	21 ± 11	–
	infliximab every 8 weeks + MTX (86)	81	54 ± 11	10 ± 8	MTX, other DMARDs not specified, mean no. 2.8 ± 1.5 (excluding MTX), mean duration not specified (MTX ≥ 3 months)	84	16 ± 4	79	63	–	3.9 ± 3.4	1.8 ± 0.6	32 ± 18	22 ± 12	–
	placebo + MTX (88)	80	51 ± 12	11 ± 8	MTX, other DMARDs not specified, mean no. 2.5 ± 1.4 (excluding MTX), mean duration not specified (MTX ≥ 3 months)	77	16 ± 4	72	64	–	4.0 ± 4.2	1.7 ± 0.6	31 ± 18	21 ± 12	–
START	infliximab + MTX (360)	80.0	53.0 (45, 61)[7]	7.8 (3, 15)[7]	MTX, chloroquine, sulfasalzine, leflunomide, gold, azathioprine, cyclosporine, mean no. not specified, mean duration not specified (MTX ≥ 3 months)	82.8	15.0 (10–18)[7]	43.3	59.2	–	1.6 (1, 3)[7]	1.5 (1, 2)[7]	22 (15, 31)[7]	15 (11, 21)[7]	6.1 (5, 8)[7]
	placebo + MTX (361	83.2	52.0 (44, 61)[7]	8.4 (4, 15)[7]	as above	80.7	15.0 (10–15)[7]	39.4	59.2	–	1.2 (1, 3)[7]	1.5 (1, 2)[7]	22 (15, 32)[7]	15 (10, 21)[7]	5.9 (5, 7)[7]
(Abe *et al.*, 2006)	infliximab + MTX (49)	81.6	55.2 ± 10.9	9.1 ± 7.4	MTX, other DMARDs not specified, mean no. not specified, mean duration not specified (MTX ≥ 3 months)	–	7.1 ± 1.9	89.8	85.7	–	4.2 ± 3.1	–	19.0 ± 11.8	15.1 ± 9.0	–

Trial	*Interventions compared (number of patients)*	*Gender (% female)*	*Age (yrs)*	*Disease duration (yrs)*	*Prior DMARD treatment (names, mean no., mean duration)*	*Positive rheumatoid factor (%)*	*Dose MTX (mg/ week)*	*Concomitant NSAID therapy (%)*	*Concomitant glucocorticoid therapy (%)*	*Mean DAS28*	*CRP (mg /dL)*	*HAQ-DI*	*Tender joint count*	*Swollen joint count*	*Patient assessment of pain (VAS)*
	placebo + MTX (47)	74.5	55.1 ± 7.6	7.5 ± 5.0	as above	–	7.4 ± 2.2	95.7	89.4	–	4.1 ± 2.4	–	17.8 ± 8.7	13.5 ± 7.6	–
(Maini *et al.*, 1998)	infliximab + placebo (14)	86	47.0 ± 15.0	7.8 ± 4.3	MTX, other DMARDs not specified, mean no. 2.5, mean duration not specified (MTX ≥ 24 weeks)	85.7	10.0 (7.5, 12.5)[7]	–	50.0	–	1.8 (1.5, 5.8)[7]	1.8 (1.5, 2.0)[7]	31 (23, 39)[7]	17 (11, 32)[7]	5.9 (3.5, 7.6)[7]
	infliximab + MTX (15)	67	58.9 ± 10.0	12.1 ± 9.0	MTX, other DMARDs not specified, mean no. 2.0, mean duration not specified (MTX ≥ 24 weeks)	66.7	10.0 (7.5, 15.0)[7]	–	60.0	–	4.2 (2.3, 6.6)[7]	2.0 (1.5, 2.5)[7]	21 (12, 31)[7]	16 (13, 22)[7]	6.0 (5.5, 7.7)[7]
	placebo + MTX (14)	71	48.8 ± 12.3	7.6 ± 4.0	MTX, other DMARDs not specified, mean no. 2.0, mean duration not specified (MTX ≥ 24 weeks)	76.9	15 (10.0, 15.0)[7]	–	50.0	–	5.1 (3.5, 5.8)[7]	2.0 (1.6, 2.1)[7]	28 (22, 47)[7]	17 (12, 25)[7]	6.7 (5.9, 8.2)[7]
(Edwards *et al.*, 2004)	placebo + MTX (40)	80	54 ± 11	11 ± 7	MTX, other DMARDs not specified, mean no. 2.6 ± 1.3, mean duration not specified (MTX ≥ 3 months)	–	–	–	–	6.9 ± 0.75	3.2 ± 4.3	2.0 ± 0.5	32 ± 13	19 ± 10	6.26 ± 1.61
	rituximab + placebo (40)	73	54 ± 10	9 ± 6	MTX, other DMARDs not specified, mean no. 2.5 ± 1.6, mean duration not specified (MTX ≥ 3 months)	–	–	–	–	6.8 ± 0.97	2.6 ± 2.2	2.0 ± 0.6	34 ± 15	21 ± 11	6.20 ± 2.02
	rituximab + cyclophosphamide (41)	83	53 ± 10	10 ± 6	MTX, other DMARDs not specified, mean no. 2.6 ± 1.4, mean duration not specified (MTX ≥ 3 months)	–	–	–	–	6.9 ± 0.84	4.0 ± 4.0	1.8 ± 0.7	33 ± 14	19 ± 10	5.75 ± 2.00
	rituximab + MTX (40)	75	54 ± 12	12 ± 7	MTX, other DMARDs not specified, mean no. 2.5 ± 1.4, mean duration not specified (MTX ≥ 3 months)	–	–	–	–	6.8 ± 0.92	2.9 ± 3.2	1.8 ± 0.6	32 ± 16	23 ± 13	5.46 ± 1.78

Trial	*Interventions compared (number of patients)*	*Gender (% female)*	*Age (yrs)*	*Disease duration (yrs)*	*Prior DMARD treatment (names, mean no., mean duration)*	*Positive rheumatoid factor (%)*	*Dose MTX (mg/ week)*	*Concomitant NSAID therapy (%)*	*Concomitant glucocorticoid therapy (%)*	*Mean DAS28*	*CRP (mg /dL)*	*HAQ-DI*	*Tender joint count*	*Swollen joint count*	*Patient assessment of pain (VAS)*
OPTION	tocilizumab + MTX (205)	85.4	50.8	7.5 ± 7.3	MTX, other DMARDs not specified, mean no. 1.5 ± 1.4 excluding MTX, mean duration not specified (MTX ≥ 12 weeks)	83	14.5 ± 4.4	66	55	6.8 ± 0.9	2.6 ± 2.6	1.6 ± 0.6	31.9 ± 15.5	19.5 ± 11.3	5.99 ± 2.24
	placebo + MTX (204)	77.9	50.6	7.8 ± 7.2	MTX, other DMARDs not specified, mean no. 1.7 ± 1.5 excluding MTX, mean duration not specified (MTX ≥ 12 weeks)	71	14.8 ± 4.2	68	54	6.8 ± 0.9	2.4 ± 2.8	1.5 ± 0.6	32.8 ± 16.1	20.7 ± 11.7	5.73 ± 2.22
SATORI	tocilizumab + placebo (61)	90.2	52.6 ± 10.6	8.5 ± 8.4	MTX, other DMARDs not specified, mean no. 3.3 (range 1–8), mean duration not specified (MTX ≥ 8 weeks)	–	8.0	–	–	6.1 ± 0.9	3.0 ± 2.0	–	13.8 ± 7.5	12.4 ± 5.9	–
	placebo + MTX (64)	75.0	50.8 ± 12.2	8.7 ± 7.1	MTX, other DMARDs not specified, mean no. 3.6 (range 1–8), mean duration not specified (MTX ≥ 8 weeks)	–	8.0	–	–	6.2 ± 0.9	3.2 ± 2.6	–	14.2 ± 8.6	12.7 ± 7.5	–
TOWARD	tocilizumab + DMARDs (803)	81	53 ± 13	9.8 ± 8.8	MTX, chloroquine, sulfasalzine, leflunomide, gold, azathioprine, mean no 1.6 ± 1.6, mean duraction not specified (all DMARDs ≥ 8 weeks)	–	14.7	71.4	51.2	6.7 ± 1.0	2.6 ± 3.2	1.5 ± 0.6	30.1 ± 16.0	19.7 ± 11.6	5.8 ± 2.3
	placebo + DMARDs (413	84	54 ± 13	9.8 ± 9.1	as above	–	15.0	77.1	54.6	6.6 ± 1.0	2.6 ± 4.7	1.5 ± 0.6	29.1 ± 14.8	18.7 ± 10.8	5.9 ± 2.3

Data presented as mean ± standard deviation unless otherwise indicated. [1]mean (range); [2]median (range); [3]geometric mean (95% CI); [4]median (minimum, maximum)
[5]mean (coefficient of variation); [6]median; [7]median (interquartile range)

Table 17. Characteristics of participants in the RCTs across randomised groups (TNFα experienced population)

Trial	*Interventions compared (N, number of patients)*	*Gender (% female)*	*Age (yrs)*	*Disease duration (yrs)*	*Prior DMARD treatment (names, mean no., mean duration)*	*Positive rheumatoid factor (%)*	*Dose MTX (mg/ week)*	*Concomitant NSAID therapy (%)*	*Concomitant glucocorticoid therapy (%)*	*Mean DAS28*	*CRP (mg/ dL)*	*HAQ-DI*	*Tender joint count*	*Swollen joint count*	***Patient's assessment of pain (VAS)***
ATTAIN	abatacept + DMARDs (256)	77.1	53.4 ± 12.4	12.2 ± 8.5	MTX, azathioprine, penicillamine, gold, [hydroxy]chloroquine, leflunomide, sulfasalazine, or anakinra, mean no. not specified, mean duration not specified (all DMARDs ≥ 3 months)	73.3	15.2 ± 5.3	70.2	70.2	6.5 ± 0.9	4.6 ± 4.0	1.8 ± 0.6	31.2 ± 13.0	22.3 ± 10.2	7.08 ± 1.98
	placebo + DMARDs (133)	79.7	52.7 ± 11.3	11.4 ± 8.9	as above	72.9	14.4 ± 6.1	71.4	64.7	6.5 ± 0.8	4.0 ± 3.6	1.8 ± 0.6	32.8 ± 13.4	22.0 ± 10.0	6.99 ± 1.90
ASSURE	abatacept + DMARDs (103)	75.7	54.6 ± 11.2	11.3 ± 8.9	MTX, (hydroxy)chloroquine, sulfasalazine, leflunomide, gold, azathioprine, etanercept, infliximab, adalimumab, anakinra, mean no. not specified, mean duration not specified (≥ 3 months)	–	–	–	74.8	–	1.4 ± 1.9	1.5 ± 0.6	–	–	6.22 ± 2.03
	placebo + DMARDs (64)	75.0	52.8 ± 11.4	11.3 ± 9.6	as above	–	–	–	79.7	–	1.5 ± 1.9	1.6 ± 0.6	–	–	6.15 ± 2.00
GO-AFTER	golimumab (153)	74	55.0 (46.0, 63.0)[1]	9.6 (5.6, 17.2)[1]	MTX, sulfasalazine, hydroxychloroquine, mean no. not specified, mean duration not specified (≥ 3 months)	72	–	–	–	6.3 (5.6, 7.2)[1]	8 (3, 27)[1]	1.6 (1.1, 2.0)[1]	27.0 (16.0, 42.0)[1]	14.0 (9.0, 25.0)[1]	6.9 (5.3, 8.8)[1]
	placebo (155)	85	54.0 (46.0, 64.0)[1]	9.8 (4.9, 17.6)[1]	as above	73	–	–	–	6.3 (5.5, 7.1)[1]	10 (3, 21)[1]	1.8 (1.3, 2.1)[1]	26.0 (15.0, 43.0)[1]	14.0 (9.0, 23.0)[1]	7.0 (5.2, 8.4)[1]
REFLEX	rituximab + MTX (298)	81	52.2 ± 12.2	12.1 ± 8.3	MTX, other DMARDs not specified, mean no. 2.6 ± 1.8 (excluding MTX), mean duration not specified (MTX ≥ 12 weeks)	79	16.4 ± 8.8	–	65	6.9 ± 1.0	3.7 ± 3.8	1.9 ± 0.6	33.9 ± 15.1	23.4 ± 11.8	–
	placebo + MTX (201)	81	52.8 ± 12.6	11.7 ± 7.7	MTX, other DMARDs not specified, mean no. 2.4 ± 1.8 (excluding MTX), mean duration not specified (MTX ≥ 12 weeks)	79	16.7 ± 9.9	–	61	6.8 ± 1.0	3.8 ± 4.1	1.9 ± 0.5	33.0 ± 15.6	22.9 ± 12.7	–

RADIATE	tocilizumab + MTX (170)	84	53.9 ± 12.7	12.6 ± 9.3	MTX, leflunomide, other conventional DMARDs not specified, etanercept, adalimumab, infliximab, mean no. 1.9 ± 1.7, mean duration – (MTX ≥ 12 weeks)	79	15.7 ± 4.4	–	52	6.79 ± 0.93	2.80 ± 3.37	1.7 ± 0.6	31.7 ± 15.4	18.9 ± 10.9	6.47 ± 2.06
	placebo + MTX (158)	79	53.4 ± 13.3	11.4 ± 9.2	MTX, leflunomide, other conventional DMARDs not specified, etanercept, adalimumab, infliximab, mean no. 2.1 ± 1.6, mean duration not specified (MTX ≥ 12 weeks)	75	16.5 ± 4.8	–	58	6.80 ± 1.06	3.71 ± 4.12	1.7 ± 0.6	30.4 ± 16.8	18.9 ± 11.1	6.41 ± 2.18

Data presented as mean ± standard deviation unless otherwise indicated. [1]median (interquartile range)

Outcomes

5.3.5 Provide details of the outcomes investigated and the measures used to assess those outcomes. Indicate which outcomes were specified in the trial protocol as primary or secondary, and whether they are relevant with reference to the decision problem. This should include therapeutic outcomes, as well as patient-related outcomes such as assessment of health-related quality of life, and any arrangements to measure compliance. Data provided should be from pre-specified outcomes rather than post-hoc analyses. When appropriate, also provide evidence of reliability or validity, and current status of the measure (such as use within UK clinical practice). The following table provides a suggested format for presenting primary and secondary outcomes when there is more than one RCT.

Primary and secondary outcomes of the RCTs are available in 8.14 in Table 178.

Statistical analysis and definition of study groups

5.3.6 State the primary hypothesis or hypotheses under consideration and the statistical analysis used for testing hypotheses. Also provide details of the power of the study and a description of sample size calculation, including rationale and assumptions. Provide details of how the analysis took account of patients who withdrew (for example, a description of the intention-to-treat analysis undertaken, including censoring methods; whether a per-protocol analysis was undertaken). The following table provides a suggested format for presenting the statistical analyses in the trials when there is more than one RCT.

Summary of statistical analyses in RCTs are available in Appendix 8.14 in Table 179.

5.3.7 Provide details of any subgroup analyses that were undertaken and specify the rationale and whether they were pre-planned or post-hoc.

For golimumab, subgroup analyses were conducted based on demographic features, geographic region, baseline disease characteristics, and baseline medications for RA in GO-FORWARD and GO-AFTER trials. These analyses were pre-planned. Separate post-hoc analyses were conducted comparing individual golimumab doses with placebo on some of the baseline demographics and disease characteristics.

Participant flow

5.3.8 Provide details of the numbers of patients who were eligible to enter the RCT(s), randomised, and allocated to each treatment. Provide details of, and the rationale for, patients who crossed over treatment groups and/or were lost to follow-up or withdrew from the RCT. This information should be presented as a CONSORT flow chart.

Participant flow diagrams are available in Appendix 8.14 in Figure 7.

5.4 Critical appraisal of relevant RCTs

5.4.1 The validity of the results of an individual study will depend on the robustness of its overall design and execution, and its relevance to the decision problem. Each study that meets the criteria for inclusion should therefore be critically appraised. Whenever possible, the criteria for assessing published studies should be used to assess the validity of unpublished and part-published studies. The critical appraisal will be validated by the ERG. The following are the minimum criteria for assessment of risk of bias in RCTs, but the list is not exhaustive.

- Was the method used to generate random allocations adequate?
- Was the allocation adequately concealed?
- Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?
- Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?
- Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?
- Is there any evidence to suggest that the authors measured more outcomes than they reported?
- Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?

5.4.2 Please provide as an appendix a complete quality assessment for each RCT. See section 9.3, appendix 3 for a suggested format.

The quality assessment for the GO FORWARD and GO AFTER trials is available in Appendix 8.3 , in Table 169 and Table 170 respectively.

5.4.3 If there is more than one RCT, tabulate a summary of the responses applied to each of the critical appraisal criteria. A suggested format for the quality assessment results is shown below.

The quality assessment for all trials is available in Appendix 8.3 in Table 171.

5.5 Results of the relevant RCTs

5.5.1 Provide the results for all relevant outcome measure(s) pertinent to the decision problem. Data from intention-to-treat analyses should be presented whenever possible and a definition of the included patients provided. If patients have been excluded from the analysis, the rationale for this should be given. **If there is more than one RCT, tabulate the responses.**

Please refer to 5.5.3.

5.5.2 The information may be presented graphically to supplement text and tabulated data. If appropriate, please present graphs such as Kaplan-Meier plots.

Not applicable.

5.5.3 For each outcome for each included RCT, the following information should be provided.

- The unit of measurement.
- The size of the effect; for dichotomous outcomes, the results ideally should be expressed as both relative risks (or odds ratios) and risk (or rate) differences. For time-to-event analysis, the hazard ratio is an equivalent statistic. Both absolute and relative data should be presented.
- A 95% confidence interval.
- Number of participants in each group included in each analysis and whether the analysis was by 'intention to treat'. State the results in absolute numbers when feasible.
- When interim RCT data are quoted, this should be clearly stated, along with the point at which data were taken and the time remaining until completion of that RCT. Analytical adjustments should be described to cater for the interim nature of the data.
- Other relevant data that may assist in interpretation of the results may be included, such as adherence to medication and/or study protocol.
- Discuss and justify definitions of any clinically important differences.
- Report any other analyses performed, including subgroup analysis and adjusted analyses, indicating those pre-specified and those exploratory.

DMARD Experienced Population

ACR20 response at 24 weeks (DMARD experienced population)

The analysis of ACR20 response at 24 weeks included 25 studies for the conventional DMARD experienced population: 2 golimumab studies, 2 abatacept studies, 6 adalimumab studies, 2 certolizumab pegol studies, 4 etanercept studies, 4 infliximab studies, 1 rituximab study and 3 tocilizumab studies. Three trials did not report ACR20 response at 24 weeks: ASSURE (abatacept), Chen *et al.* (2009) (adalimumab) and Abe *et al.* (2006) (infliximab). The data extracted from the 25 studies that contributed to the present analysis are presented in Table 18.

Table 18. ACR20 response at 24 weeks: data contributing to the analysis (DMARD experienced population)

Trial	*Intervention*	*Placebo*		*Study drug*	
		Patients with positive response	*Number in treatment arm*	*Patients with positive response*	*Number in treatment arm*
GO-FORWARD	Golimumab	37	133	53	89
(Kay *et al.*, 2008)	Golimumab	16	35	26	35
(Kremer *et al.*, 2003)	Abatacept	42	119	69	115
AIM	Abatacept	87	219	294	433
ATTEST	Abatacept	46	110	104	156
(Kim *et al.*, 2007)	Adalimumab	23	63	40	65
(van de Putte *et al.*, 2004)	Adalimumab + placebo	21	110	52	113
ARMADA	Adalimumab	9	62	45	67
CHANGE	Adalimumab + placebo	12	87	40	91
STAR	Adalimumab	111	318	168	318
DE019	Adalimumab	59	200	131	207
RAPID 1	Certolizumab pegol	27	199	231	393
RAPID 2	Certolizumab pegol	11	127	141	246
(Combe *et al.*, 2006)	Etanercept + sulfasalazine	14	50	74	101
(Combe *et al.*, 2006)	Etanercept + placebo	14	50	76	103
TEMPO	Etanercept + placebo	164	228	156	223
TEMPO	Etanercept	164	228	187	231
(Weinblatt *et al.*, 1999)	Etanercept	27	30	71	59
(Moreland *et al.*, 1999)	Etanercept + placebo	9	80	46	78
(Maini *et al.*, 1998)	Infliximab + placebo	1	14	4	14

Trial	*Intervention*	*Placebo*		*Study drug*	
		Patients with positive response	*Number in treatment arm*	*Patients with positive response*	*Number in treatment arm*
(Maini *et al.*, 1998)	Infliximab	1	14	6	15
START	Infliximab	87	361	199	360
ATTRACT	Infliximab (3mg/kg/4 weeks)	18	88	43	86
ATTRACT	Infliximab (3mg/kg/8 weeks)	18	88	47	86
ATTEST	Infliximab	46	110	98	165
(Edwards *et al.*, 2004)	Rituximab + placebo	15	40	26	40
(Edwards *et al.*, 2004)	Rituximab + cyclophosphoamide	15	40	31	41
(Edwards *et al.*, 2004)	Rituximab	15	40	29	40
OPTION	Tocilizumab	54	204	120	205
SATORI	Tocilizumab + placebo	16	64	49	61
TOWARD	Tocilizumab	101	413	488	803

ACR50 response at 24 weeks (DMARD experienced population)

The analysis of ACR50 response at 24 weeks included a total of 25 studies for the conventional DMARD-experienced population: 2 golimumab studies, 2 abatacept studies, 6 adalimumab studies, 2 certolizumab pegol studies, 4 etanercept studies, 4 infliximab studies, 1 rituximab study and 3 tocilizumab studies. Three trials did not report ACR20 response at 24 weeks: ASSURE (abatacept), Chen *et al.* (2009) (adalimumab) and Abe *et al.* (2006) (infliximab). The data extracted from the 25 studies which were included in the analysis are presented in Table 19.

Table 19. ACR50 response at 24 weeks: data contributing to the analysis (DMARD experienced population)

Trial	*Intervention*	*Placebo*		*Study drug*	
		Patients with positive response	*Number in treatment arm*	*Patients with positive response*	*Number in treatment arm*
GO-FORWARD	Golimumab	18	133	33	89
(Kay *et al.*, 2008)	Golimumab	4	35	14	35
(Kremer *et al.*, 2003)	Abatacept	14	119	42	115
AIM	Abatacept	37	219	173	433
ATTEST	Abatacept	22	110	63	156

Trial	*Intervention*	*Placebo*		*Study drug*	
		Patients with positive response	*Number in treatment arm*	*Patients with positive response*	*Number in treatment arm*
(Kim *et al.*, 2007)	Adalimumab	9	63	28	65
(van de Putte *et al.*, 2004)	Adalimumab + placebo	9	110	25	113
ARMADA	Adalimumab	5	62	37	67
CHANGE	Adalimumab + placebo	5	87	22	91
STAR	Adalimumab	36	318	92	318
DE019	Adalimumab	19	200	81	207
RAPID 1	Certolizumab pegol	15	199	146	393
RAPID 2	Certolizumab pegol	4	127	80	246
(Combe *et al.*, 2006)	Etanercept + sulfasalazine	7	50	48	103
(Combe *et al.*, 2006)	Etanercept + placebo	7	50	52	101
TEMPO	Etanercept + placebo	91	228	89	223
TEMPO	Etanercept	91	228	136	231
(Weinblatt *et al.*, 1999)	Etanercept	3	30	39	59
(Moreland *et al.*, 1999)	Etanercept + placebo	4	80	31	78
(Maini *et al.*, 1998)	Infliximab + placebo	0	14	2	14
(Maini *et al.*, 1998)	Infliximab	0	14	5	15
START	Infliximab	33	361	110	360
ATTRACT	Infliximab (3mg/kg/4 weeks)	4	88	25	86
ATTRACT	Infliximab (3mg/kg/8 weeks)	4	88	22	86
ATTEST	Infliximab	22	110	61	165
(Edwards *et al.*, 2004)	Rituximab + placebo	5	40	13	40
(Edwards *et al.*, 2004)	Rituximab + cyclophosphoamide	5	40	17	41
(Edwards *et al.*, 2004)	Rituximab	5	40	17	40
OPTION	Tocilizumab	22	204	90	205
SATORI	Tocilizumab + placebo	7	64	30	61
TOWARD	Tocilizumab	37	413	301	803

TNFα Experienced Population

ACR20 response at 24 weeks (TNFα experienced population)

The analysis of ACR20 response at 24 weeks included 4 studies for the TNFα experienced population: 1 golimumab study, 1 abatacept study, 1 rituximab study and 1 tocilizumab study. The ASSURE study was not included because no ACR20 response data were reported. The data extracted from the 4 studies which contributed to the present analysis are presented in Table 22.

Table 20. ACR20 response at 24 weeks: data contributing to the analysis (TNFα experienced population)

Trial	Intervention	Placebo		Study drug	
		Patients with positive response	*Number in treatment arm*	*Patients with positive response*	*Number in treatment arm*
GO-AFTER	Golimumab +/- MTX	26	155	52	153
ATTAIN	Abatacept + cDMARDs	26	133	129	258
REFLEX	Rituximab	36	201	152	298
RADIATE	Tocilizumab	16	158	85	170

ACR50 response at 24 weeks (TNFα experienced population)

The analysis of ACR50 response at 24 weeks for the TNFα experienced population included 4 studies: 1 golimumab study, 1 abatacept study, 1 rituximab study and 1 tocilizumab study. The ASSURE study was not included because no ACR50 response data were reported. The data extracted from the 4 studies which contributed to the present analysis are presented in Table 21. ACR50 response at 24 weeks: data contributing to the analysis

Table 21. ACR50 response at 24 weeks: data contributing to the analysis (TNFα experienced population)

Trial	*Intervention*	*Placebo*		*Intervention*	
		Patients with positive response	*Number in treatment arm*	*Patients with positive response*	*Number in treatment arm*
GO-AFTER	Golimumab +/- MTX	8	155	28	153
ATTAIN	Abatacept + cDMARDs	5	133	52	258
REFLEX	Rituximab	10	201	80	298
RADIATE	Tocilizumab	6	158	49	170

Meta-analysis

5.5.4 The following steps should be used as a minimum when presenting a meta-analysis.

- Perform a statistical assessment of heterogeneity. If the visual presentation and/or the statistical test indicate that the RCT results are heterogeneous, try to provide an explanation for the heterogeneity.
- Statistically combine (pool) the results for both relative risk reduction and absolute risk reduction using both the fixed effects and random effects models (giving four combinations in all).
- Provide an adequate description of the methods of statistical combination and justify their choice.
- Undertake sensitivity analysis when appropriate.
- Tabulate and/or graphically display the individual and combined results (such as through the use of forest plots).

Meta-analyses methodology

The summary measures (RR for dichotomous outcomes, WMD for continuous outcomes) and associated precisions are calculated for each outcome in each study. These are then pooled to produce a pooled estimate of the effect of each treatment versus placebo for each outcome. Pooling is done using both fixed-effect and random-effect models. The fixed-effect model is run using the Mantel-Haenszel method. The random-effect model is run using the DerSimonian and Laird method, with the estimate of heterogeneity being taken from the Mantel-Haenszel model.

Exclusion of monotherapy trials

In the majority of studies reporting data for the conventional DMARD experienced population the biologic DMARD was administered with MTX. However, in four studies no concomitant MTX was permitted (van der Putte *et al.* [2004], CHANGE, Moreland *et al.* [1999] and SATORI), and three studies included a monotherapy treatment arm (Combe *et al.* [2006], TEMPO and Maini *et al.* [1998]). To investigate the effect of the small group of monotherapy studies, and monotherapy treatment arms on the RR estimate additional fixed- and random-effects meta-analyses were performed. The RR of the monotherapy group versus the original group (all studies) was calculated.

Exclusion of RCTs from meta-analyses

Although the systematic literature review identified trials for all biologics, only those that are licensed for the particular patient population and have received NICE Technology Appraisal Guidance as of the submission date were included within the meta-analyses.

Rituximab, abatacept and tocilizumab trials were excluded from the DMARD experienced meta-analyses as NICE TAG is still pending and thus final positioning

within the treatment sequence and reimbusement in the UK is unknown as of submission date.

Trials for abatacept and tocilizumab were excluded from the TNFα inhibitor experienced meta-analyses for the same reasons explained above. The excluded trials are presented in Table 51.

ACR20 at 24 weeks meta-analyses (DMARD experienced population)

Adalimumab

Table 22 presents the adalimumab studies which were included within the meta-analyses. Table 23 presents the global analysis of the adalimumab meta-analyses. There is a high amount of heterogeneity in this meta-analysis, which is highly significant. The random-effect model is therefore the most appropriate. It shows that patients on adalimumab are 2.22 times more likely to achieve an ACR20 response at 6 months than patients on placebo. Table 24 presents the meta-analyses results for adalimumab excluding the monotherapy arms. Both meta-analyses found adalimumab to be statistically superior (p<0.001) to placebo.

Table 22. Adalimumab studies included within meta-analysis (ACR20 at 24 wks in DMARD experienced population)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
ARMADA	ADA	placebo	45	67	9	62
CHANGE (monotherapy)	ADA	placebo	40	91	12	87
DE019	ADA	placebo	131	207	59	200
Kim	ADA	placebo	40	65	23	63
STAR	ADA	placebo	168	318	111	318
Van de Putte (monotherapy)	ADA	placebo	52	113	21	110

Table 23. Adalimumab meta-analysis results (ACR20 at 24 wks in DMARD experienced population)

Study	**RR**	**95% CI**	**Weights fixed-effect meta-analysis (%)**	**Weights random-effect meta-analysis (%)**
ARMADA	4.63	2.47, 8.66	3.9	11.2
CHANGE	3.19	1.79, 5.66	5.2	12.3
DE019	2.15	1.69, 2.72	25.3	21.1
Kim	1.69	1.15, 2.46	9.8	17.2
STAR	1.51	1.26, 1.82	46.8	22.4
Van de Putte	2.41	1.56, 3.72	9.0	15.7
Pooled RR			**1.98 (1.75, 2.24)**	**2.22 (1.67, 2.95)**
p-value pooled RR			**<0.001**	**<0.001**
Heterogeneity			**I^2=74.9%, chi-square p-value=<0.001**	

Table 24. Adalimumab meta-analysis excluding the monotherapy arms (ACR20 at 24 wks in DMARD experienced population)

Study	RR	95% CI	Weights fixed-effect meta-analysis (%)	Weights random-effect meta-analysis (%)
ARMADA	4.63	2.47, 8.66	4.6	15.7
DE019	2.15	1.69, 2.72	29.5	29.2
Kim	1.69	1.15, 2.46	11.5	23.9
STAR	1.51	1.26, 1.82	54.5	31.1
Pooled RR			**1.86 (1.63, 2.13)**	**2.05 (1.46, 2.87)**
p-value pooled RR			**<0.001**	**<0.001**
Heterogeneity			**I^2=79.6%, chi-square p-value=0.002**	

Certolizumab

Table 25 presents the certolizumab studies which were included within the meta-analyses. Table 26 presents the global analysis of the certolizumab meta-analyses. There is a moderate amount heterogeneity in this meta-analysis, which does not reach significance. The fixed-effect is therefore the preferred one in this situation. It shows that patients on certolizumab are 5 times more likely to achieve an ACR20 response at 6 months than patients on placebo. The meta-analyses found certolizumab to be statistically superior (p<0.001) to placebo.

Table 25. Certolizumab studies included within meta-analysis (ACR20 at 24 wks in DMARD experienced population)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
RAPID 1	CTZ	placebo	231	393	27	199
RAPID 2	CTZ	placebo	141	246	11	127

Table 26. Certolizumab meta-analysis results (ACR20 at 24 wks in DMARD experienced population)

Study	RR	95% CI	Weights fixed-effect meta-analysis (%)	Weights random-effect meta-analysis (%)
RAPID 1	4.33	3.02, 6.21	71.2	64.4
RAPID 2	6.62	3.72, 11.76	28.8	35.6
Pooled RR			**4.99 (3.66, 6.78)**	**5.04 (3.38, 7.52)**
p-value pooled RR			**<0.001**	**<0.001**
Heterogeneity			**I^2=34.2%, chi-square p-value=0.218**	

Etanercept

Table 27 presents the etanercept studies which were included within the meta-analyses. Table 28 presents the global analysis of the etanercept meta-analyses. There is a high amount of heterogeneity in this meta-analysis, which is highly significant. The random-effect model is therefore the most appropriate. It shows that patients on etanercept are 2.43 times more likely to achieve an ACR20 response at 6 months than patients on placebo. However this estimate just fails to reach significance, as the high amount of heterogeneity translates into a large standard error (reflecting uncertainty), and therefore a large confidence interval.

Table 29 presents the meta-analyses results for etanercept excluding the monotherapy arms. There is a high amount of heterogeneity in this meta-analysis, which is highly significant. The random-effect model is therefore the most appropriate. It shows that patients on etanercept are 1.93 times more likely to achieve an ACR20 response at 6 months than patients on placebo. However this estimate just fails to reach significance, as the high amount of heterogeneity translates into a large standard error (reflecting uncertainty), and therefore a large confidence interval.

Table 27. Etanercept studies included within meta-analysis (ACR20 at 24 wks in DMARD experienced population)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
Combe	ETN	placebo	150	204	14	50
Combe (excluding monotherapy)	ETN	placebo	74	101	14	50
Moreland (monotherapy)	ETN	placebo	46	78	9	80
TEMPO	ETN	placebo	343	454	164	228
TEMPO (excluding monotherapy)	ETN	placebo	187	231	164	228
Weinblatt	ETN	placebo	42	59	8	30

Table 28. Etanercept meta-analysis results (ACR20 at 24 wks in DMARD experienced population)

Study	**RR**	**95% CI**	**Weights fixed-effect meta-analysis (%)**	**Weights random-effect meta-analysis (%)**
Combe	2.63	1.67, 4.13	8.6	25.3
Moreland	5.24	2.76, 9.97	3.4	23.8
TEMPO	1.05	0.95, 1.16	83.9	26.9
Weinblatt	2.67	1.44, 4.94	4.1	24.0
Pooled RR			**1.40 (1.26, 1.55)**	**2.43 (0.97, 6.07)**
p-value pooled RR			**<0.001**	**0.058**
Heterogeneity			**I^2=95.1%, chi-square p-value=<0.001**	

Table 29. Etanercept meta-analysis excluding the monotherapy arms (ACR20 at 24 wks in DMARD experienced population)

Study	**RR**	**95% CI**	**Weights fixed-effect meta-**	**Weights random-effect meta-**

			analysis (%)	analysis (%)
Combe	2.62	1.65, 4.14	9.6	32.9
TEMPO	1.13	1.02, 1.25	84.9	36.9
Weinblatt	2.67	1.44, 4.94	5.5	30.2
Pooled RR			**1.35 (1.21, 1.51)**	**1.93 (0.88, 4.22)**
p-value pooled RR			**<0.001**	**0.100**
Heterogeneity			**I^2=92.0%, chi-square p-value=<0.001**	

Golimumab

Table 30 presents the golimumab studies which were included within the meta-analyses. Table 31 presents the global analysis of the golimumab meta-analyses. There is a low level of heterogeneity in this meta-analysis, which is not significant. The fixed-effect model is therefore the most appropriate. It shows that patients on golimumab are 1.96 times more likely to achieve an ACR20 response at 6 months than patients on placebo.

Table 30. Golimumab studies included within meta-analysis (ACR20 at 24 wks in DMARD experienced population)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
GO-FORWARD	GOL	placebo	53	89	37	133
Kay	GOL	placebo	26	35	16	35

Table 31. Golimumab meta-analysis results (ACR20 at 24 wks in DMARD experienced population)

Study	RR	95% CI	Weights fixed-effect meta-analysis (%)	Weights random-effect meta-analysis (%)
GO-FORWARD	2.14	1.55, 2.96	65.0	60.8
Kay	1.63	1.08, 2.45	35.0	39.2
Pooled RR			**1.96 (1.52, 2.53)**	**1.92 (1.47, 2.51)**
p-value pooled RR			**<0.001**	**<0.001**
Heterogeneity			**I^2=8.3%, chi-square p-value=0.297**	

Infliximab

Table 32 presents the infliximab studies which were included within the meta-analyses. Table 33 presents the global analysis of the infliximab meta-analyses. There is a high amount of heterogeneity in this meta-analysis, which is significant. The random-effect model is therefore the most appropriate. It shows that patients on infliximab are 2 times more likely to achieve an ACR20 response at 6 months than patients on placebo.

Table 34 presents the meta-analyses results for infliximab excluding the monotherapy arms. There is a high amount of heterogeneity in this meta-analysis, which is significant. The random-effect model is therefore the most appropriate. It shows that patients on infliximab are 2 times more likely to achieve an ACR20 response at 6 months than patients on placebo.

Table 32. Infliximab studies included within meta-analysis (ACR20 at 24 wks in DMARD experienced population)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
ATTEST	IFX	placebo	98	165	46	110
ATTRACT	IFX	placebo	90	172	18	88
Maini	IFX	placebo	10	29	1	14
Maini (excluding monotherapy)	IFX	placebo	6	15	1	14
START	IFX	placebo	199	360	87	361

Table 33. Infliximab meta-analysis results (ACR20 at 24 wks in DMARD experienced population)

Study	**RR**	**95% CI**	**Weights fixed-effect meta-analysis (%)**	**Weights random-effect meta-analysis (%)**
ATTEST	1.42	1.1, 1.83	33.0	34.5
ATTRACT	2.56	1.65, 3.96	14.2	25.7
Maini	4.83	0.68, 34.07	0.8	3.1
START	2.29	1.87, 2.82	51.9	36.7
Pooled RR			**2.06 (1.77, 2.40)**	**2.05 (1.43, 2.92)**
p-value pooled RR			**<0.001**	**<0.001**
Heterogeneity			**I^2=72.7%, chi-square p-value=0.012**	

Table 34. Infliximab meta-analysis excluding the monotherapy arms (ACR20 at 24 wks in DMARD experienced population)

Study	**RR**	**95% CI**	**Weights fixed-effect meta-analysis (%)**	**Weights random-effect meta-analysis (%)**
ATTEST	1.42	1.1, 1.83	33.1	34.5
ATTRACT	2.56	1.65, 3.96	14.3	25.9
Maini	5.60	0.77, 40.88	0.6	3.1
START	2.29	1.87, 2.82	52.0	36.6
Pooled RR			**2.06 (1.77, 2.40)**	**2.06 (1.43, 2.95)**
p-value pooled RR			**<0.001**	**<0.001**
Heterogeneity			**I^2=73.3%, chi-square p-value=0.011**	

ACR50 at 24 weeks meta-analyses (DMARD experienced population)

Adalimumab

Table 35 presents the adalimumab studies which were included within the meta-analyses. Table 36 presents the global analysis of the adalimumab meta-analyses. There is a low amount of heterogeneity in this meta-analysis, which is not significant. The fixed-effect model is therefore the most appropriate. It shows that patients on adalimumab are 3.35 times more likely to achieve an ACR50 response at 6 months than patients on placebo.

Table 37 presents the meta-analyses results for adalimumab excluding the monotherapy arms. There is a moderate to high level of heterogeneity in this meta-analysis, but it does not reach significance. The fixed-effect model is therefore the preferred one, although it is worth looking at the random-effects results as well. In this case, both models concur on the magnitude of the effect and its significance, and show that patients on adalimumab are around 3.4 times more likely to achieve an ACR50 response at 6 months than patients on placebo.

Table 35. Adalimumab studies included within meta-analysis (ACR50 at 24 wks in DMARD experienced population)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
ARMADA	ADA	placebo	37	67	5	62
CHANGE (monotherapy)	ADA	placebo	22	91	5	87
DE019	ADA	placebo	81	207	19	200
Kim	ADA	placebo	28	65	9	63
STAR	ADA	placebo	92	318	36	318
Van de Putte (monotherapy)	ADA	placebo	25	113	9	110

Table 36. Adalimumab meta-analysis results (ACR50 at 24 wks in DMARD experienced population)

Study	**RR**	**95% CI**	**Weights fixed-effect meta-analysis (%)**	**Weights random-effect meta-analysis (%)**
ARMADA	6.85	2.88, 16.31	6.2	8.6
CHANGE	4.21	1.67, 10.61	6.1	7.7
DE019	4.12	2.6, 6.53	23.0	24.1
Kim	3.02	1.55, 5.87	10.9	13.6
STAR	2.56	1.8, 3.64	42.9	33.8
Van de Putte	2.70	1.32, 5.53	10.9	12.1
Pooled RR			**3.35 (2.67, 4.20)**	**3.34 (2.55, 4.38)**
p-value pooled RR			**<0.001**	**<0.001**
Heterogeneity			**I^2=20.9%, chi-square p-value=0.277**	

Table 37. Adalimumab meta-analysis excluding the monotherapy arms (ACR50 at 24 wks in DMARD experienced population)

Study	**RR**	**95% CI**	**Weights fixed-**	**Weights**

			effect meta-analysis (%)	random-effect meta-analysis (%)
ARMADA	6.85	2.88, 16.31	7.5	13.9
DE019	4.12	2.6, 6.53	27.7	29.7
Kim	3.02	1.55, 5.87	13.1	19.9
STAR	2.56	1.8, 3.64	51.7	36.5
Pooled RR			**3.37 (2.64, 4.31)**	**3.49 (2.40, 5.08)**
p-value pooled RR			**<0.001**	**<0.001**
Heterogeneity			**I^2=48.0%, chi-square p-value=0.124**	

Certolizumab

Table 38 presents the certolizumab studies which were included within the meta-analyses. Table 39 presents the global analysis of the certolizumab meta-analyses. There is a moderate to high level of heterogeneity in this meta-analysis, but it does not reach significance. The fixed-effect model is therefore the preferred one, although it is worth looking at the random-effects results as well. In this case, both models concur on the magnitude of the effect and its significance, and show that patients on certolizumab are over 6 times more likely to achieve an ACR50 response at 6 months than patients on placebo.

Table 38. Certolizumab studies included within meta-analysis (ACR50 at 24 wks in DMARD experienced population)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
RAPID 1	CTZ	placebo	146	393	15	199
RAPID 2	CTZ	placebo	80	246	4	127

Table 39. Certolizumab meta-analysis results (ACR50 at 24 wks in DMARD experienced population)

Study	RR	95% CI	Weights fixed-effect meta-analysis (%)	Weights random-effect meta-analysis (%)
RAPID 1	4.93	2.98, 8.15	79.1	66.4
RAPID 2	10.33	3.87, 27.54	20.9	33.6
Pooled RR			**6.06 (3.87, 9.48)**	**6.32 (3.15, 12.66)**
p-value pooled RR			**<0.001**	**<0.001**
Heterogeneity			**I^2=43.8%, chi-square p-value=0.182**	

Etanercept

Table 40 presents the etanercept studies which were included within the meta-analyses. Table 41 presents the global analysis of the etanercept meta-analyses. There

is a very high amount of heterogeneity in this meta-analysis, which is highly significant. The random-effect model is therefore the most appropriate. It shows that patients on etanercept are 3 times more likely to achieve an ACR50 response at 6 months than patients on placebo.

Table 40. Etanercept studies included within meta-analysis (ACR50 at 24 wks in DMARD experienced population)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
Combe	ETN	placebo	100	204	7	50
Combe (excluding monotherapy)	ETN	placebo	52	101	7	50
Moreland (monotherapy)	ETN	placebo	31	78	4	80
TEMPO	ETN	placebo	225	454	91	228
TEMPO (excluding monotherapy)	ETN	placebo	136	231	91	228
Weinblatt	ETN	placebo	23	59	1	30

Table 41. Etanercept meta-analysis results (ACR50 at 24 wks in DMARD experienced population)

Study	**RR**	**95% CI**	**Weights fixed-effect meta-analysis (%)**	**Weights random-effect meta-analysis (%)**
Combe	3.68	1.80, 7.50	9.2	37.6
TEMPO	1.48	1.22, 1.79	89.5	45.1
Weinblatt	11.69	1.66, 82.47	1.3	17.3
Pooled RR			**1.81 (1.49, 2.19)**	**2.98 (1.06, 834)**
p-value pooled RR			**<0.001**	**0.038**
Heterogeneity			**I^2=82.8%, chi-square p-value=0.003**	

Golimumab

Table 42 presents the golimumab studies which were included within the meta-analyses. Table 43 presents the global analysis of the golimumab meta-analyses. There is no heterogeneity in this meta-analysis. The fixed-effect model is therefore the most appropriate. It shows that patients on golimumab are 2.9 times more likely to achieve an ACR50 response at 6 months than patients on placebo.

Table 42. Golimumab studies included within meta-analysis (ACR50 at 24 wks in DMARD experienced population)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
GO-FORWARD	GOL	placebo	33	89	18	133
Kay	GOL	placebo	14	35	4	35

Table 43. Golimumab meta-analysis results (ACR50 at 24 wks in DMARD experienced population)

Study	RR	95% CI	Weights fixed-effect meta-analysis (%)	Weights random-effect meta-analysis (%)
GO-FORWARD	2.74	1.65, 4.55	78.3	79.8
Kay	3.50	1.28, 9.59	21.7	20.2
Pooled RR			**2.90 (1.84, 4.58)**	**2.88 (1.83, 4.53)**
p-value pooled RR			**<0.001**	**<0.001**
Heterogeneity			**I^2=0%, chi-square p-value=0.669**	

Infliximab

Table 44 presents the infliximab studies which were included within the meta-analyses. Table 45 presents the global analysis of the infliximab meta-analyses. There is a moderate to high amount of heterogeneity in this meta-analysis, which barely fails to reach significance. The fixed-effect model is therefore the preferred one, although it is worth looking at the random-effects results as well. In this case, both models concur on the magnitude of the effect and its significance, and show that patients on infliximab are 3 times more likely to achieve an ACR50 response at 6 months than patients on placebo.

Table 46 presents the meta-analyses results for infliximab excluding the monotherapy arms. There is a moderate to high amount of heterogeneity in this meta-analysis, which is significant. The random-effect model is therefore the most appropriate. It shows that patients on infliximab are 3.1 times more likely to achieve an ACR50 response at 6 months than patients on placebo.

Table 44. Infliximab studies included within meta-analysis (ACR50 at 24 wks in DMARD experienced population)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
ATTEST	IFX	placebo	61	165	22	110
ATTRACT	IFX	placebo	47	172	4	88
Maini	IFX	placebo	7	29	0	14
Maini (excluding monotherapy)	IFX	placebo	5	15	0	14
START	IFX	placebo	110	360	33	361

Table 45. Infliximab meta-analysis results (ACR50 at 24 wks in DMARD experienced population)

Study	RR	95% CI	Weights fixed-effect meta-analysis (%)	Weights random-effect meta-analysis (%)
ATTEST	1.85	1.21, 2.82	40.4	37.8
ATTRACT	6.01	2.24, 16.15	8.1	18.4
Maini	7.50	0.46, 122.7	1.0	3.4
START	3.34	2.33, 4.79	50.5	40.4

Pooled RR			**3.00 (2.30, 3.90)**	**3.06 (1.79, 5.23)**
p-value pooled RR			**<0.001**	**<0.001**
Heterogeneity			**I^2=60.9%, chi-square p-value=0.053**	

Table 46. Infliximab meta-analysis excluding the monotherapy arms (ACR50 at 24 wks in DMARD experienced population)

Study	**RR**	**95% CI**	**Weights fixed-effect meta-analysis (%)**	**Weights random-effect meta-analysis (%)**
ATTEST	1.85	1.21, 2.82	40.5	37.6
ATTRACT	6.01	2.24, 16.15	8.1	18.9
Maini	10.31	0.62, 170.96	0.8	3.5
START	3.34	2.33, 4.79	50.6	40.0
Pooled RR			**3.01 (2.31, 3.92)**	**3.11 (1.80, 5.39)**
p-value pooled RR			**<0.001**	**<0.001**
Heterogeneity			**I^2=62.7%, chi-square p-value=0.045**	

ACR20 at 24 weeks (TNFα experienced population)

Trials for abatacept and tocilizumab were excluded from the meta-analyses as NICE Technology Appraisal Guidance is still pending and thus the final positioning and availability of these products witin the treatment sequence is unknown as of the submission date.

Data are only available in one study each for golimumab and rituximab. Therefore no meta-analyses are needed. Table 47 presents the assessed TNFα experienced studies. Table 48 presents the RR for each treatment. This data shows that patients on both treatments are significantly more likely to achieve an ACR20 response at 6 months than patients on placebo.

Table 47. TNFα experienced studies assessed for ACR20 at 24 wks

			treatment		**comparator**	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
GO-AFTER	GOL	placebo	52	153	26	155
REFLEX	RTX	placebo	152	298	36	201

Table 48. RR for each treatment (ACR20 at 24 wks for TNFα experienced)

Treatment	**RR**	**95% CI**	**p-value**
golimumab	2.03	1.34, 3.07	0.001
rituximab	2.85	2.08, 3.91	<0.001

ACR50 at 24 weeks (TNFα experienced population)

Abatacept and tocilizumab RCTs were excluded as both are still awaiting NICE Technology Appraisal Guidance. Data are only available in one study for both rituximab and golimumab. Therefore no meta-analyses are needed. Table 49 presents the assessed TNFα experienced studies. Table 50 presents the RR for both treatments. This shows that patients on all treatments are significantly more likely to achieve an ACR50 response at 6 months than patients on placebo.

Table 49. TNFα experienced studies assessed for ACR50 at 24 wks

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
GO-AFTER	golimumab	placebo	28	153	8	155
REFLEX	rituximab	placebo	80	298	10	201

Table 50. RR for each treatment (ACR20 at 24 wks for TNFα experienced)

Treatment	**RR**	**95% CI**	**p-value**
golimumab	3.55	1.67, 7.53	0.001
rituximab	5.40	2.87, 10.16	<0.001

5.5.5 If a meta-analysis is not considered appropriate, a rationale should be given and a qualitative overview provided. The overview should summarise the overall results of the individual studies with reference to their critical appraisal.

Not applicable for DMARD experienced population. For the TNFα inhibitor experienced population, only one RCT was identified for rituximab and one RCT for golimumab and thus a meta-analysis is not appropriate.

5.5.6 If any of the relevant RCTs listed in response to section 5.2.4 (Complete list of relevant RCTs) are excluded from the meta-analysis, the reasons for doing so should be explained. The impact that each exclusion has on the overall meta-analysis should be explored.

Table 51 presents the studies which were excluded from the meta-analyses. All tocilizumab and abatacept trials were excluded as they are currently awaiting NICE Tehnology Appraisal Guidance. DMARD experienced rituximab trials were excluded from the meta-analyses as they have yet to be appraised by NICE in this patient population.

Table 51. Studies excluded from meta-analyses

(Kremer *et al.*, 2003)	(Kremer *et al.*, 2003), phase 2b	abatacept	cDMARD-experienced
(Kremer *et al.*, 2005)	(Kremer *et al.*, 2003), phase 2b	abatacept	cDMARD-experienced
(Kremer *et al.*, 2006)	AIM	abatacept	cDMARD-experienced
(Russell *et al.*, 2007)	AIM	abatacept	cDMARD-experienced
(Weinblatt *et al.*, 2006)	ASSURE	abatacept	cDMARD- and anti-TNF-experienced
(Genovese *et al.*, 2005)	ATTAIN	abatacept	anti-TNF-experienced
(Edwards *et al.*, 2004)	(Edwards *et al.*, 2004), phase 2	rituximab	cDMARD-experienced

(Strand *et al.*, 2006)	(Edwards *et al.*, 2004), phase 2	rituximab	cDMARD-experienced
(Smolen *et al.*, 2008)	OPTION	tocilizumab	cDMARD-experienced
(Emery *et al.*, 2008)	RADIATE	tocilizumab	anti-TNF-experienced
(Nishimoto *et al.*, 2009)	SATORI	tocilizumab	cDMARD-experienced
(Genovese *et al.*, 2008)	TOWARD	tocilizumab	cDMARD-experienced

Table 52 presents the pooled baseline characteristics of the included versus excluded trials within the DMARD experienced meta-analyses. The mean characteristics are comparable across the majority of parameters and therefore have minimal implications on exclusion from the analyses. The largest difference is seen between study duration, however derivation of the confidence intervals around each of these mean values would overlap largely and thus not be significant.

Table 52. Comparison of baseline characteristics (DMARD experienced population)

	Golimumab, Adalimumab, Infliximab, Etanercept, Certolizumab studies		Abatacept, Tocilizumab, Rituximab studies	
	Number of studies	**Mean (SD)**	**Number of studies**	**Mean (SD)**
study duration (weeks)	20	33.2 (15.0)	7	40.6 (13.4)
age at baseline	20	52.7 (2.2)	7	52.5 (1.6)
percentage of female	20	80.0 (3.9)	7	79.5 (4.5)
disease duration	20	8.8 (2.1)	7	9.1 (0.9)
number of previous DMARDS	11	2.3 (0.8)	4	2.3 (0.9)
HAQ-DI at baseline	18	1.6 (0.2)	6	1.5 (0.3)

A comparison of baseline characteristics within Table 53 between included and excluded RCTs within the TNFα inhibitor experienced population found negligible differences and thus it may be concluded that the excluded studies have minimal implications on the final indirect comparison results.

Table 53. Comparison of baseline characteristics (TNFα inhibitor experienced population)

	Golimumab, Rituximab studies		Abatacept, Tocilizumab studies	
	Number of studies	**Mean (SD)**	**Number of studies**	**Mean (SD)**
study duration (weeks)	2	38.0 (19.8)	3	33.3 (16.2)
age at baseline	2	53.5 (1.3)	3	53.6(0.4)
percentage of female	2	80.5 (1.3)	3	78.3 (3.1)

disease duration	2	12.2 (0.3)	3	11.7 (0.4)
number of previous DMARDS	1	2.5 (-)	1	2 (-)
number of previous anti-TNFs	1	1.5(-)	0	-
HAQ-DI at baseline	2	1.8 (0.2)	3	1.7 (0.2)

5.6 Mixed treatment comparisons (MTC) / Indirect Comparison (IC)

5.6.1 Describe the strategies used to retrieve relevant clinical data on the comparators and common references both from the published literature and from unpublished data. The methods used should be justified with reference to the decision problem. Sufficient detail should be provided to enable the methods to be reproduced, and the rationale for any inclusion and exclusion criteria used should be provided. Exact details of the search strategy used should be provided in section 9.4, appendix 4.

The search strategy used was identical to that used in the clinical section. For details, please refer to section 5.1 and 5.2.

5.6.2 Please follow the instructions specified in sections 5.1 to 5.5 for the identification, selection and methodology of the trials, quality assessment and the presentation of results. Provide in section 9.5, appendix 5, a complete quality assessment for each comparator RCT identified.

Please see the quality assessments in Section 8.5, Table 172.

5.6.3 Provide a summary of the trials used to conduct the MTC.

Table 54 presents the studies used to conduct the mixed treatment comparison in a DMARD-experienced population. Table 55 presents the studies used to conduct the indirect comparison in a TNFα-experienced population, respectively.

Table 54. Summary of the trials used to conduct the MTC (DMARD experienced)

No. trials	References of trials	Placebo	GOL	ADA	CTZ	ETN	IFX
1	(Kim *et al.*, 2007)	✓		✓			
2	(van de Putte *et al.*, 2004)	✓		✓			
3	ARMADA	✓		✓			
4	CHANGE	✓		✓			

5	DE019	✓		✓			
6	STAR	✓		✓			
7	(Chen *et al.*, 2009)	✓		✓			
8	RAPID 1	✓			✓		
9	RAPID 2	✓			✓		
10	TEMPO	✓				✓	
11	(Combe *et al.*, 2006)	✓				✓	
12	(Moreland *et al.*, 1999)	✓				✓	
13	(Weinblatt *et al.*, 1999)	✓				✓	
14	GO-FORWARD	✓	✓				
15	(Kay *et al.*, 2008)	✓	✓				
16	ATTEST	✓					✓
27	ATTRACT	✓					✓
28	START	✓					✓
29	(Abe *et al.*, 2006)	✓					✓
20	(Maini *et al.*, 1998)	✓					✓

Table 55. Summary of the trials used to conduct the indirect comparison (TNFα experienced population)

No. trials	**References of trials**	**Placebo**	**GOL**	**RTX**
1	GO-AFTER	✓	✓	
2	REFLEX	✓		✓

5.6.4 For the selected trials, provide a summary of the data used in the analysis.

Please refer to data presented within Table 16 to Table 21.

5.6.5 Please provide a clear description of the mixed treatment comparison / indirect comparison methodology.

Please refer to Section 0 for sample source code.

DMARD experienced

The objective of this analysis is to assess the efficacy and safety of golimumab versus placebo and 4 comparators. A network analysis is a type of analysis that allows for

all the evidence on a network of treatments to be analysed at once. The situation here is perfectly suited to this with the network of treatments being golimumab, adalimumab, certolizumab, etanercept, infliximab and placebo.

Approach

The approach chosen for this analysis is a Bayesian model. The model can be written as follows:

$$n_{ib(i)} \sim \text{Bin}(p_{ib(i)}, N_{ib(i)})$$
$$n_{it(i)} \sim \text{Bin}(p_{it(i)}, N_{it(i)})$$
$$\text{logit}(p_{ib(i)}) = \lambda_i$$
$$\text{logit}(p_{it(i)}) = \lambda_i + \theta_{i,b(i),t(i)}$$

Where: i represents the study.

b(i) the baseline treatment in that study,
t(i) the comparator in that study.
$n_{ib(i)}$ (resp. $n_{it(i)}$) the number of events observed in the baseline (resp. trt) arm of study i
$N_{ib(i)}$ (resp. $N_{it(i)}$) the total number of patients in that arm.
λ_i the baseline effect (log-odds) of treatment b(i) in study i
$\theta_{i,b(i),t(i)}$ the log odds-ratio of treatment t(i) relative to treatment b(i) in study i.

This structure is common to both fixed and random-effect models. The distinction comes in the definition of theta, with:

$\theta_{i,b(i),t(i)} = \mu_{t(i)} - \mu_{b(i)}$ (fixed-effect model)

$\theta_{i,b(i),t(i)} \sim N(\mu_{t(i)} - \mu_{b(i)}, \sigma^2)$ (random-effect model)

Homogeneity of the variances is assumed in the random-effect approach (the parameter σ^2 is the same for all treatment comparisons). The variance parameter is given a vague prior, with the precision (1/variance) following a gamma distribution. Vague non-informative normal priors are given to all other parameters.

The ATTEST study reports data on 3 treatment groups: placebo, infliximab and abatacept. Whenever this trial is included in the network analysis, an adjustment should be made to take this correlation into account. In order to do so, the effects of infliximab and abatacept compared to placebo are modelled using a multivariate normal distribution as shown below:

$$\begin{pmatrix} \theta_{j,e,k} \\ \theta_{j,e,ke} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_k - \mu_e \\ \mu_{ke} - \mu_e \end{pmatrix}, \sigma^2 \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix} \right)$$

All other specifications remain unchanged.

Implementation

These models are implemented in WinBugs, version 1.4.3. For each analysis, an initial 30,000 iterations is run. If convergence has not been achieved after this, the burn-in period is extended until convergence is achieved. Auto-correlations between results from successive simulations are checked. In case of the presence of such correlations, a thinning process is applied. Results are based on a further 50,000 iterations.

Both fixed-effect and random-effect models are run for each outcome and the most appropriate model in each case is selected based on the Deviance Information Criterion (DIC). The DIC measures the performance of the model in terms of goodness of fit while penalising complexity, a small DIC indicating a better performance.

Reporting and interpretation of results

For each outcome, results are reported as:

- mean, median RR of each treatment versus placebo and associated 95% credible interval.
- mean, median RR of golimumab versus each treatment and associated 95% credible interval.
- probability that each treatment is the most effective

A 95% credible interval can be interpreted as the range of values within which the parameter has a 95% probability of falling. For example, in this analysis, if the 95% credible interval does not include 1, it can be interpreted as there being less than 5% chance that there is no difference between those two treatments. In other words, if the credible interval does not include 1 then the two treatments can be considered as significantly different. Conversely, if the credible interval includes 1 then it cannot be concluded that there is a higher risk or chance of experiencing the outcome with one treatment or the other.

Bayesian inference also offers the possibility of ranking treatments, and the probability of each treatment being the best for each outcome is reported. This is calculated as the proportion of simulations in which this treatment is ranked 'best' in terms of relative efficacy/safety.

TNFα inhibitor experienced

Methods

The efficacy of golimumab and rituximab is indirectly compared using Placebo as common comparator, following the method developed by Bucher at al.

The summary measures (RR) and their precision are calculated for each study and each outcome. Given only one study is available for each treatment of interest, no meta-analysis is required.

The indirect effect of golimumab versus rituximab and its associated 95% bilateral

confidence interval is then calculated using the formulas below. Due to the mathematical characteristics of relative risks, it is necessary to perform the analysis on the logarithmic scale and then back-transform (exponentiate) the results at the end.

$$\ln(RR)_{G \text{ vs } R} = \ln(RR)_{G \text{ vs } P} - \ln(RR)_{R \text{ vs } P}$$
$$SE(\ln(RR)_{G \text{ vs } R}) = [Var(\ln(RR)_{G \text{ vs } P}) + Var(\ln(RR)_{R \text{ vs } P})]^{1/2}$$

The 95% CI around the logarithm of the indirect effect is calculated as:

$$\ln(RR)_{G \text{ vs } R} \pm 1.96 * SE(\ln(RR)_{G \text{ vs } R})$$

5.6.6 Please present the results of the analysis.

Within a DMARD experienced population, all TNFα inhibitors were found to be similar for ACR20 and ACR50 response. No statistically significant differences were found in either efficacy outcome. Golimumab was found to be statistically superior to placebo for both ACR20 and ACR50 responders.

Table 56. ACR20 at 24 weeks MTC (DMARD experienced population)

	FIXED EFFECT MODEL (DIC=389.9)		**RANDOM EFFECT MODEL (DIC=340.1)**	
	median	**95% credibility interval**	**median**	**95% credibility interval**
golimumab	1.00	-	1.00	-
adalimumab	1.06	0.83, 1.29	0.98	0.55, 1.46
certolizumab	0.74	0.58, 0.90	0.72	0.41, 1.06
etanercept	1.17	0.91, 1.47	0.93	0.51, 1.43
infliximab	1.05	0.82, 1.29	1.05	0.57, 1.65
placebo	2.11	1.67, 2.53	2.17	1.27, 3.00

Table 57. ACR50 at 24 weeks MTC (DMARD experienced population)

	FIXED EFFECT MODEL (DIC=344.1)		**RANDOM EFFECT MODEL (DIC=320.9)**	
	median	**95% credibility interval**	**median**	**95% credibility interval**
golimumab	1.00	-	1.00	-
adalimumab	0.92	0.60, 1.36	0.90	0.40, 1.76
certolizumab	0.65	0.41, 0.99	0.63	0.27, 1.31
etanercept	1.45	0.92, 2.19	0.98	0.40, 1.99
infliximab	1.03	0.66, 1.54	0.99	0.42, 2.04
placebo	3.02	2.00, 4.35	3.22	1.54, 5.74

Within a TNFα inhibitor experienced population, golimumab was found to be statistically similar to rituximab.

Table 58. ACR20 & ACR50 at 24 weeks Indirect Comparison (TNFα inhibitor experienced population)

Outcome	Mean indirect estimate Golimumab vs Rituximab	95% confidence interval
ACR20 at 6 months	0.71	0.42, 1.20
ACR50 at 6 months	0.66	0.25, 1.76

For both patient populations, MTCs and IC found golimumab to be superior to placebo and statistically similar to other biologics.

5.6.7 Please provide the statistical assessment of heterogeneity undertaken. The degree of, and the reasons for, heterogeneity should be explored as fully as possible.

Heterogeneity is assessed by the Cochran's Q test and the I^2 statistics. Cochran's Q is the standard test for heterogeneity and examines the null hypothesis that all studies are evaluating the same effect. However, the power of this test is low when the meta-analysis includes a small number of studies, which is the case in this analysis.

The I^2 statistics is a measure that lies between 0% and 100% and gives an indication of the magnitude of heterogeneity in the meta-analysis, a value of 0% representing no observed heterogeneity and larger values indicating increasing heterogeneity. This measure can be interpreted as the percentage of the total variability in a set of effect sizes due to true heterogeneity. For example, a meta-analysis with a I^2value of 50% means that half of the total variability among effect sizes is caused not by sampling error, but by true heterogeneity between studies. The I^2 results are interpreted according to the classification proposed by Higgins and Thompson in 2002. Values of around 25% represent a low level of heterogeneity, values around 50% a medium level and values around 75% indicates a high amount of heterogeneity. Both heterogeneity measures are reported for each model and a recommendation on which model is the most appropriate is made.

5.6.8 If there is doubt about the relevance of a particular trial, please present separate sensitivity analyses in which these trials are excluded.

Exclusion of the TEMPO trial (Klareskog et al 2004)

Following a recent submission to NICE by the manufacturers of tocilizumab, the expert review group commented that the inclusion of the TEMPO study in the analysis had a substantial effect on the RR of tocilizumab versus etanercept (Meads *et al.*, 2009). In the TEMPO trial the response in the placebo arm was high compared with other studies. Therefore, in the present analysis, the TEMPO trial was removed from the data set and further fixed- and random-effects meta-analyses were performed for ACR20 and ACR50 at 24 weeks).

ACR20

Based on the random effect model, which is most appropriate due to the lower DIC score, the general conclusions remain similar. Whilst the point estimates are slightly altered from Table 59, adalimumab, etanercept and infliximab do not have a statistical difference. Certolizumab becomes statistically superior to golimumab with the exclusion of TEMPO.

Table 59. ACR20 at 24 weeks MTC: Excluding TEMPO trial (DMARD experienced population)

			FIXED EFFECT MODEL (DIC=236.8)			RANDOM EFFECT MODEL (DIC=227.5)		
			mean	median	95% credibility interval	mean	median	95% credibility interval
golimumab	vs	placebo	2.30	2.30	1.76, 2.88	2.32	2.31	1.51, 3.17
golimumab	vs	golimumab	1.00	1.00	-	1.00	1.00	-
golimumab	vs	adalimumab	1.08	1.07	0.81, 1.36	1.00	1.00	0.62, 1.42
golimumab	vs	certolizumab	0.70	0.70	0.53, 0.88	0.69	0.68	0.43, 0.97
golimumab	vs	etanercept	0.75	0.74	0.56, 0.96	0.74	0.73	0.46, 1.06
golimumab	vs	infliximab	1.05	1.05	0.79, 1.34	1.06	1.05	0.65, 1.56

ACR50

Table 60 presents the adjusted MTC figures for ACR50 at 24 weeks after the exclusion of the TEMPO trial. All of the comparators remain statistically similar to golimumab within the DMARD experienced population (random effect model due to the lower DIC score).

Table 60. ACR50 at 24 weeks MTC: Excluding TEMPO trial (DMARD experienced population)

			FIXED EFFECT MODEL (DIC=214.4)			RANDOM EFFECT MODEL (DIC=211.7)		
			mean	median	95% credibility interval	mean	median	95% credibility interval
golimumab	vs	placebo	3.28	3.22	2.06, 4.84	3.42	3.32	1.76, 5.78
golimumab	vs	golimumab	1.00	1.00	-	1.00	1.00	-
golimumab	vs	adalimumab	0.94	0.92	0.57, 1.42	0.94	0.90	0.44, 1.64
golimumab	vs	certolizumab	0.63	0.61	0.37, 0.98	0.64	0.61	0.29, 1.19
golimumab	vs	etanercept	0.63	0.60	0.35, 1.04	0.64	0.60	0.28, 1.19
golimumab	vs	infliximab	1.04	1.02	0.62, 1.61	1.05	1.01	0.47, 1.92

Exclusion of monotherapy trials

In the majority of studies reporting data for the DMARD experienced population the biologic DMARD was administered with MTX. However, in four studies no concomitant MTX was permitted (van der Putte *et al.* [2004], CHANGE, Moreland *et al.* [1999] and SATORI), and three studies included a monotherapy treatment arm

(Combe *et al.* [2006], TEMPO and Maini *et al.* [1998]). To investigate the effect of the small group of monotherapy studies, and monotherapy treatment arms, on the RR estimate additional fixed- and random-effects meta-regressions were performed.

ACR20

Table 61 shows that the exclusion of monotherapy arms from the MTC does not alter the conclusions drawn: no statistical difference for all TNFα inhibitors versus golimumab still holds true.

Table 61. ACR20 at 24 weeks MTC: Excluding monotherapy arms (DMARD experienced population)

			FIXED EFFECT MODEL (DIC=306.0)			RANDOM EFFECT MODEL (DIC=287.4)		
			mean	median	95% credibility interval	mean	median	95% credibility interval
golimumab	vs	placebo	2.06	2.06	1.65, 2.47	2.09	2.10	1.36, 2.77
golimumab	vs	golimumab	1.00	1.00	-	1.00	1.00	-
golimumab	vs	adalimumab	1.09	1.09	0.86, 1.32	1.01	1.01	0.62, 1.44
golimumab	vs	certolizumab	0.75	0.75	0.6, 0.9	0.74	0.74	0.47, 1.01
golimumab	vs	etanercept	1.13	1.12	0.86, 1.42	1.02	1.01	0.61, 1.5
golimumab	vs	infliximab	1.05	1.05	0.83, 1.27	1.05	1.04	0.64, 1.52

ACR50

Similar to Table 61, Table 62 presents data which suggests that the exclusion of monotherapy arms does not change the conclusion that golimumab is statistically similar to all of the other TNFα inhibitors.

Table 62. ACR50 at 24 weeks MTC: Excluding monotherapy arms (DMARD experienced population)

			FIXED EFFECT MODEL (DIC=279.7)			RANDOM EFFECT MODEL (DIC=272.7)		
			mean	median	95% credibility interval	mean	median	95% credibility interval
golimumab	vs	placebo	3.03	2.99	2.01, 4.27	3.20	3.11	1.66, 5.24
golimumab	vs	golimumab	1.00	1.00	-	1.00	1.00	-
golimumab	vs	adalimumab	0.92	0.91	0.59, 1.34	0.90	0.87	0.42, 1.57
golimumab	vs	certolizumab	0.66	0.65	0.42, 0.99	0.67	0.64	0.31, 1.20
golimumab	vs	etanercept	1.29	1.26	0.80, 1.93	1.10	1.06	0.47, 2.01
golimumab	vs	infliximab	1.05	1.03	0.67, 1.52	1.03	0.99	0.47, 1.83

5.6.9 Please discuss any heterogeneity between results of pairwise comparisons and inconsistencies between the direct and indirect evidence on the technologies.

Not applicable.

5.7 Non-RCT evidence

5.7.1 If non-RCT evidence is considered (see section 5.2.7), please repeat the instructions specified in sections 5.1 to 5.5 for the identification, selection and methodology of the trials, and the presentation of results. For the quality assessments of non-RCTs, use an appropriate and validated quality assessment instrument. Key aspects of quality to be considered can be found in 'Systematic reviews: CRD's guidance for undertaking reviews in health care' (www.york.ac.uk/inst/crd). Exact details of the search strategy used and a complete quality assessment for each trial should be provided in sections 9.6 and 9.7, appendices 6 and 7.

Not applicable.

5.8 *Adverse events*

5.8.1 If any of the main trials are designed primarily to assess safety outcomes (for example, they are powered to detect significant differences between treatments with respect to the incidence of an adverse event), please repeat the instructions specified in sections 5.1 to 5.5 for the identification, selection, methodology and quality of the trials, and the presentation of results. Examples for search strategies for specific adverse effects and/or generic adverse-effect terms and key aspects of quality criteria for adverse-effects data can found in 'Systematic reviews: CRD's guidance for undertaking reviews in health care' (www.york.ac.uk/inst/crd).

There were no trials identified which were designed to primaritly assess the safety outcomes of the interventions discussed herein.

5.8.2 Please provide details of all important adverse events for each intervention group. For each group, give the number with the adverse event, the number in the group and the percentage with the event. Then present the relative risk and risk difference and associated 95% confidence intervals for each adverse event.

Meta-analyses and MTCs were conducted for key safety features within the DMARD experienced and TNFα inhibitor experienced populations. Results are presented below.

DMARD experienced patient population

Serious Adverse Events (DMARD experienced)

Table 63. Adalimumab SAE data (DMARD experienced)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
CHANGE	adalimumab	placebo	17	91	8	87
Chen	adalimumab	placebo	5	35	1	12
Kim	adalimumab	placebo	7	65	6	63
STAR	adalimumab	placebo	17	318	22	318
Van de Putte	adalimumab	placebo	13	113	16	110

Table 64. Adalumumab SAE meta-analyses (DMARD experienced)

Study	**RR**	**95% CI**	**Weights fixed-effect meta-analysis (%)**	**Weights random-effect meta-analysis (%)**
CHANGE	2.03	0.92, 4.46	15.2	21.9
Chen	1.71	0.22, 13.24	2.8	3.7
Kim	1.13	0.4, 3.18	11.3	13.5
STAR	0.77	0.42, 1.43	40.8	33.1
Van de Putte	0.79	0.4, 1.57	30.0	27.8
Pooled RR			**1.04 (0.73, 1.48)**	**1.04 (0.70, 1.55)**

p-value pooled RR			**0.848**	**0.840**
Heterogeneity			**I^2=12.1%, chi-square p-value=0.337**	

There is a very low level of heterogeneity in the adalimumab meta-analysis, which does not come out as significant. The fixed-effect model is therefore the most appropriate. It shows that patients on adalimumab are 1.04 times more likely to experience a serious AE than patients on placebo and that this difference is not significant.

Table 65. Certolizumab SAE data (DMARD experienced)

			treatment		**comparator**	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
RAPID 1	certolizumab	placebo	45	392	11	199
RAPID 2	certolizumab	placebo	18	248	4	125

Table 66. Adalumumab SAE meta-analyses (DMARD experienced)

Study	**RR**	**95% CI**	**Weights fixed-effect meta-analysis (%)**	**Weights random-effect meta-analysis (%)**
RAPID 1	2.08	1.1, 3.93	73.3	73.6
RAPID 2	2.27	0.78, 6.56	26.7	26.4
Pooled RR			**2.13 (1.23, 3.67)**	**2.13 (1.23, 3.67)**
p-value pooled RR			**0.007**	**0.007**
Heterogeneity			**I^2=0%, chi-square p-value=0.889**	

There is no heterogeneity in the certolizumab meta-analysis, which shows that patients on certolizumab are 2.13 times more likely to experience a serious AE than patients on placebo and that this difference is significant.

Table 67. Etanercept SAE data (DMARD experienced)

			treatment		**comparator**	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
Combe	etanercept	placebo	8	204	1	50
TEMPO	etanercept	placebo	44	454	27	228

Table 68. Etanercept SAE meta-analyses (DMARD experienced)

Study	**RR**	**95% CI**	**Weights fixed-effect meta-analysis (%)**	**Weights random-effect meta-analysis (%)**
Combe	1.96	0.25, 15.32	4.3	4.6
TEMPO	0.82	0.52, 1.29	95.7	95.4
Pooled RR			**0.87 (0.56, 1.35)**	**0.85 (0.55, 1.32)**
p-value pooled RR			**0.526**	**0.477**
Heterogeneity			**I^2=0%, chi-square p-value=0.414**	

There is no heterogeneity in the etanercept meta-analysis. The fixed-effect model is therefore the most appropriate. It shows that patients on etanercept are 0.87 times more likely to experience a serious AE than patients on placebo and that this difference is not significant. This is equivalent to saying that patients on placebo are 1.15 times more likely to experience a serious AE than patients on etanercept.

Table 69. Golimumab SAE data (DMARD experienced)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
GO-FORWARD	golimumab	placebo	9	212	5	134
Kay	golimumab	placebo	4	37	2	34

Table 70. Golimumab SAE meta-analyses (DMARD experienced)

Study	**RR**	**95% CI**	**Weights fixed-effect meta-analysis (%)**	**Weights random-effect meta-analysis (%)**
GO-FORWARD	1.14	0.39, 3.32	74.6	69.9
Kay	1.84	0.36, 9.4	25.4	30.1
Pooled RR			**1.32 (0.54, 3.20)**	**1.31 (0.54, 3.22)**
p-value pooled RR			**0.546**	**0.550**
Heterogeneity			**I^2=0%, chi-square p-value=0.630**	

There is no heterogeneity in the golimumab meta-analysis. The fixed-effect model is therefore the most appropriate. It shows that patients on golimumab are 1.32 times more likely to experience a serious AE than patients on placebo and that this difference is not significant.

Table 71. Infliximab SAE data (DMARD experienced)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
ATTEST	infliximab	placebo	19	165	13	110
ATTRACT	infliximab	placebo	24	174	18	86
Abe	infliximab	placebo	0	49	1	47
START	infliximab	placebo	28	360	27	361

Table 72. Infliximab SAE meta-analyses (DMARD experienced)

Study	**RR**	**95% CI**	**Weights fixed-effect meta-analysis (%)**	**Weights random-effect meta-analysis (%)**
Abe	0.32	0.01, 7.66	2.2	1.0
ATTEST	0.97	0.5, 1.89	22.9	23.9
ATTRACT	0.66	0.38, 1.15	35.3	34.3
START	1.04	0.63, 1.73	39.5	40.7
Pooled RR			**0.87 (0.63, 1.21)**	**0.86 (0.63, 1.20)**

p-value pooled RR			**0.415**	**0.380**
Heterogeneity			**I^2=0%, chi-square p-value=0.586**	

There is no heterogeneity in the infliximab meta-analysis. The fixed-effect model is therefore the most appropriate. It shows that patients on infliximab are 0.87 times more likely to experience a serious AE than patients on placebo and that this difference is not significant. This is equivalent to saying that patients on placebo are 1.15 times more likely to experience a serious AE than patients on infliximab.

The MTC for serious adverse events found no statistically significant differences between golimumab , placebo and all TNFα inhibitors.

Table 73. SAE MTC (DMARD experienced)

			FIXED EFFECT MODEL (DIC=263.2)			RANDOM EFFECT MODEL (DIC=263.1)		
			mean	**median**	**95% credibility interval**	**mean**	**median**	**95% credibility interval**
golimumab	vs	placebo	1.46	1.32	0.55, 3.12	1.49	1.33	0.51, 3.39
golimumab	vs	golimumab	1.00	1.00	-	1.00	1.00	-
golimumab	vs	adalimumab	1.44	1.28	0.5, 3.23	1.43	1.25	0.44, 3.48
golimumab	vs	certolizumab	0.72	0.63	0.23, 1.7	0.74	0.63	0.2, 1.92
golimumab	vs	etanercept	1.73	1.53	0.57, 4.06	1.73	1.46	0.46, 4.52
golimumab	vs	infliximab	1.52	1.36	0.54, 3.44	1.61	1.39	0.49, 3.96

Serious Infections (DMARD experienced)

Table 74. Adalimumab SI data (DMARD experienced)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
CHANGE	adalimumab	placebo	6	91	1	87
DE019	adalimumab	placebo	11	207	1	200
STAR	adalimumab	placebo	4	318	6	318

Table 75. Adalimumab SI meta-analyses (DMARD experienced)

Study	**RR**	**95% CI**	**Weights fixed-effect meta-analysis (%)**	**Weights random-effect meta-analysis (%)**
CHANGE	5.74	0.7, 46.68	12.7	30.1
DE019	10.63	1.38, 81.56	12.7	30.7
STAR	0.67	0.19, 2.34	74.6	39.3
Pooled RR			**2.57 (1.14, 5.80)**	**2.98 (0.44, 19.92)**
p-value pooled RR			**0.023**	**0.261**

Heterogeneity			**I^2=70.9%, chi-square p-value=0.032**

There is a very high level of heterogeneity in the adalimumab meta-analysis, which comes out as significant. The random-effect model is therefore the most appropriate one. It shows that patients on adalimumab are 2.98 times more likely to experience a serious infection than patients on placebo but that this difference is not significant.

Table 76. Certolizumab SI data (DMARD experienced)

			treatment		**comparator**	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
RAPID 2	certolizumab	placebo	8	248	0	125

No meta-analysis is needed for certolizumab.

Table 77. Certolizumab serious infections RR (DMARD experienced)

Study	**RR**	**95% CI**	**p-value**
RAPID 2	8.60	0.50, 147.84	0.138

This shows that patients on certolizumab are 8.6 times more likely to experience a serious infection than patients on placebo but that this difference is not significant. This estimate is very imprecise due to the fact that no events were observed in the placebo arm.

Table 78. Etanercept SI data (DMARD experienced)

			treatment		**comparator**	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
TEMPO	etanercept	placebo	20	454	10	228

No meta-analysis is needed for certolizumab.

Table 79. Etanercept serious infections RR (DMARD experienced)

Study	**RR**	**95% CI**	**p-value**
TEMPO	1.00	0.48, 2.11	0.991

This shows that there is very little difference between etanercept and placebo in terms of serious infections.

Table 80. Golimumab SI data (DMARD experienced)

			treatment		**comparator**	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
GO-FORWARD	golimumab	placebo	2	212	1	134
Kay	golimumab	placebo	1	37	1	34

Table 81. Golimumab SI meta-analyses (DMARD experienced)

Study	RR	95% CI	Weights fixed-effect meta-analysis (%)	Weights random-effect meta-analysis (%)
GO-FORWARD	1.26	0.12, 13.81	54.0	56.6
Kay	0.92	0.06, 14.12	46.0	43.4
Pooled RR			**1.11 (0.18, 6.65)**	**1.10 (0.18, 6.65)**
p-value pooled RR			**0.913**	**0.917**
Heterogeneity			**I^2=0%, chi-square p-value=0.863**	

There is no heterogeneity in the golimumab meta-analysis. The fixed-effect model is therefore the most appropriate. It shows that patients on golimumab are 1.11 times more likely to experience a serious infection than patients on placebo and that this difference is not significant.

Table 82. Inflximab SI data (DMARD experienced)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
ATTEST	infliximab	placebo	7	165	3	110
ATTRACT	infliximab	placebo	8	174	7	86
START	infliximab	placebo	6	360	6	361

Table 83. Infliximab SI meta-analyses (DMARD experienced)

Study	RR	95% CI	Weights fixed-effect meta-analysis (%)	Weights random-effect meta-analysis (%)
ATTEST	1.56	0.41, 5.89	19.0	23.5
ATTRACT	0.56	0.21, 1.51	49.4	43.3
START	1.00	0.33, 3.08	31.6	33.1
Pooled RR			**0.89 (0.47, 1.68)**	**0.87 (0.45, 1.65)**
p-value pooled RR			**0.722**	**0.665**
Heterogeneity			**I^2=0%, chi-square p-value=0.462**	

There is no heterogeneity in this meta-analysis. The fixed-effect model is therefore the most appropriate. It shows that patients on infliximab are 0.89 times more likely to experience a serious infection than patients on placebo and that this difference is not significant. This is equivalent to saying that patients on placebo are 1.12 times more likely to experience a serious infection than patients on infliximab.

The MTC for serious infections in a DMARD experienced population have wide confidene intervals due to the none or few events observed which lead to highly imprecise estimates (especially for certolizumab).

Table 84. Serious Infections MTC (DMARD experienced)

			FIXED EFFECT MODEL (DIC=167.3)			RANDOM EFFECT MODEL (DIC=165.6)		
			mean	median	95% credibility interval	mean	median	95% credibility interval
golimumab	vs	placebo	1.90	1.11	0.17, 8.49	2.18	1.13	0.13, 10.46
golimumab	vs	golimumab	1.00	1.00	-	1.00	1.00	-
golimumab	vs	adalimumab	0.78	0.43	0.05, 3.7	0.92	0.40	0.03, 4.8
golimumab	vs	certolizumab	0.10	0.02	0, 0.7	0.16	0.02	0, 0.93
golimumab	vs	etanercept	2.00	1.09	0.14, 9.51	3.79	1.10	0.07, 17.77
golimumab	vs	infliximab	1.85	1.03	0.14, 8.56	2.27	0.99	0.09, 11.71
golimumab	vs	abatacept	1.58	0.89	0.12, 7.25	2.16	0.94	0.09, 11.34
golimumab	vs	rituximab	1.24	0.42	0.02, 7.16	2.36	0.42	0.01, 11.68
golimumab	vs	tocilizumab	1.18	0.65	0.08, 5.5	1.4	0.6	0.05, 7.25

Injection Site Reactions (DMARD experienced)

Table 85. Adalimumab ISR data (DMARD experienced)

			treatment		comparator	
study	treatment	comparator	n	total	n	total
ARMADA	adalimumab	placebo	8	67	2	62
CHANGE	adalimumab	placebo	28	91	2	87
Chen	adalimumab	placebo	1	35	0	12
DE019	adalimumab	placebo	54	207	48	200
STAR	adalimumab	placebo	62	318	37	318
Van de Putte	adalimumab	placebo	11	113	1	110

Table 86. Adalimumab ISR meta-analyses (DMARD experienced)

Study	RR	95% CI	Weights fixed-effect meta-analysis (%)	Weights random-effect meta-analysis (%)
ARMADA	3.70	0.82, 16.76	2.3	13.0
CHANGE	13.38	3.29, 54.5	2.2	14.2
Chen	1.08	0.05, 24.96	0.8	4.4
DE019	1.09	0.78, 1.52	53.2	30.0
STAR	1.68	1.15, 2.44	40.4	29.5
Van de Putte	10.71	1.41, 81.55	1.1	8.8
Pooled RR			**1.76 (1.40, 2.23)**	**2.53 (1.25, 5.14)**
p-value pooled RR			**<0.001**	**0.010**
Heterogeneity			**I^2=75.1%, chi-square p-value=0.001**	

There is a very high level of heterogeneity in the adalimumab meta-analysis, which comes out as significant. The random-effect model is therefore the most appropriate one. It shows that patients on adalimumab are 2.53 times more likely to experience an injection site reaction than patients on placebo and that this difference is significant.

Table 87. Certolizumab ISR data (DMARD experienced)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
RAPID1	certolizumab	placebo	9	392	0	199
RAPID 2	certolizumab	placebo	3	248	0	125

Table 88. Certolizumab ISR meta-analyses (DMARD experienced)

Study	**RR**	**95% CI**	**Weights fixed-effect meta-analysis (%)**	**Weights random-effect meta-analysis (%)**
RAPID1	9.67	0.57, 165.28	50.0	52.0
RAPID 2	3.54	0.18, 68.04	50.0	48.0
Pooled RR			**6.60 (0.87, 50.24)**	**5.97 (0.77, 46.26)**
p-value pooled RR			**0.068**	**0.087**
Heterogeneity			**I^2=0%, chi-square p-value=0.624**	

There is no heterogeneity in the certolizumab meta-analysis. The fixed-effect model is therefore the most appropriate one. It shows that patients on certolizumab are 6.60 times more likely to experience an injection site reaction than patients on placebo but that this difference is not significant. This estimation is very imprecise due to the fact that no events were observed in the placebo arms.

Table 89. Etanercept ISR data (DMARD experienced)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
Combe	etanercept	placebo	49	204	1	50
Moreland	etanercept	placebo	38	78	10	80
TEMPO	etanercept	placebo	69	454	4	228
Weinblatt	etanercept	placebo	25	59	2	30

Table 90. Etanercept ISR meta-analyses (DMARD experienced)

Study	**RR**	**95% CI**	**Weights fixed-effect meta-analysis (%)**	**Weights random-effect meta-analysis (%)**
Combe	12.01	1.7, 84.88	8.3	6.0
Moreland	3.90	2.09, 7.27	50.7	58.6

TEMPO	8.66	3.2, 23.44	27.4	23.1
Weinblatt	6.36	1.61, 25.05	13.6	12.2
Pooled RR			**6.21 (3.75, 10.26)**	**5.33 (3.30, 8.61)**
p-value pooled RR			**<0.001**	**<0.001**
Heterogeneity			**I^2=0.3%, chi-square p-value=0.390**	

There is almost no heterogeneity in the etanercept meta-analysis. The fixed-effect model is therefore the most appropriate one. It shows that patients on etanercept are 6.21 times more likely to experience an injection site reaction than patients on placebo and that this difference is significant.

Table 91. Golimumab ISR data (DMARD experienced)

			treatment		**comparator**	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
GO-FORWARD	golimumab	placebo	5	212	4	134
Kay	golimumab	placebo	5	37	4	34

Table 92. Golimumab ISR meta-analyses (DMARD experienced)

Study	**RR**	**95% CI**	**Weights fixed-effect meta-analysis (%)**	**Weights random-effect meta-analysis (%)**
GO-FORWARD	0.79	0.22, 2.89	54.0	47.3
Kay	1.15	0.34, 3.93	46.0	52.7
Pooled RR			**0.95 (0.39, 2.32)**	**0.96 (0.39, 2.35)**
p-value pooled RR			**0.919**	**0.933**
Heterogeneity			**I^2=0%, chi-square p-value=0.681**	

There is no heterogeneity in the golimumab meta-analysis. The fixed-effect model is therefore the most appropriate. It shows that patients on golimumab are 0.95 times more likely to experience an injection site reaction than patients on placebo and that this difference is not significant. It is equivalent to saying that patients on placebo are 1.05 times more likely to experience one than patients on golimumab.

The MTC for injection site reactions in a DMARD experienced population have wide confidence intervals due to none or few events observed which lead to highly imprecise estimates.

Table 93. Injection Site Reactions (DMARD experienced)

	FIXED EFFECT MODEL (DIC=173.4)			**RANDOM EFFECT MODEL (DIC=157.9)**		
	mean	**median**	**95% credibility interval**	**mean**	**median**	**95% credibility interval**

golimumab	vs	placebo	1.08	0.95	0.36, 2.53	1.31	0.96	0.2, 4.52
golimumab	vs	golimumab	1.00	1.00	-	1.00	1.00	-
golimumab	vs	adalimumab	0.55	0.48	0.18, 1.34	0.42	0.29	0.04, 1.57
golimumab	vs	certolizumab	0.07	0.04	0.01, 0.36	0.11	0.03	0, 0.53
golimumab	vs	etanercept	0.15	0.13	0.04, 0.38	0.17	0.11	0.02, 0.67
golimumab	vs	infliximab	-	-	-	-	-	-

Discontinuation due to AE (DMARD experienced)

Table 94. Adalimumab discontinuation due to AE data (DMARD experienced)

			treatment		**comparator**	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
ARMADA	adalimumab	placebo	0	67	2	62
CHANGE	adalimumab	placebo	12	91	4	87
Chen	adalimumab	placebo	3	35	0	12
DE019	adalimumab	placebo	26	207	13	200
Kim	adalimumab	placebo	4	65	4	63
STAR	adalimumab	placebo	9	318	8	318
Van de Putte	adalimumab	placebo	6	113	1	110

Table 95. Adalimumab discontinuation due to AE meta-analyses (DMARD experienced)

Study	**RR**	**95% CI**	**Weights fixed-effect meta-analysis (%)**	**Weights random-effect meta-analysis (%)**
AMADA	0.19	0.01, 3.79	7.7	2.0
CHANGE	2.87	0.96, 8.55	12.1	15.4
Chen	2.53	0.14, 45.69	2.2	2.2
DE019	1.93	1.02, 3.65	39.2	45.3
Kim	0.97	0.25, 3.71	12.0	10.2
STAR	1.13	0.44, 2.88	23.7	20.8
Van de Putte	5.84	0.71, 47.73	3.0	4.2
Pooled RR			**1.73 (1.15, 2.62)**	**1.72 (1.12, 2.64)**
p-value pooled RR			**0.009**	**0.013**
Heterogeneity			**I^2=0%, chi-square p-value=0.432**	

There is no heterogeneity in the adalimumab meta-analysis. The fixed-effect model is therefore the most appropriate. It shows that patients on adalimumab are 1.73 times more likely to discontinue treatment because of AE than patients on placebo and that this difference is significant.

Table 96. Certolizumab discontinuation due to AE data (DMARD experienced)

			treatment		**comparator**	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
RAPID 1	certolizumab	placebo	17	393	3	199

RAPID 2	certolizumab	placebo	11	246	2	127

Table 97. Certolizumab discontinuation due to AE meta-analyses (DMARD experienced)

Study	RR	95% CI	Weights fixed-effect meta-analysis (%)	Weights random-effect meta-analysis (%)
RAPID 1	2.87	0.85, 9.68	60.2	60.1
RAPID 2	2.84	0.64, 12.62	39.8	39.9
Pooled RR			**2.86 (1.11, 7.33)**	**2.86 (1.11, 7.33)**
p-value pooled RR			**0.029**	**0.029**
Heterogeneity			**I^2=0%, chi-square p-value=0.991**	

There is no heterogeneity in the certolizumab meta-analysis, which shows that patients on certolizumab are 2.86 times more likely to discontinue treatment because of AE than patients on placebo and that this difference is significant.

Table 98. Etanercept discontinuation due to AE data (DMARD experienced)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
Combe	etanercept	placebo	7	204	3	50
Moreland	etanercept	placebo	5	78	3	80
TEMPO	etanercept	placebo	49	454	32	228
Weinblatt	etanercept	placebo	2	59	0	30

Table 99. Etanercept discontinuation due to AE meta-analyses (DMARD experienced)

Study	RR	95% CI	Weights fixed-effect meta-analysis (%)	Weights random-effect meta-analysis (%)
Combe	0.57	0.15, 2.13	9.4	8.3
Moreland	1.71	0.42, 6.91	5.8	7.3
TEMPO	0.77	0.51, 1.17	83.5	82.8
Weinblatt	2.58	0.13, 52.16	1.3	1.6
Pooled RR			**0.83 (0.57, 1.21)**	**0.81 (0.56, 1.18)**
p-value pooled RR			**0.326**	**0.279**
Heterogeneity			**I^2=0%, chi-square p-value=0.570**	

There is no heterogeneity in the etanercept meta-analysis. The fixed-effect model is therefore the most appropriate. It shows that patients on etanercept are 0.83 times more likely to discontinue treatment because of AE than patients on placebo and that this difference is not significant. This is equivalent to saying that patients on placebo are 1.20 times more likely to discontinue treatment than patients on etanercept.

Table 100. Golimumab discontinuation due to AE data (DMARD experienced)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
GO-FORWARD	golimumab	placebo	2	89	5	133
Kay	golimumab	placebo	2	35	3	35

Table 101. Golimumab discontinuation due to AE meta-analyses (DMARD experienced)

Study	RR	95% CI	Weights fixed-effect meta-analysis (%)	Weights random-effect meta-analysis (%)
GO-FORWARD	0.60	0.12, 3.01	57.2	53.3
Kay	0.67	0.12, 3.75	42.8	46.7
Pooled RR			**0.63 (0.19, 2.04)**	**0.63 (0.19, 2.05)**
p-value pooled RR			**0.439**	**0.442**
Heterogeneity			**I^2=0%, chi-square p-value=0.928**	

There is no heterogeneity in the golimumab meta-analysis. The fixed-effect model is therefore the most appropriate. It shows that patients on golimumab are 0.63 times more likely to discontinue treatment because of AE than patients on placebo and that this difference is not significant. This is equivalent to saying that patients on placebo are 1.59 times more likely to discontinue treatment than patients on golimumab.

Table 102. Infliximab discontinuation due to AE data (DMARD experienced)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
ATTEST	infliximab	placebo	8	165	1	110
ATTRACT	infliximab	placebo	14	172	7	88
Abe	infliximab	placebo	1	49	1	47
Maini	infliximab	placebo	1	29	0	14
START	infliximab	placebo	18	360	8	361

Table 103. Infliximab discontinuation due to AE meta-analyses (DMARD experienced)

Study	RR	95% CI	Weights fixed-effect meta-analysis (%)	Weights random-effect meta-analysis (%)
ATTEST	5.33	0.68, 42.05	6.0	7.2
ATTRACT	1.02	0.43, 2.44	46.0	40.3
Abe	0.96	0.06, 14.9	5.1	4.1
Maini	1.50	0.06, 34.66	3.3	3.1
START	2.26	0.99, 5.12	39.7	45.4
Pooled RR			**1.78 (1.04, 3.06)**	**1.66 (0.96, 2.89)**
p-value pooled RR			**0.036**	**0.071**
Heterogeneity			**I^2=0%, chi-square p-value=0.530**	

There is no heterogeneity in the infliximab meta-analysis. The fixed-effect model is therefore the most appropriate. It shows that patients on infliximab are 1.78 times more likely to discontinue treatment because of AE than patients on placebo and that this difference is significant.

The MTC for discontinuation due to AE shows that golimumab is the treatment that performs the best on this outcome (ie, the treatment with the least chance of discontinuation because of an AE). The RR show that golimumab performs better on this outcome than all other treatments and the difference compared to certolizumab is significant.

Table 104. Discontinuation due to AE (DMARD experienced)

			FIXED EFFECT MODEL (DIC=273.3)			RANDOM EFFECT MODEL (DIC=274.5)		
			mean	median	95% credibility interval	mean	median	95% credibility interval
golimumab	vs	placebo	0.70	0.59	0.15, 1.92	0.71	0.59	0.14, 2.02
golimumab	vs	golimumab	1.00	1.00	-	1.00	1.00	-
golimumab	vs	adalimumab	0.40	0.33	0.08, 1.14	0.41	0.33	0.07, 1.26
golimumab	vs	certolizumab	0.27	0.20	0.04, 0.88	0.27	0.20	0.04, 0.95
golimumab	vs	etanercept	0.88	0.72	0.17, 2.54	0.86	0.68	0.14, 2.67
golimumab	vs	infliximab	0.37	0.30	0.07, 1.08	0.37	0.29	0.06, 1.17

TNFα inhibitor experienced

Table 105. SAE data (TNFα inhibitor experienced)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
GO-AFTER	golimumab	placebo	11	152	15	155
REFLEX	rituximab	placebo	23	309	21	208

Table 106. SAE RR (TNFα inhibitor experienced)

Treatment	RR	95% CI	p-value
golimumab	0.75	0.35, 1.58	0.445
rituximab	0.74	0.42, 1.30	0.290

The RR for golimumab and rituximab shows that patients on both treatments are less likely to experience a serious AE than patients on placebo, but neither of these differences are significant.

The serious adverse event MTC found no statistically significant difference between golimumab and either placebo or rituximab.

Table 107. Serious Adverse Event MTC (TNFα inhibitor experienced)

			FIXED EFFECT MODEL		
			mean	median	95% credibility interval
golimumab	vs	placebo	0.79	0.74	0.34, 1.53
golimumab	vs	golimumab	1.00	1.00	-
golimumab	vs	rituximab	1.12	1.00	0.38, 2.55

Table 108. Serious infection data (TNFα inhibitor experienced)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
GO-AFTER	golimumab	placebo	5	152	5	155
REFLEX	rituximab	placebo	7	309	3	208

Table 109. Serious infection RR (TNFα inhibitor experienced)

Treatment	RR	95% CI	p-value
golimumab	1.02	0.30, 3.45	0.975
rituximab	1.57	0.41, 6.00	0.509

This shows that there is almost no difference between golimumab and placebo. Patients on rituximab are more likely to experience a serious infection than patients on placebo, but none of these differences are significant.

The serious infection MTC found no statistically significant difference between golimumab and either placebo or rituximab.

Table 110. Serious Infection MTC (TNFα inhibitor experienced)

			FIXED EFFECT MODEL		
			mean	**median**	**95% credibility interval**
golimumab	vs	placebo	1.25	1.02	0.28, 3.62
golimumab	vs	golimumab	1.00	1.00	-
golimumab	vs	rituximab	0.94	0.61	0.09, 3.75

Table 111. Injection site reaction data (TNFα inhibitor experienced)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
GO-AFTER	golimumab	placebo	9	152	6	155

Data are only available in the golimumab study for this outcome. Results are shown below.

Table 112. Injection site reaction RR (TNFα inhibitor experienced)

Treatment	RR	95% CI	p-value
golimumab	1.53	0.56, 4.19	0.409

This shows that patients on golimumab are more likely to experience an injection site reaction than patients on placebo, but this difference is not significant.

Table 113. Discontinuation due to AE (TNFα inhibitor experienced)

			treatment		comparator	
study	**treatment**	**comparator**	**n**	**total**	**n**	**total**
GO-AFTER	golimumab	placebo	4	152	9	155
REFLEX	rituximab	placebo	8	309	2	208

A meta-analysis is not needed as data are only available in one study for the both treatments.

Table 114. Injection site reaction RR (TNFα inhibitor experienced)

Treatment	RR	95% CI	p-value
golimumab	0.45	0.14, 1.44	0.180
rituximab	2.69	0.58, 12.55	0.207
tocilizumab	1.14	0.46, 2.82	0.772

This shows that patients on golimumab are less likely to discontinue treatment because of an AE than patients on placebo. Patients on rituximab are on the other hand more likely to discontinue because of an AE than patients on placebo. However, none of these differences are significant.

The discontinuation due to AE MTC found no statistically significant difference between golimumab and placebo. Golimumab was found to have less discontinuations due to AE than rituximab. This result was significant.

Table 115. Discontinuation due to AE MTC (TNFα inhibitor experienced)

			FIXED EFFECT MODEL		
			mean	median	95% credibility interval
golimumab	vs	placebo	0.50	0.43	0.11, 1.33
golimumab	vs	golimumab	1.00	1.00	-
golimumab	vs	rituximab	0.23	0.15	0.02, 0.91

5.8.3 Give a brief overview of the safety of the technology in relation to the decision problem.

As presented in more detail within Section 5.9, golimumab in combination with MTX was generally well tolerated within a DMARD experienced and TNFα experienced patient population. The proportion of patients having at least 1 adverse events was similar across the placebo and treatment arms within GO-FORWARD and GO-AFTER. Section 5.8.2 presents meta-analyses and MTCs for both patient populations which generally conclude there is no statistically significant differences between golimumab and the available biologics and placebo.

5.9 Interpretation of clinical evidence

5.9.1 Please provide a statement of principal findings from the clinical evidence highlighting the clinical benefit and harms from the technology.

GO-FORWARD: DMARD Experienced Population

GO-FORWARD Design

GO-FORWARD (Keystone 2009) is a multi-centre, randomised, double-blind, placebo-controlled study designed to assess the efficacy and safety of golimumab plus MTX compared with MTX alone in patients with active RA despite ongoing MTX therapy. The golimumab dose regimens chosen for this study, 50 mg and 100 mg administered by subcutaneous injection every 4 weeks were selected based on non-clinical data of the Phase 2 RA golimumab dose-ranging study (Kay et al 2008).

GO-FORWARD Methods

Four hundred forty-four patients from 60 sites across twelve countries were recruited. All had been diagnosed with active RA (persistent disease activity with at least four swollen and four tender joints) according to the American College of Rheumatology (ACR) criteria (Arnett et al 1988) despite a stable MTX dose of at least 15 mg/week for at least 4 weeks prior to screening. The enrolled patients must have tolerated MTX (at least 15 mg/week) for at least 3 months and were excluded from the study if an anti-TNFα inhibitor, rituximab, natalizumab or cytotoxic agents had ever been received. Patients were also excluded from the study if anakinra, DMARDs other than MTX, or corticosteroids had been received within four weeks before the first dose of the study agent. Patients on non-steroidal anti-inflammatory drugs (NSAIDs) or other RA analgesics had to be taking a stable dose (at least 2 weeks) before the first dose of the study.

The patients were randomised 3:3:2:2 into the following four treatment groups:

Group 1: placebo injection + MTX (n=133),
Group 2: golimumab 100 mg injection + placebo capsules (n=133),
Group 3: golimumab 50 mg injection + MTX (n=89),
Group 4: golimumab 100 mg injection + MTX (n=89).

The randomisation was stratified by study site. At week 16, patients with an inadequate response in Group 1, 2 or 3 (less than 20% improvement from baseline in both tender and swollen joint counts) entered a double-blinded rescue therapy phase: patients in Group 1, 2 or 3 subsequently received active 50 mg golimumab and active MTX, 100 mg golimumab and active MTX, or increased dose to 100 mg golimumab and active MTX, respectively. Patients in Group 4 received no dose or treatment adjustments.

GO-FORWARD Primary endpoints

A number of health-related outcomes were measured in GO-FORWARD to assess the efficacy of golimumab in patients with active RA despite MTX therapy including the reduction of the signs and symptoms of RA at Week 14 and the improvement in physical function at Week 24.

The two co-primary endpoints were:

- ≥20% improvement in ACR criteria for the assessment of RA (ACR20) at Week 14
- Improvement from baseline in Health Assessment Questionnaire (HAQ) at Week 24

A wide range of secondary endpoints were also measured; major secondary endpoints included:

- ACR 20 at Week 24
- ACR 50 at Week 14 and Week 24
- DAS 28 (Disease Activity Score) at Week 14 and Week 24
- Improvement from baseline in HAQ at Week 14

ACR

A health outcome measure on the American College of Rheumatology (ACR) 20 clinical criteria defined as achieving (Kwoh et al 2002):

- At least 20% improvement in the tender joint count, and
- At least 20% improvement in the swollen joint count, and
- At least 20% improvement in 3 of the following 5 assessments:
- Patient's global assessment of disease activity (VAS).
- Physician's global assessment if disease activity (VAS).
- Evaluator's global assessment of disease activity (VAS).
- Patient's assessment of physical function as measured by the Health Assessment Questionnaire disability index.
- Level of acute-phase reactant (CRP).

The ACR50 is defined as a 50% improvement in the parameters as described above for ACR20.

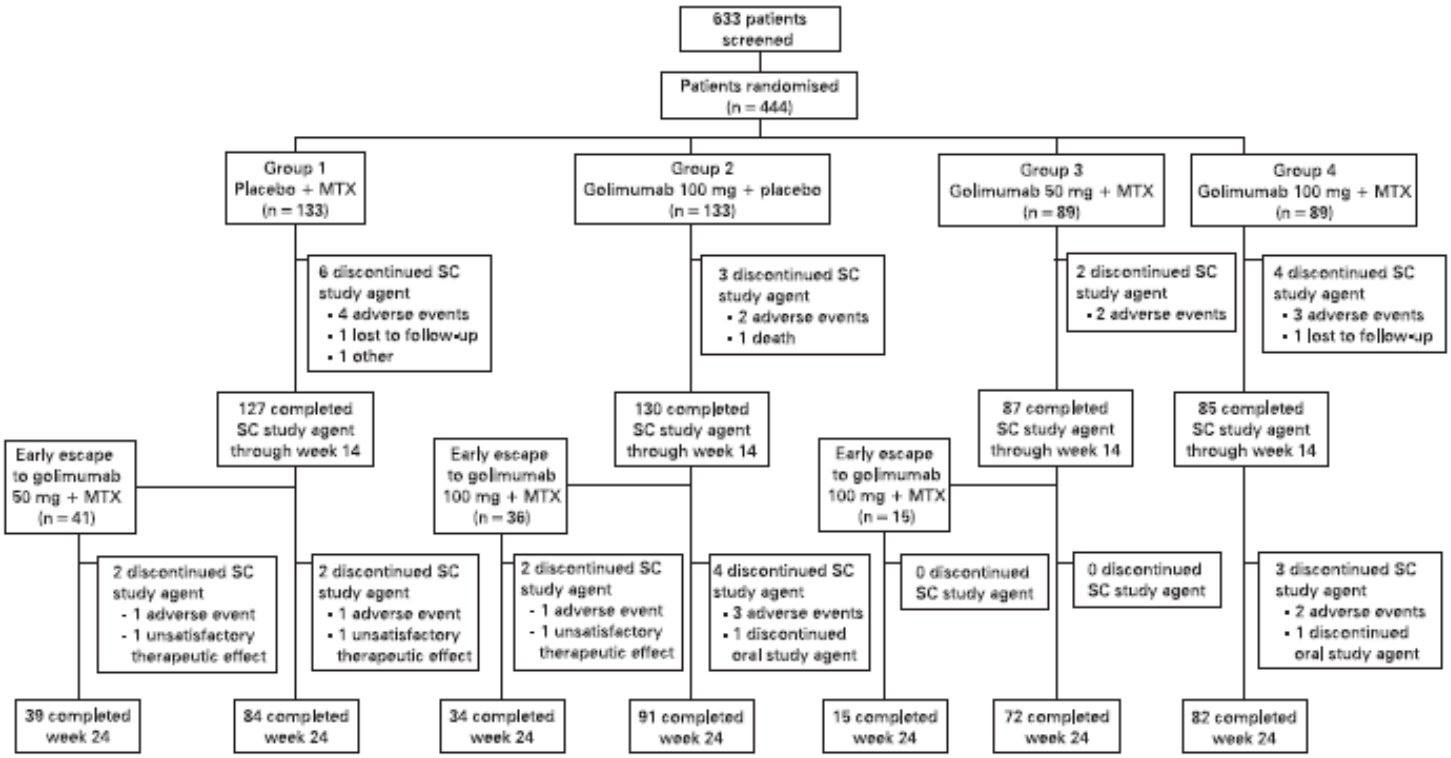
HAQ

The Health Assessment Questionnaire disability index score (HAQ-DI) is a patient reported outcome instrument based on eight dimensions: dressing and grooming, arising, eating, walking, hygiene, reach, grip and activities (Fries et al 1980).

DAS28

The Disease Activity Index Score 28 (DAS28) using C-reactive protein (CRP) is a statistically derived index combining the number of tender and swollen joints (based on a 28 joint count), CRP, and GH (van der Linden et al 2004). A comparison of measures in clinical trials of TNFα inhibitors (infliximab and etanercept) found that the DAS28 performed better than single clinical variables (ie, swollen joint count or physician's global assessment) in regards to discerning differences between the placebo and active drug groups) (Fransen et al 2004).

Figure 2. GO-FORWARD patient flow



GO-FORWARD Efficacy Results

Four hundred forty-four patients were randomised, of whom 133 were assigned to the placebo group (Group 1), and 89 were assigned to the 50 mg golimumab group (Group 3). Baseline clinical disease characteristics were generally well balanced across the groups, and generally indicated the presence of long-standing disease of substantial impact.

Note: for the purpose of this appraisal, only the placebo and 50 mg groups will be discussed. Monotherapy (Group 2) is not within golimumab's licence. Golimumab 100mg is only approved within a small sub-population: patients >100kg who do not respond to 50mg.

Table 116. Baseline demographics of placebo and golimumab 50mg groups

Demographic	Characteristic	Placebo (Group 1)	Golimumab 50mg (Group 3)
Patients randomised	n	133	89
Age (years)	Mean (SD)	[redacted]	[redacted]
Women (years)	n (%)	109 (82.0%)	72 (80.9%)
Height (cm)	Mean (SD)	[redacted]	[redacted]
Weight (kg)	Mean (SD)	[redacted]	[redacted]
RA disease duration (years)	Mean (SD)	[redacted]	[redacted]

	Median	6.50	4.50
Oral corticosteroids use	n (%)	87 (65.4%)	67 (75.3%)
NSAIDs use	n (%)	114 (85.7%)	77 (86.5%)
Methotrexate dose (mg/wk)	Mean (SD)	[redacted]	[redacted]
Duration of previous MTX use, years			
< 1	n (%)	33 (24.8%)	20 (22.5%)
≥ 1 to > 3	n (%)	30 (22.6)	32 (36.0)
≥ 3	n (%)	68 (51.1)	37 (41.6)
Corticosteroid use	n (%)	83 (53.5%)	92 (60.5%)
Prednisolone or equivalent dose (mg/wk)	Mean (SD)	[redacted]	[redacted]
Swollen joint count (0-66)	Mean (SD)	[redacted]	[redacted]
	Median	9.0	8.0
Tender joint count (0-68)	Mean (SD)	[redacted]	[redacted]
	Median	12.0	15.0
Patient assessment of pain	Mean (SD)	[redacted]	[redacted]
(VAS, 0-10cm)	Median	5.70	6.10
Patient's global assessment of disease	Mean (SD)	[redacted]	[redacted]
(VAS, 0-10cm)	Median	5.30	6.00
Physician's global assessment of disease	Mean (SD)	[redacted]	[redacted]
(VAS, 0-10cm)	Median	5.65	6.10
(HAQ-DI, 0-3)	Median	1.25	1.38
C-reactive protein (mg/dL)	Mean (SD)	[redacted]	[redacted]
	Median	0.80	1.00

GO-FORWARD Primary endpoint

Both co-primary endpoints were met. A significantly greater proportion of patients in the golimumab 50 mg + MTX group achieved an ACR20 response at week 14 than the placebo +MTX group (Table 117). Improvement in HAQ at week 24 was significantly greater in the golimumab 50 mg + MTX group compared with the placebo + MTX group (median: 0.3750 vs 0.1250; p<0.001).

Table 117. Co-primary endpoints, randomised patients

Response	Characteristic	Placebo (Group 1)	Golimumab 50mg (Group 3)	p-value
Patients Randomised		n=133	n=89	
ACR20 response at week 14				
patients in response	n (%)	44 (33.1%)	49 (55.1%)	p=0.001
Improvement from baseline in HAQ score at week 24				
	Mean (SD)	[redacted]	[redacted]	
	Median	0.1250	0.3750	p<0.001

GO-FORWARD Major secondary endpoints

All secondary endpoints in Table 118 had a significantly greater proportion of patients respond in the golimumab 50 mg + MTX group compared with placebo + MTX.

Table 118. Secondary endpoints

Response	Characteristic	Placebo (Group 1)	Golimumab 50mg (Group 3)	p-value
Patients Randomised		n=133	n=89	
ACR20 at week 24	n (%)	37 (27.8%)	53 (59.6%)	p<0.001
ACR50 at week 14	n (%)	13 (9.8%)	31 (34.8%)	p<0.001
ACR50 at week 24	n (%)	18 (13.5%)	33 (37.1%)	p<0.001
DAS28 at week 14	n (%)	69 (51.9%)	66 (74.2%)	p<0.001
DAS28 at week 24	n (%)	62 (46.6%)	66 (74.2%)	p<0.001
Improvement from baseline in HAQ at week 14	Mean (SD)	0.158 (0.494)	0.421 (0.504)	
	Median	0.125	0.375	p<0.001

GO-FORWARD Efficacy Conclusion

GO-FORWARD provides evidence that golimumab reduces the signs and symptoms of RA and improves physical function in patients with active RA despite treatment with MTX, which suggests first line treatment with golimumab following exposure to DMARDs was effective.

In patients with active RA who had previously received methotrexate:

- A significantly greater proportion of patients receiving golimumab achieved the primary endpoints (ACR 20 response at week 14, improvement from baseline HAQ at week 24) compared with placebo.
- Golimumab was superior to placebo in the following secondary endpoints:
    - ACR 20 and 50 response at week 24
    - DAS28 at week 14 and 24
    - Improvement from baseline in HAQ at week 14

GO-FORWARD Safety Background

GO-FORWARD reported the incidence and type of adverse events (AEs) by treatment group at week 16 and week 24. Comparisons of AE incidence between groups of patients meeting criteria for early escape and changing treatment regimens through week 24 are not presented due to the discrepancy in duration of follow-up in those treatment groups compared with randomised treatment groups.

Whilst week 24 adverse events data is presented in , data presented through week 16 may be more appropriate. Data presented for week 24 may underestimate the

proportion of patients with safety events in the placebo + MTX group compared with the golimumab 50 mg + MTX group as those patients in the placebo + MTX group had a lower average duration of follow-up and a lower average number of administrations than patients treated with golimumab 50 mg + MTX due to the greater number of patients meeting criteria for early escape at week 16 in the placebo + MTX group.

GO-FORWARD Safety Results

Prior to early escape at week 16, the adverse event profiles were similar for the placebo + MTX and golimumab 50mg +MTX groups. The proportions of patients with adverse events were 60.9% in group 1 and 68.5% in group 3 (Table 119). There were no substantial differences in the frequency of common adverse events, infections, serious adverse events, serious infections, injection site disorders, or malignancies between the two groups.

Table 119. Adverse events to week 16 (before early escape)

Assessment	**Placebo (Group 1)**	**Golimumab 50mg (Group 3)**
	n=133	n=89
Average duration of follow-up (weeks)	15.9	16.1
Average exposure (no of administrations)	3.9	3.9
Adverse events		
Urinary tract infections	1 (0.8%)	0 (0.0%)
Cellulitis	0 (0.0%)	1 (1.1%)
Subcutaneous abscess	0 (0.0%)	1 (1.1%)
Bursitis	1 (0.8%)	0 (0.0%)
Rheumatoid arthritis	0 (0.0%)	1 (1.1%)
Myocardial infarction	1 (0.8%)	0 (0.0%)
Goitre	7 (5%)	5 (3%)
Hypertension	2 (1%)	5 (3%)
Infections	32 (24.1%)	25 (28.1%)
Serious adverse events	3 (2.3%)	5 (5.6%)
Serious infections	1 (0.8%)	2 (2.2%)
Injection-site disorders	3 (2.3%)	4 (4.5%)
Malignancies	0 (0.0%)	0 (0.0%)
Data are number of patients (%) unless stated		

Through week 24, the proportion of patients experiencing at least 1 AE was 67.7% in Group 1 and 73.0% in group 3 (Table 120). Upper respiratory tract infections were reported at a higher frequency in group 3 than group 1 whereas all other assessments were not substantially different between the two groups. A larger proportion of patients discontinued treatment through week 24 due to one or more adverse events in group 1 (n=6, 4.5%) than group 3 (n=2, 2.2%).

Table 120. Adverse events to week 24

Assessment	Placebo (Group 1)	Golimumab 50mg (Group 3)
	n=133	n=89
Average duration of follow-up (weeks)	21.1	22.6
Average exposure (no of administrations)	5.1	5.5
Adverse events	90 (67.7%)	65 (73.0%)
Upper respiratory tract infection	9 (6.8%)	11 (12.4%)
Cough	7 (5.3%)	6 (6.7%)
Headache	5 (3.8%)	5 (5.6%)
Nasopharyngitis	6 (4.5%)	4 (4.5%)
Rash	4 (3.0%)	5 (5.6%)
Bronchitis	3 (2.3%)	3 (3.4%)
Abdominal pain upper	4 (3.0%)	3 (3.4%)
Diarrhoea	4 (3.0%)	4 (4.5%)
Infections	37 (27.8%)	28 (31.5%)
Serious adverse events	5 (3.8%)	6 (6.7%)
Serious infections	1 (0.8%)	2 (2.2%)
Injection-site reactions	0 (0.0%)	3 (3.4%)
Data are number of patients (%) unless stated		

GO-FORWARD Safety Summary

Golimumab in combination with MTX was generally well tolerated. The proportion of patients experiencing at least 1 AE was similar across group 1 and 3 at week 16 and week 24. Serious adverse events and serious infections through week 16 and through week 24 were slightly more common in group 3 than group 1. Common adverse events (i.e., cough, headache) were not substantially different between the two groups.

GO-FORWARD Long Term Data: Open Label Trial

Long term data from GO-FORWARD (Keystone 2010) found GOL 50mg to be clinically effective over year 1 with ACR20, 50 and 70 response rates of 64.0%, 43.8% and 24.7%, respectively. Patients with active RA despite MTX therapy continued to benefit from treatment with GOL 50mg; 90.6% of those patients achieving ACR20 response at week 24 maintained the response at week 52. 61.4% showed DAS28 remission (≤2.6) at week 52 with 36.8% achieving sustained DAS28 remission.

GO-AFTER: TNFα Inhibitor Experienced Population

GO-AFTER Design

The GO-AFTER trial was a multicentre, randomised, double-blind, placebo-controlled phase III trial, designed to investigate the efficacy of the new TNFα inhibitor, golimumab, in patients with active rheumatoid arthritis (RA) who had previously received one or more other TNFα inhibitors. GO-AFTER was considered a landmark trial, as previous evidence for sequential use of TNFα inhibitors has been based upon analyses of registry data or small-scale unblinded observational studies

(Cohen et al 2005; Nikas et al 2006; Bennett et al 2005; Hyrich et al 2007; Furst et al 2007).

GO-AFTER Methods

461 patients from 82 sites across ten countries were recruited. All had been diagnosed with active RA (persistent disease activity with at least four swollen and four tender joints) according to the American College of Rheumatology (ACR) criteria (Arnett et al 1988) at least 3 months before screening. All had been treated with at least one dose of a TNFα inhibitor —etanercept, adalimumab, or infliximab— at least 8 weeks (adalimumab or etanercept) or 12 weeks (infliximab) before the first dose of the study drug. The previous TNFα inhibitor could have been discontinued for any reason, and the reason was categorised by investigators as 'lack of effectiveness', 'intolerance', or 'other'.

Concomitant disease-modifying anti-rheumatic drug (DMARD) treatment with methotrexate, sulfasalazine, and hydroxychloroquine (alone or in combination) was permitted but not required, as long as the DMARD had been tolerated for 12 weeks, in a dose that had been stable for 4 weeks. Patients receiving DMARDs at baseline were allowed to discontinue these drugs before starting the study, although if continued, the dose had to be maintained throughout the study. Oral corticosteroids (not exceeding the equivalent of 10 mg of prednisone per day) or non-steroidal anti-inflammatory drugs were also allowed if the doses had been stable for at least 2 weeks before the first dose of study drug.

The patients were randomised 1:1:1 into three groups: placebo injections, 50mg golimumab, and 100mg golimumab, all taken 4-weekly from week 0 to week 20. The randomisation was stratified by study site and baseline methotrexate usage. At week 16, patients with an inadequate response (less than 20% improvement from baseline in both tender and swollen joint counts) entered a double-blinded rescue therapy phase: patients receiving placebo, 50mg golimumab, or 100mg golimumab subsequently received 50mg golimumab, 100mg golimumab, or100mg golimumab, respectively.

GO-AFTER Endpoints

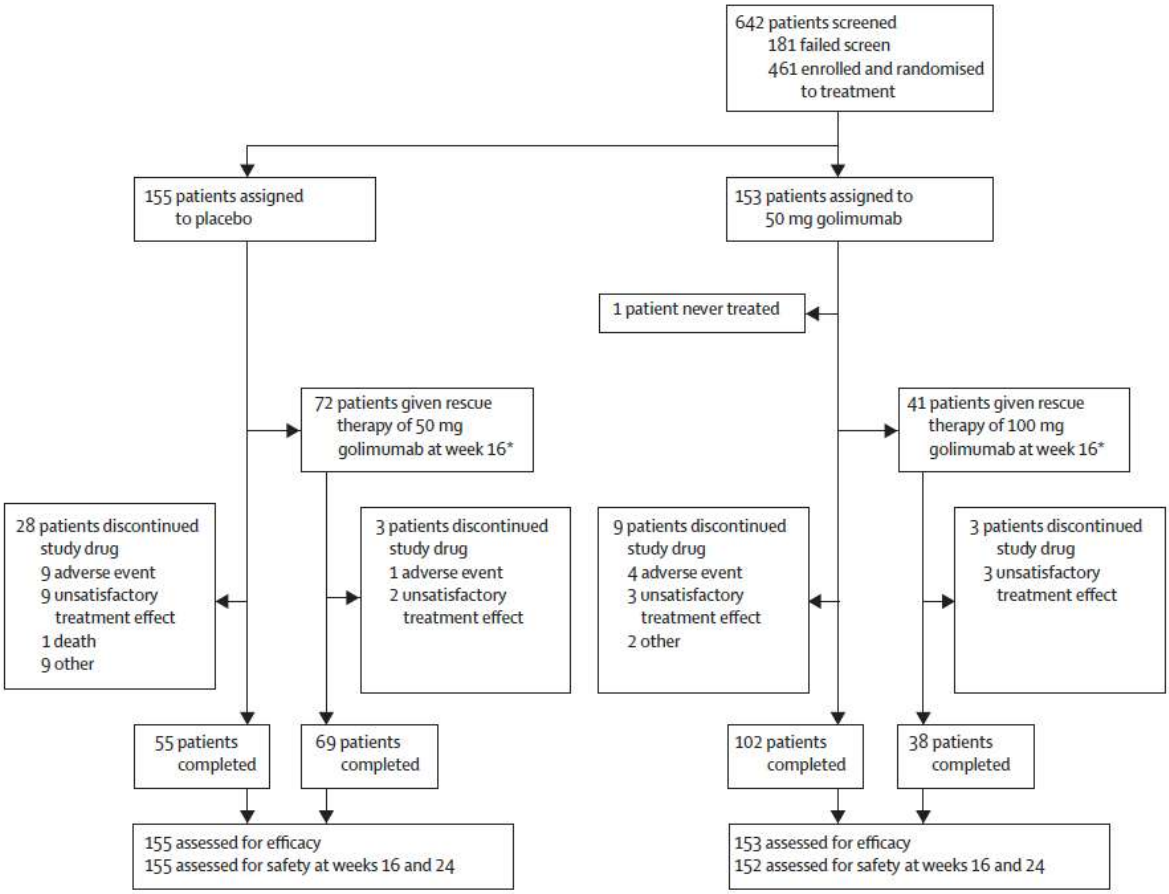
The primary efficacy endpoint was a ≥20% improvement in the ACR criteria for the assessment of RA (ACR20).

A wide range of secondary endpoints were also considered; the endpoints applicable to this submission are:

- ACR 50, 70, and 90 at week 14
- ACR 20, 50, 70, and 90 at week 24
- Improvement from baseline in HAQ score at week 24

GO-AFTER Patient flow

Figure 3. GO-AFTER patient flow



GO-AFTER Efficacy Results

461 patients were randomised, of whom 155 were assigned to the placebo group, and 153 were assigned to the 50mg golimumab group. One patient in the golimumab 50mg group was randomised in error, and excluded prior to receiving the study agent.

Note: for the purpose of this appraisal, only the placebo and 50mg groups will be discussed.

Table 121. Baseline demographics of placebo and golimumab 50mg groups

	Characteristic	Placebo	Golimumab 50mg
Subjects randomised	n	155	153
Age (years)	Mean (SD)	[redacted]	[redacted]
Women (years)	n (%)	132 (85.1%)	113 (73.8%)
Height (cm)†	Mean (SD)	[redacted]	[redacted]
Weight (kg)	Mean (SD)	[redacted]	[redacted]
RA disease duration (years)	Mean (SD)	[redacted]	[redacted]
	Median	9.80	9.55
Past TNFα inhibitor use			
Adalimumab	n (%)	85 (54.8%)	72 (47.1%)
Etanercept	n (%)	73 (47.1%)	76 (49.7%)
Infliximab	n (%)	83 (53.5%)	64 (41.8%)
Methotrexate treatment	n (%)	102 (65.8%)	103 (67.8%)
Methotrexate dose (mg/wk)	Mean (SD)	[redacted]	[redacted]
Corticosteroid use	n (%)	83 (53.5%)	92 (60.5%)
Prednisolone or equivalent dose (mg/wk)	Mean (SD)	6.9 (2.83)	6.9 (2.75)
Swollen joint count (0-66)	Mean (SD)	[redacted]	[redacted]
	Median	14	14
Tender joint count (0-68)	Mean (SD)	[redacted]	[redacted]
	Median	26	27
Patient assessment of pain	Mean (SD)	[redacted]	[redacted]
(VAS, 0-10cm)	Median	7	6.9

Patient's global assessment of disease	Mean (SD)	[redacted]	[redacted]
(VAS, 0-10cm)	Median	6.5	6.8
Physician's global assessment of disease	Mean (SD)	[redacted]	[redacted]
(VAS, 0-10cm)	Median	6.3	6.3
Assessment of physical function	Mean (SD)	[redacted]	[redacted]
(HAQ-DI, 0-3)	Median	1.7500	1.6250
C-reactive protein (mg/dL)	Mean (SD)	[redacted]	[redacted]
	Median	1	0.8

[redacted]

[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]		

GO-AFTER Primary endpoint results

Significantly more patients achieved the primary endpoint of ACR20 at week 14 in the golimumab 50mg group than the placebo group (35.3% vs 18.1%; p<0.001).

Among patients who discontinued one or more prior TNFα inhibitors due to lack of efficacy, a greater proportion of subjects achieved an ACR20 response in the golimumab 50mg group than the placebo group (35.7% vs 17.7%; p=0.006). A greater proportion of patients who discontinued one or more prior TNFα inhibitors due to intolerance achieved an ACR 20 response in the golimumab 50mg group than in the placebo group (34.6% vs 16.7%; p=0.154). Note the small numbers in this subgroup (Table 7).

Table 122. ACR20 responses at week 14, with stratification by reasons for discontinuation

ACR20 response at week 14	Placebo	Golimumab 50mg	p-value
All subjects			
n	155	153	
subjects in response, n (%)	28 (18.1%)	54 (35.3%)	p<0.001
Prior TNFα inhibitor discontinued due to lack of efficacy			
n	96	84	
subjects in response, n (%)	17 (17.1%)	30 (35.7%)	p=0.006
Prior TNFα inhibitor discontinued due to intolerance			
n	24	20	
subjects in response, n (%)	4 (16.7%)	7 (35.0%)	P=0.154

GO-AFTER Secondary endpoints: ACR responses at weeks 14 and 24

Significantly more patients achieved the secondary endpoints of ACR 50, 70, and 90 at week 14 and ACR 20, 50, 70, and 90 at week 24 in the golimumab 50mg group than the placebo group.

Table 123. Secondary endpoint ACR responses

Response	Placebo	Golimumab 50mg	p-value
	n=155	n=153	
ACR50 at week 14			
subjects in response, n (%)	10 (6.5%)	25 (16.3%)	p=0.006
ACR70 at week 14			
subjects in response, n (%)	3 (1.9%)	16 (10.5%)	p=0.002
ACR90 at week 14			
subjects in response, n (%)	0	3 (2.0%)	p=0.081
ACR20 at week 24			
subjects in response, n (%)	26 (16.8%)	52 (34.0%)	p<0.001
ACR50 at week 24			
subjects in response, n (%)	8 (5.2%)	28 (18.3%)	p<0.001
ACR70 as week 24			
subjects in response, n (%)	5 (3.2%)	18 (11.8%)	p=0.004
ACR90 at week 24			
subjects in response, n (%)	2 (1.3%)	7 (4.6%)	p=0.086

GO-AFTER Secondary endpoint: Improvement from baseline in HAQ-DI score at week 24

At week 24, significantly more patients in the 50mg golimumab group had a clinically important reduction in HAD-DI than in the placebo group (50.0% vs 34%, p=0.0044).

Table 124. Improvement from baseline HAQ-DI at week 24

HAQ-DI - improvement from baseline, at week 24	Placebo	Golimumab 50mg	p-value
	n=155	n=153	
Improvement from baseline			
Mean (SD)	[redacted]	[redacted]	[redacted]
Median	0	0.25	

GO-AFTER Safety Background

Safety was assessed by summarising the incidence and type of adverse events (AEs) by treatment goup (actual treatment received). A subject with an AE was counted in a

treatment goup based on the study agent the subject was receiving at the time of onset of the event.

GO-AFTER Safety Results

The adverse event profiles to week 24 were similar for the placebo and golimumab 50mg groups. Table 10 provides the adverse events profile to week 24. Fewer patients reported adverse events in the golimumab 50mg group, and there were no substantial differences in the frequency of common adverse events, infections, serious adverse events, serious infections, or malignancies between the two groups.

Table 125. Adverse events to week 24

	Placebo	Golimumab 50mg
	n=155	n=152*
Number of injections, mean (SD)	4.4 (1.3)	5.2 (1.2)
Patients reporting adverse events	112 (72%)	101 (66%)
Common adverse events		
Upper respiratory tract infection	10 (6%)	11 (7%)
Nasopharyngitis	11 (7%)	12 (8%)
Rheumatoid arthritis	16 (10%)	9 (6%)
Cough	5 (3%)	11 (7%)
Diarrhoea	7 (5%)	5 (3%)
Arthralgia	8 (5%)	6 (4%)
Sinusitis	7 (5%)	5 (3%)
Hypertension	2 (1%)	5 (3%)
Infections	51 (33%)	53 (35%)
Serious adverse events	15 (10%)	11 (7%)
Serious infections	5 (3%)	5 (3%)
Injection-site reactions	6 (4%)	9 (6%)
Malignancies	1 (1%)	1 (1%)
Data are number of patients (%) unless stated		
* patient randomised but excluded prior to treatment not included		

GO-AFTER Clinical Effectiveness Conclusion

GO-AFTER shows that a TNFα inhibitor reduced the signs and symptoms of active RA and improved physical function in patients who had previously received a TNFα

inhibitor(s), which suggests that switching patients from one TNFα inhibitor to another was effective and generally well tolerated.

In subjects with active RA who had previously received at least 1 dose of a TNFα inhibitor(s):

- A significantly greater proportion of subjects receiving golimumab achieved the primary endpoint (ie, an ACR 20 response at Week 14) compared with placebo. These patients further increased or stabilised during the remaining 24 week study period.
- Among subjects who had discontinued 1 or more previous TNFα inhibitors due to lack of efficacy, a significantly greater proportion in the golimumab group than in the placebo group achieved an ACR 20 response at Week 14.
- Golimumab was superior to placebo in all major secondary endpoints.
- Golimumab was generally well tolerated. In the placebo-controlled portion of the study, approximately equal proportions of subjects in the placebo and golimumab groups had at least 1 AE.

5.9.2 Please provide a summary of the strengths and limitations of the clinical-evidence base of the intervention.

Strengths of the clinical-evidence base include:

Inclusion of RCTs for analysis, the large number of parameters considered, TNFα inhibitors considered safe and efficacious based on earlier data.

RCT data for two populations: DMARD experienced and TNFα inhibitor experienced, the latter population which previously did not have RCT evidence for TNFα inhibitors.

Weakness:

Golimumab is a relatively new drug hence long-term data is awaited. In addition, there are no non-RCTs or observational studies either at this point of time.

5.9.3 Please provide a brief statement of the relevance of the evidence base to the decision problem. Include a discussion of the relevance of the outcomes assessed in clinical trials to the clinical benefits experienced by patients in practice.

The evidence laid out relates to the usefulness of golimumab in the treatment of rheumatoid arthritis with a DMARD- and TNFα inhibitor-experienced. In this regard, the intervention has been indirectly compared with other biologic treatments which are currently reimbursed in the UK. Clinical and safety benefits of the interventions have been compared on parameters such as ACR, HAQ, discontinuation and adverse events, all of which are of high clinical significance.

5.9.4 Identify any factors that may influence the external validity of study results to patients in routine clinical practice; for example, how the technology was used in the trial, issues relating to the conduct of the trial compared with clinical practice, or the choice of eligible patients. State any criteria that would be used in clinical practice to select patients for whom treatment would be suitable based on the evidence submitted. What proportion of the evidence base is for the dose(s) given in the SPC?

Golimumab has been reviewed on the basis of RCT evidence for this appraisal (GO-FORWARD & GO-AFTER); this may influence the ability to generalise the findings.

In the RCT considered, golimumab has been considered as the first line of treatment as well as after the exposure to a TNFα inhibitor; however in clinical practice patients may only be offered another biologicf treatment after failure on a TNFα inhibitor. Whilst the majority of GO-AFTER patients have been exposed to at least one TNFα inhibitor, a sub-section has failed on TNFα inhibitor treatment.

6 Cost effectiveness

6.1 Published cost-effectiveness evaluations

Identification of studies

6.1.1 Describe the strategies used to retrieve relevant cost-effectiveness studies from the published literature and from unpublished data held by the manufacturer or sponsor. The methods used should be justified with reference to the decision problem. Sufficient detail should be provided to enable the methods to be reproduced, and the rationale for any inclusion and exclusion criteria used should be provided. The search strategy used should be provided as in section 9.10, appendix 10.

A systematic review of the economic literature was conducted to identify published economic evaluation studies on therapies used in the treatment of rheumatoid arthritis from the following bibliographic databases:

- MEDLINE (R) In-Process & Other Non-indexed Citations (OVID)
- EMBASE (OVID)
- NHS Economic Evaluation Database, NHS EED (Centre for Reviews & Dissemination, CRDWeb)
- Database of Abstracts of Reviews of Effects, DARE (CRDWeb)
- Health Technology Assessment Database (CRDWeb)

Full details of the conducted search strategies are contained in Appendix 10, section 9.10. Inclusion and exclusion criteria applied for economic searches are presented in Table 126.

Table 126. Inclusion & exclusion criteria applied for cost-effectiveness review

Study Design	Cost-consequence analysis, cost-benefit analysis, cost-effectiveness analysis, cost-utility analysis, cost studies (UK only)
Population	Adults with Rheumatoid Arthritis; other forms of arthritis are excluded
Intervention	Golimumab, Infliximab, Etanercept, Adalimumab
Comparator	Standard care, other biologics
Outcome	Cost-effectiveness, cost-estimates (UK only)
Exclusion Criteria	Studies that did not fit within the inclusion criteria or studies with a juvenile population (aged 0-17 years)

Description of identified studies

6.1.2 Provide a brief overview of each study, stating the aims, methods, results and relevance to decision-making in England and Wales. Each study's results should be interpreted in light of a critical appraisal of its methodology. When studies have been identified and not included, justification for this should be provided. If more than one study is identified, please present in a table as suggested below.

Table 127. Summary list of other cost-effectiveness evaluations

Author	Year	Title	Summary of model	TNF Considered	Comparator	ICER
Bansback	2005	Cost effectiveness of adalimumab in the treatment of patients with moderate to severe rheumatoid arthritis in Sweden	Markov	Adalimumab, Etanercept, Infliximab	DMARD sequence	ACR50/DAS28 good: [Adal+MTX]:€34,167/QALY, [Adal]: €41,567/QALY, [Etan+MTX]: €37,760/QALY, [Etan]: €36,927/QALY, [Infl+MTX]: €48,333/QALY ACR20/DAS28 moderate: [Adal+MTX]: €40,875/QALY, [Adal]: €65,499/QALY, [Etan+MTX]: €51,976/QALY, [Etan]: €42,480/QALY, [Infl+MTX]€64,935/QALY
Barbieri	2005	The cost-effectiveness of infliximab for severe treatment resistant rheumatoid arthritis in the UK	Markov	Infliximab	Methotrexate	[Infl+MTX]: £33,618/QALY
Brennan	2007	Modelling the cost effectiveness of TNF-alpha antagonists in the management of rheumatoid arthritis: results from the British Society for Rheumatology Biologics Registry	Decision Tree	Adalimumab, Etanercept, Infliximab (considered as a class)	DMARD sequence	£23,882/QALY
Brennan	2004	Modelling the cost-effectiveness of etanercept in adults with rheumatoid arthritis in the UK	Patient level simulation	Etanercept	DMARD sequence	£16,330/QALY

Chen	2006	A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness	Birmingham Rheumatoid Arthritis Model (BRAM)	Adalimumab, Etanercept, Infliximab	Placebo	As last active therapy: [Etan]: £24,000/QALY, [Adal]: £30,000/QALY, [Infl]: £38,000/QALY
Chiou	2004	Cost-effectiveness analysis of biological treatments for rheumatoid arthritis	Decision Tree	Adalimumab, Etanercept, Infliximab	Anakinra	[Adal]: Dominated, [Adal+MTX]: Dominated, [Etan]: US$13,387/QALY, [Etan+MTX]: US$7,925/QALY, [Infl+MTX}: Dominated
Choi	2002	A cost effectiveness analysis of treatment options for methotrexate-naive rheumatoid arthritis	Decision Tree	Etanercept	Leflunomide methotrexate, sulfasalazine, no second line agent	US$41900/ACR20, US$40800/ACR70WR
Jobanputra	2002	The effectiveness of infliximab and etanercept for the treatment of rheumatoid arthritis: a systematic review and economic evaluation	Patient level simulation	Etanercept, Infliximab	DMARD sequence	[Etan]: £83095/QALY, [Infl]: £115937/QALY
Kobelt	2003	The cost-effectiveness of infliximab (Remicade) in the treatment of rheumatoid arthritis in Sweden and the United Kingdom based on the ATTRACT study	Markov	Infliximab	Methotrexate	€34,800/QALY
Kobelt	2004	TNF inhibitors in the treatment of rheumatoid arthritis in clinical practice: Costs and outcomes in a follow up study of patients with Ra treated with etanercept or infliximab in southern Sweden	N/A	Etanercept, Infliximab	Baseline = failed at least 2 DMARDs including methotrexate	After 3 mths treatment: €43,500/QALY, After 6 mths treatment: €36,900/QALY
Spalding	2006	Cost effectiveness of tumour necrosis factor-alpha inhibitors as first-line agents in rheumatoid	Markov	Adalimumab, Etanercept, Infliximab	Methotrexate	[Adal]: US$63,679/QALY, [Etan]: US$89,772/QALY, [Adal+MTX]: US$194,589/QALY, [Infl+MTX]:

		arthritis				$409,523/QALY
Wailoo	2008	Biologic drugs for rheumatoid arthritis in the Medicare program: a cost-effectiveness analysis	Decision Tree	Adalimumab, Etanercept, Infliximab	Adalimumab / Anakinra	[Etan]: v. Adalimumab: $92,058/QALY, [Adal]: v. Anakinra: $142,726/QALY. [Infl]: Dominated by both etanercept and adalimumab
Wong	2002	Estimating the cost-effectiveness of 54 weeks of infliximab for rheumatoid arthritis	Markov	Infliximab	Placebo & MTX	US$30,500/QALY

6.1.3 Please provide a complete quality assessment for each cost-effectiveness study identified. Use an appropriate and validated instrument, such as those of Drummond and Jefferson (1996) or Philips et al. (2004). For a suggested format based on Drummond and Jefferson (1996), please see section 9.11, appendix 11.

The quality assessment for each identified study is available in section 9.11, Appendix 11 as per the format of Drummond and Jefferson (1996).

6.2 *De novo analysis*

Patients

6.2.1 What patient groups are included in the economic evaluation? Do they reflect the licensed indication/CE marking or the population from the trials in sections 1.4 and 5.3.3, respectively? If not, how and why are there differences? What are the implications of this for the relevance of the evidence base to the specification of the decision problem? For example, the population in the economic model is more restrictive than that described in the (draft) SPC/IFU and included in the trials.

The model is designed to assess the cost-effectiveness of golimumab at different stages of the treatment pathway. The first analysis reflects the GO-FORWARD population and the second analysis reflects the GO-AFTER population. These Phase III trials have been conducted to determine the safety and efficacy of golimumab for the treatment of moderate to severe, active RA in adult patients. Both of these placebo-controlled trials make an assessment in patients at different stages of the treatment pathway in line with the licensed indication as well as the patient group included in the scope of this appraisal. There are therefore no specific implications of available evidence base to the specification of the decision problem.

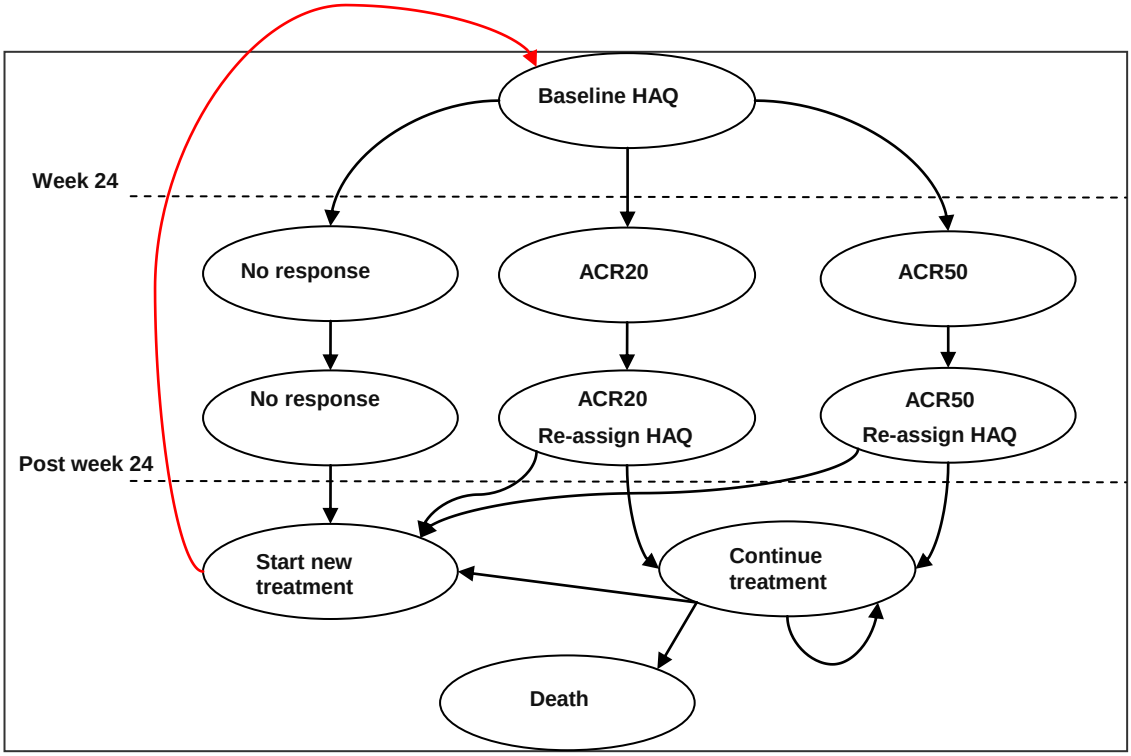
- GO-FORWARD (Keystone 2009) investigates the use of golimumab in patients who have failed treatment with MTX but have not been exposed to a TNFα inhibitor.
- GO-AFTER (Smolen 2009) makes an assessment of golimumab in patients who failed treatment with methotrexate and have been exposed to at least one TNFα inhibitor.

Model structure

6.2.2 Please provide a diagrammatical representation of the model you have chosen.

A diagram representing the sequence of events in model can be found in Figure 4.

Figure 4: Patient flow diagram



6.2.3 Please justify the chosen structure in line with the clinical pathway of care identified in section 2.4.

A Markov model was used to estimate the expected costs and QALYs of rheumatoid arthritis patients treated with golimumab as they are often utilised to model disease progression in chronic diseases. Many other economic models, including the Birmingham Rheumatoid Arthritis Model, have adopted patient simulation methods for rheumatoid arthritis. However, it is felt that the Markov model maintains a greater degree of simplicity and transparency. In this respect the assumptions of the model should be relatively easily communicated and the mechanics of the model straightforward to describe.

The model structure has been designed to simulate the treatment pathways of RA patients. On entering the model patients are assumed to have already received prior treatment for RA. The specific choice of treatment before entry into the model is dependent on the patient population. Patients can switch between a maximum of seven treatments within the model.

In line with published NICE TA guidance as of June 2010, the Markov model assesses approved treatment options within both of the DMARD experienced and TNFα inhibitor experienced populations.

6.2.4 Please define what the health states in the model are meant to capture.

The Markov model assigns patients to ACR response health states defined by treatment response in clinical trials. ACR, a health outcome measure on the American College of Rheumatology (ACR) 20 clinical criteria is defined as achieving (Rosenberg et al 2005):

- At least 20% improvement in the tender joint count, and
- At least 20% improvement in the swollen joint count, and
- At least 20% improvement in 3 of the following 5 assessments:
- Patient's global assessment of disease activity (VAS).
- Physician's global assessment if disease activity (VAS).
- Evaluator's global assessment of disease activity (VAS).
- Patient's assessment of physical function as measured by the Health Assessment Questionnaire disability index.
- Level of acute-phase reactant (CRP).

The ACR50 is defined as a 50% improvement in the parameters as described above for ACR20.

The use of ACR response rates to estimate effectiveness is similar to that employed in a recently published economic model for rituximab (Kielhorn et al 2008). Patient responses were linked to the Health Assessment Questionnaire Disability Index (HAQ-DI) score in order to estimate utilities and disease progression over the long term (Barton et al 2004; Scott et al 2002). HAQ-DI is a patient reported outcome instrument based on eight dimensions: dressing and grooming, arising, eating, walking, hygiene, reach, grip and activities (Fries et al 1980).

The DAS28 was not chosen as a measure of efficacy because insufficient studies had reported all the information required to conduct the mixed treatment comparison analysis. The ACR response rates capture the relative efficacy of RA treatment options. Disease severity at baseline and long-term progression is accounted for by the HAQ score in the model.

6.2.5 How does the model structure capture the main aspects of the condition for patients and clinicians as identified in section 2 (Context)? What was the underlying disease progression implemented in the model? Or what treatment was assumed to reflect underlying disease progression?

Disease Progression

At any point in the model a patient can be in a state of no treatment response, an ACR20-49 response, or an ACR50+ response. Numerous health states were defined to allow patients to change treatments over the model time horizon. This allows for the long-term extrapolation of patient pathways whilst utilising the short-term follow-up of clinical trials.

On entry into the model patients receive either golimumab or the appropriate comparator for the patient population under consideration. All patients in the cohort are allocated a baseline HAQ score based on the baseline characteristics of the GO-FORWARD and GO-AFTER trials. The patients are exposed to the first treatment in the sequence and are allocated to one of the two responder groups (ACR 20-49, ACR 50+) or to the non-responder group.

After 24 weeks of treatment patients are assigned to one of three health states: no response, ACR 20- 49 response, or ACR50 and greater response. At week 24 patients' HAQ scores are redefined and are dependent on response health state. Patients in the ACR 50+ health state will have the greatest improvement in HAQ. The changes in HAQ score for each health state were obtained through analysis of the golimumab trials.

After week 24 all patients who have not responded are started on a new treatment. In the next model cycle (6 months) a proportion of patients in the ACR20-49 and ACR50+ health states will withdraw due to adverse events and loss of efficacy. Those who maintain their response status will continue on treatment until loss of efficacy or death. Those patients who switch treatment due to loss of response, loss of efficacy or adverse event re-enter the model and are allocated a response status after the first cycle on the new treatment.

Within the base case, the model assumes that disease progression does not occur for patients responding to treatment. This was assumed within the BRAM as no differential deterioration was found between etanercept and infliximab, and that HAQ progression was found to be halted in patient who continued to receive etanercept or infliximab for 48 and 34 weeks respectively after the break of randomisation. However the uncertainty around this assumption is addressed in the one way sensitivity analysis where the HAQ of responders is assumed to progress at the same rate as natural history after the initial HAQ improvement.

The treatment pathway assumed that patients withdrawing from TNF-α inhibitor treatment move to palliation and not placebo. In common with the BRAM model, palliation was assumed to experience natural progression.

Treatment sequence

Patients who withdraw from a treatment due to no response, adverse event, or loss of response start a new treatment in the next cycle of the model. The treatment sequence for the BRAM model was used as an initial treatment sequence to be validated and edited by local clinicians. Current NICE guidelines were also consulted. The treatment sequences of the NICE model can be found in Table 128.

Table 128: BRAM treatment sequence

	Base case	**Adalimumab**	**Certolizumab**	**Etanercept**	**Infliximab**
1st line treatment	Methotrexate	Methotrexate	Methotrexate	Methotrexate	Methotrexate
2nd line treatment	Sulfasalazine + methotrexate	Sulfasalazine + methotrexate	Sulfasalazine + methotrexate	Sulfasalazine + methotrexate	Sulfasalazine + methotrexate
3rd line treatment	Leflunomide	Adalimumab + methotrexate	Certolizumab + methotrexate	Etanercept + methotrexate	Infliximab + methotrexate
4th line treatment	Gold	Leflunomide	Leflunomide	Leflunomide	Leflunomide

5th line treatment	Azathioprine	Gold	Gold	Gold	Gold
6th line treatment	Ciclosporin	Azathioprine	Azathioprine	Azathioprine	Azathioprine
7th line treatment	Palliative care	Ciclosporin	Ciclosporin	Ciclosporin	Ciclosporin
8th line treatment		Palliative care	Palliative care	Palliative care	Palliative care

In 2007 NICE issued guidance that adalimumab, etanercept and infliximab were recommended as options for treatment of rheumatoid arthritis in patients who had failed two DMARDs. Subsequently, in 2010, NICE recommended the use of certolizumab pegol in the same RA patient population. Therefore, current treatment practice in the UK for third line therapy is one of these four TNFα inhibitors.

A comparison with methotrexate has been chosen to estimate the cost-effectiveness of golimumab compared with DMARDs. Methotrexate was chosen because there is direct comparative data from the clinical trial and is considered standard practice in the UK. The treatment sequences used in the UK model are detailed in Table 129 and
Table 130. The treatment sequences represent a complete list of treatments initiated after diagnosis. The model starts at the initiation of golimumab. For the methotrexate experienced population the model begins at third line therapy (Table 129). For the TNFα experienced population the model begins at fourth line therapy (
Table 130).

The model allows patients to cycle through two palliative care settings (depending on whether the reintroduction of DMARDs begins at the 4th or 5th line of treatment). The methotrexate arm of the model will always allow patients to cycle through two palliative care settings. This structure was maintained in the model for two reasons. Firstly, the structure allows for a comparable number of treatment sequences for patients receiving methotrexate to the anti-TNF arms. Secondly, in clinical practice it is unlikely that patients would be left without treatment, therefore the additional cycle of palliative care accounts for instances where patients respond to other unspecified treatments. The impact of this assumption is likely to increase total QALYs and decrease total costs because more patients will respond to treatment, however the costs of drugs administered in palliative care are assumed to be zero.

Table 129: Golimumab model treatment sequence UK (DMARD experienced)

Treatment stage	**Golimumab arm**	**Other TNFα inhibitor arm**	**Methotrexate arm**
1st line treatment	*Methotrexate*	*Methotrexate*	*Methotrexate*
2nd line treatment	*Sulfasalazine + Methotrexate*	*Sulfasalazine + Methotrexate*	*Sulfasalazine + Methotrexate*
3rd line treatment	Golimumab + methotrexate	Anti-TNF + Methotrexate	Methotrexate
4th line treatment	Leflunomide	Leflunomide	Leflunomide

5th line treatment	Gold	Gold	Gold
6th line treatment	Azathioprine	Azathioprine	Azathioprine
7th line treatment	Ciclosporin	Ciclosporin	Ciclosporin
8th line treatment	Palliative care	Palliative care	Palliative care
9th line treatment	Palliative care	Palliative care	Palliative care

Table 130: Golimumab model treatment sequence UK (TNFα inhibitor experienced)

Treatment stage	**Golimumab arm**	**Other TNFα inhibitor arm**	**Methotrexate arm**
1st line treatment	*Methotrexate*	*Methotrexate*	*Methotrexate*
2nd line treatment	*Sulfasalazine + Methotrexate*	*Sulfasalazine + Methotrexate*	*Sulfasalazine + Methotrexate*
3rd line treatment	*Anti-TNF + Methotrexate*	*Anti-TNF + Methotrexate*	*Methotrexate*
4th line treatment	Golimumab + Methotrexate	Rituximab + Methotrexate	Methotrexate
5th line treatment	Leflunomide	Leflunomide	Leflunomide
6th line treatment	Gold	Gold	Gold
7th line treatment	Azathioprine	Azathioprine	Azathioprine
8th line treatment	Ciclosporin	Ciclosporin	Ciclosporin
9th line treatment	Palliative care	Palliative care	Palliative care
10th line treatment	Palliative care	Palliative care	Palliative care

6.2.6 Please provide a table containing the following information and any additional features of the model not previously reported. A suggested format is presented below.

Table 131. Key features of analysis

Factor	**Chosen values**	**Justification**
Time horizon	45 years	Lifetime model with patient starting age of 50 (DMARD experienced) and 54 (TNFα experienced)
Cycle length	24 weeks	Clinical practice and previous NICE RA guidance
Half-cycle correction	Yes	
Were health effects measured in QALYs; if not, what was used?	Yes	NICE reference case
Discount of 3.5% for utilities and costs	Yes	NICE reference case
Perspective (NHS/PSS)	UK NHS	NICE reference case

Each Markov cycle is set to 24 weeks, and the model can be run for the following time horizons: 1 year, 5 years, 10 years or lifetime. The base case time horizon is lifetime. Costs and utilities (calculated from the HAQ scores) were assigned to each

health state in the model. Costs and utilities were aggregated for the patient cohort over the time horizon. The analysis is conducted from the UK NHS perspective. Direct costs included the drug cost, administration cost, and heath care resource use. A half cycle correction is applied to patient transitions in the model. An incremental cost per ACR 20 responder at week 24 and cost per QALY analysis is conducted using the total cost and utilities aggregated for golimumab and its comparators.

Technology

6.2.7 Are the intervention and comparator(s) implemented in the model as per their marketing authorisations/CE marking and doses as stated in sections 1.3 and 1.5? If not, how and why are there differences? What are the implications of this for the relevance of the evidence base to the specified decision problem?

The comparators for golimumab 50 mg, administered monthly were biologic treatments that are currently reimbursed for use in the treatment of RA and have robust evidence of efficacy at the appropriate stage of the treatment pathway.

For the DMARD experienced population in the GO-FORWARD trial, the comparators comprise:

- Methotrexate (7.5mg once weekly);
- Adalimumab (40mg on alternate weeks);
- Certolizumab pegol (400mg repeated 2 weeks and 4 weeks after initial injection, then 200mg every 2 weeks);
- Etanercept (25mg twice weekly); and
- Infliximab (2.67 mg/kg, repeated 2 weeks and 6 weeks after initial infusion, then every 8 weeks).

For the TNFα experienced population in the GO-AFTER trial the following were considered the most appropriate comparators:

- Methotrexate (7.5mg once weekly); and
- Rituximab (2 infusions of 1000mg each, two weeks apart, repeated every 6 months)

Rituximab is currently approved for use within this population by NICE and has RCT evidence of efficacy in this stage of the treatment pathway (Cohen et al 2006). Because no head to head studies have been conducted, mixed treatment comparisons were performed to assess the relative safety and efficacy of golimumab against the five comparators identified above (adalimumab, certolizumab pegol, etanercept, infliximab and rituximab).

The intervention and all comparators were assessed within their marketing authorisations and approved doses. The two exceptions to this rule were in regards to the dosing schedule for infliximab and rituximab.

Infliximab dosing is based upon 2.67 vials to conservatively determine the average number of *full* vials that are used per patient. This figure was derived from the weight distributions of 3,208 patients registered within the BSRBR. The below table presents data which was used to derive the weighted average vials per infusion (2.67 vials).

Table 132. BSRBR weight distributions

	Weight (kg)			
Patient weight	<66 kg	66-100 kg	101-133kg	>134kg
% of patients	41%	52%	6%	1%
Vials per infusion	2	3	4	5
Weighted average vials per infusion:		**2.67**		

Rituximab dosing frequency was estimated as every 6 months based on current clinical practice. Two international surveys assessing rituximab found the majority of RA patients being treated with rituximab are re-dosed at least every six months (Section **Error! Reference source not found.**).

6.2.8 Please note that the following question refers to clinical continuation rules and not patient access schemes. Has a treatment continuation rule been assumed? If the rule is not stated in the (draft) SPC/IFU, this should be presented as a separate scenario by considering it as an additional treatment strategy alongside the base-case interventions and comparators. Consideration should be given to the following.

- The costs and health consequences of factors as a result of implementing the continuation rule (for example, any additional monitoring required).
- The robustness and plausibility of the endpoint on which the rule is based.
- Whether the 'response' criteria defined in the rule can be reasonably achieved.
- The appropriateness and robustness of the time at which response is measured.
- Whether the rule can be incorporated into routine clinical practice.
- Whether the rule is likely to predict those patients for whom the technology is particularly cost effective.

- Issues with respect to withdrawal of treatment from non-responders and other equity considerations.

Withdrawal from the first treatment in the model can be achieved in three ways:

- Patients who have not responded to treatment after 24 weeks may discontinue therapy
- Patients who achieved an ACR20 or ACR50 response after the first 24 weeks may start a different treatment in the next sequence due to intolerable adverse events.
- In subsequent cycles a long term drop-out rate from therapy is estimated using the Weibull distribution (details within section 8.15)

Withdrawal from subsequent treatments can occur in two ways in the model:

- Patients who have not responded to treatment after 24 weeks may discontinue treatment.
- In subsequent cycles patients who had previously responded may withdraw from treatment at a constant probability. These probabilities were estimated using long-term studies of rheumatoid arthritis treatment.

6.3 *Clinical parameters and variables*

6.3.1 Please demonstrate how the clinical data were implemented into the model.

Efficacy of golimumab

The proportion of patients achieving an ACR 20 or ACR 50 response after 24 weeks of treatment in each of the patient populations (MTX experienced and TNFα inhibitor experienced) were taken from the clinical study reports of the GO-FORWARD and GO-AFTER clinical trials. It should be noted, that patients who achieve an ACR 50 response rate will by definition have satisfied the criteria for an ACR 20 response. Therefore, in order to estimate the probability of transition into the mutually exclusive ACR 20 and ACR 50 health states it is necessary to subtract the number of ACR50 responders from the ACR 20.

Table 133. Efficacy of golimumab 50 mg group by trial

	MTX experienced (GO-FORWARD)	**TNFα experienced (GO- AFTER)**
No. randomized patients	89	153
No. ACR 20 responders	53	52
No. ACR 50 responders	34	28
Probability ACR20-49	0.213	0.157
Probability ACR50	0.382	0.183

Efficacy of comparators

The mixed treatment comparison computed risk ratios for the number of ACR 20 and ACR 50 responders. These ratios were applied to the number of ACR 20 and ACR 50 responders in the golimumab trials to derive the relative proportion of patients with an ACR 20 and ACR 50 response for each TNF comparator and rituximab. The same system of subtracting the ACR 50 responders from the ACR 20 group is used to estimate the risk ratios. The risk ratios used in the model are found in Section 5.5 and Section 5.6.

In all of the analyses, golimumab is compared with the placebo +/- MTX arm of the trial. All patients in GO-FORWARD and most patients (68%) in GO-AFTER were receiving methotrexate in the placebo arm. For the purpose of the model the placebo results are assumed to be equivalent to a methotrexate arm. The efficacy data for these analyses are detailed in Table 134.

Table 134: Efficacy of placebo +/- methotrexate

	MTX experienced	**TNFα experienced**
No. patients	133	155
No. ACR 20 responders	37	26
No. ACR 50 responders	18	8
Probability ACR20	0.143	0.116
Probability ACR50	0.135	0.052

6.3.2 Demonstrate how the transition probabilities were calculated from the clinical data. If appropriate, provide the transition matrix, details of the transformation of clinical outcomes or other details here.

Transition probabilities were calculated from golimumab RCT data and the mixed treatment comparisons. In order to estimate the probability of transition into the mutually exclusive ACR 20 and ACR 50 health states it is necessary to subtract the number of ACR50 responders from the ACR 20.

$$pnonrespon\ \ se = 1 - pACR\ 20 - pACR\ 50$$

Where pnonresponse = probability of non-response, pACR20 = probability of ACR20-49 response, pACR50 = probability of ACR50+ response. This ensures that the groups always sum to 100%.

The transition probabilities for golimumab and placebo are presented in Table 133 and Table 134, respectively. Comparator transition probabilities were derived by multiplying the comparator's MTC point estimate by the total number of responders within the placebo or treatment group of the GO-FORWARD/GO-AFTER trials. The model allows for both meta-analyses and MTC figures to be applied depending on the model control parameters.

Table 135 details the ACR 20 and ACR 50 transition probabilities associated with each treatment option within a TNFα inhibitor experienced population.

Table 135. Efficacy of UK treatments after anti-TNF failure

Treatment	No Response	ACR 20	ACR 50	Source
Rituximab	0.68	0.14	0.18	Cohen et al 2006
Leflunomide	0.54	0.20	0.26	Kremer et al 2002
Azathioprine	0.83	0.12	0.05	Assumed equal to MTX
Gold	0.83	0.12	0.05	Assumed equal to MTX
Ciclosporin	0.83	0.12	0.05	Assumed equal to MTX
Palliative care	0.83	0.12	0.05	Assumed equal to MTX
Methotrexate	0.83	0.12	0.05	GO-AFTER

6.3.3 Is there evidence that (transition) probabilities should vary over time for the condition or disease? If so, has this been included in the evaluation? If there is evidence that this is the case, but it has not been included, provide an explanation of why it has been excluded.

Long term withdrawal

No studies with sufficient follow-up were identified for golimumab, adalimumab or rituximab. The long term drop-out rates for golimumab were assumed equivalent to those for infliximab treated patients. This is a very conservative assumption given that the drop-out rate after 52 weeks of golimumab 50mg is very low in the GO-FORWARD clinical trial, only 6% at week 52. The long-term drop-out rates for rituximab and adalimumab from clinical trials are more aligned with the evidence available for infliximab. Only 55% and 18% of patients returned for a second and third course of rituximab in the safety and efficacy trial (Keystone et al 2007). Keystone (2004) report comparable drop-out rates at week 52 to those observed in a 52 week trial for infliximab (Keystone 2004; Lipsky 2000).

A summary of the probability of discontinuation due long term loss of efficacy parameters can be found in Table 136.

Table 136. First treatment withdrawal parameters

Long term discontinuation due to loss of efficacy				
Treatment	**Lambda (λ)**	**Gamma (γ)**	**Mean (years)**	**Source**
Golimumab	0.103	0.532	9 years	Assumed equal to infliximab
Adalimumab	0.103	0.532	9 years	Assumed equal to infliximab
Infliximab	0.103	0.532	9 years	Kristensen et al 2006
Etanercept	0.027	0.738	12 years	Kristensen et al 2006
Rituximab	0.103	0.532	9 years	Assumed equal to infliximab
Methotrexate	0.089	0.433	20 years	Edwards et al 2005

Mortality

At any stage of the model patients are at risk of death. National life tables for the UK were used to obtain age dependent mortality rates (Interim Life Tables 2006). Furthermore, the proportion of males and females recruited in the golimumab trials were used to estimate a weighted average mortality risk by sex. The mortality rates taken from national life tables were annual rates. They were adjusted to the model cycle length rate using the following equation:

$$r = -[\ln(1 - P)] / t$$

The cycle rates were transformed into transition probabilities using the following equation:

$$p = 1 - \exp\{-rt\}$$

Table 137. Subsequent treatment withdrawal parameters UK model

Treatment	% patients on treatment	Period of follow-up	Probability drop-out	Source
Rituximab	43%	5 years	0.08	Kristensen et al 2006*
Leflunomide	85%	1 year	0.08	Geborek 2002
gold	20%	5 years	0.17	Edwards et al 2005
Azathioprine	35%	5 years	0.11	Edwards et al 2005
Ciclosporin	34%	5 years	0.11	Edwards et al 2005

*Assumed equivalent to infliximab

6.3.4 Were intermediate outcome measures linked to final outcomes (for example, was a change in a surrogate outcome linked to a final clinical outcome)? If so, how was this relationship estimated, what sources of evidence were used, and what other evidence is there to support it?

The intermediate outcomes of change in HAQ and ACR response were linked to the final outcome of QALYs. The detailed methods of elicitation are available in section 6.4.3.

6.3.5 If clinical experts assessed the applicability of values available or estimated any values, please provide the following details:

- the criteria for selecting the experts
- the number of experts approached
- the number of experts who participated
- declaration of potential conflict(s) of interest from each expert or medical speciality whose opinion was sought
- the background information provided and its consistency with the totality of the evidence provided in the submission
- the method used to collect the opinions

- the medium used to collect opinions (for example, was information gathered by direct interview, telephone interview or self-administered questionnaire?)
- the questions asked
- whether iteration was used in the collation of opinions and if so, how it was used (for example, the Delphi technique).

Validation of the model structure and assumptions was conducted with two clinicians in the UK. Dr Ostor, Consultant Rheumatologist, Addenbrooke's Hospital, Cambridge, and Dr Andrews, Consultant Rheumatologist, Southampton University Hospitals were interviewed.

Two health economics experts were also consulted at the early stages of the development of the model and to critique the draft model. They were Professor Martin Buxton, Brunel University and Professor Stephen Morris, University College London. A third health economics expert, Stephen Palmer was consulted on the structure of the model and suggested a number of amendments to the model.

Summary of selected values

6.3.6 Please provide a list of all variables included in the cost-effectiveness analysis, detailing the values used, range (distribution) and source. Provide cross-references to other parts of the submission. Please present in a table, as suggested below.

Table 138. Summary of variables applied in the economic model

	Mean	Alpha	Beta	Distribution
Efficacy of Golimumab				
Golimumab GO-FORWARD – ACR20	0.213	20	69	BETA
Golimumab GO-FORWARD – ACR50	0.382	34	55	BETA
Golimumab GO-AFTER – ACR20	0.157	24	129	BETA
Golimumab GO-AFTER – ACR50	0.183	28	125	BETA
Efficacy of placebo				
Placebo GO-FORWARD – ACR20	0.143	19	114	BETA
Placebo GO-FORWARD – ACR50	0.135	18	115	BETA
Placebo GO-AFTER – ACR20	0.116	18	137	BETA
Placebo GO-AFTER – ACR50	0.052	8	147	BETA
Mixed treatment comparison				
	Median	**2.5th CI**	**97.5th CI**	**Distribution**
GOL vs. ADA ACR20	0.98	0.55	1.46	LOGNORMAL
GOL vs. IFX ACR20	1.05	0.57	1.65	LOGNORMAL
GOL vs. ETN ACR20	0.93	0.51	1.43	LOGNORMAL
GOL vs. RTX ACR20	0.71	0.42	1.20	LOGNORMAL
GOL vs. CTZ ACR20	0.72	0.41	1.06	LOGNORMAL
GOL vs. ADA ACR50	0.90	0.40	1.76	LOGNORMAL
GOL vs. IFX ACR50	0.99	0.40	2.04	LOGNORMAL
GOL vs. ETN ACR50	0.98	0.40	1.99	LOGNORMAL
GOL vs. RTX ACR50	0.66	0.25	1.76	LOGNORMAL
GOL vs. ADA DIS. AE	0.33	0.07	1.26	LOGNORMAL

GOL vs. IFX DIS. AE	0.29	0.06	1.17	LOGNORMAL
GOL vs. ETN DIS AE.	0.68	0.14	2.67	LOGNORMAL
GOL vs. RIT DIS. AE	0.15	0.02	0.91	LOGNORMAL
GOL VS. CTZ DIS AE	0.20	0.04	0.95	LOGNORMAL
Discontinuation rate				
	Mean	**Alpha**	**Beta**	**Distribution**
Rituximab	0.081	365	275	BETA
Leflunomide	0.080	19	106	BETA
Gold	0.159	123	26	BETA
Azathioprine	0.100	438	234	BETA
Ciclosporin	0.102	109	57	BETA
HAQ score				
	Mean	**2.5th CI**	**97.5th CI**	**Distribution**
Baseline HAQ GO-FORWARD	1.410	0.056	2.764	NORMAL
ΔHAQ non resp. GO-FORWARD	0.184	-0.709	1.077	NORMAL
ΔHAQ ACR20 GO-FORWARD	0.550	-0.448	1.548	NORMAL
ΔHAQ ACR50 GO-FORWARD	0.724	-0.388	1.835	NORMAL
Baseline HAQ GO-AFTER	1.587	0.307	2.867	NORMAL
ΔHAQ non resp. GO-AFTER	0.099	-0.813	1.012	NORMAL
ΔHAQ ACR20 GO-AFTER	0.375	-0.351	1.101	NORMAL
ΔHAQ ACR50 GO-AFTER	0.701	-0.425	1.827	NORMAL
Costs – UK (£)				
Full blood test	2.71	2.03	3.39	LOGNORMAL
Erythrocyte sedimentation rate	1.42	1.07	1.78	LOGNORMAL
Biochemical profile	1.42	1.07	1.78	LOGNORMAL
TB test	3.48	2.61	4.35	LOGNORMAL
CRP	1.42	1.07	1.78	LOGNORMAL
Hep B and Hep C	3.48	2.61	4.35	LOGNORMAL
Urinalysis	1.07	0.80	1.34	LOGNORMAL
Nurse practitioner	7.34	5.51	9.18	LOGNORMAL

GP visit	17.83	13.37	22.29	LOGNORMAL
Per hour admin cost	14.99	11.24	18.74	LOGNORMAL
	Mean	Alpha	Beta	Distribution
Chest X-ray	32	4.37	7.32	GAMMA
Rheumatologist	132	13.21	10.00	GAMMA
Chest X-ray	66	4.00	16.50	GAMMA
Utility				
	Mean	**2.5th CI**	**97.5th CI**	**Distribution**
Constant	0.862	0.795	0.929	NORMAL
	-0.327	-0.366	-0.288	NORMAL
Mortality				
	Mean	**2.5th CI**	**97.5th CI**	**Distribution**
Standardised mortality ratio	1.65	1.34	1.98	NORMAL
Other Parameters				
Time horizon	43 years (Range 1 – 43 years)			
Cycle length	6 months			
Hospitalisation costs	Yes			
Rebound assumption	Equal to gain			
Discount rate Costs / Outcomes	3.5% / 3.5%			
Proportion of females	80%			
Age GO-FORWARD	50			
Age GO-AFTER	54			

6.3.7 Are costs and clinical outcomes extrapolated beyond the trial follow-up period(s)? If so, what are the assumptions that underpin this extrapolation and how are they justified? In particular, what assumption was used about the longer term difference in effectiveness between the intervention and its comparator? For the extrapolation of clinical outcomes, please present graphs of any curve fittings to Kaplan-Meier plots.

The clinical effectiveness estimated using 24 week trial data was extrapolated to the model time horizon of 43 years. Following assumptions have been made.

- In order to account for the difference in cost between initiation of treatment and maintenance treatment, the cost of the first cycle of treatment is aggregated separately to the cost of subsequent cycles of treatment
- Health care resource costs are incurred at all stages of the model.
- Patients responding to treatment at 24 weeks were assumed to continue with their current treatment with an annual probability of withdrawing from treatment and moving onto DMARDs.
- Patients not responding to treatment at 24 weeks were assumed to withdraw treatment and move to DMARDs
- Patients who achieved an ACR20 or ACR50 response after the first 24 weeks may start a different treatment in the next sequence due to intolerable adverse events.
- The model assumes that there is a constant risk of HAQ progression for RA patients. The rate of increase in the HAQ for patients receiving DMARDs is taken from the NICE appraisal model (BRAM). In this model the HAQ score declines at a rate of 0.045 per year if a patient is receiving normal DMARDs. Patients receiving palliative care have a HAQ progression two times that of patients responding to DMARDs, at 0.09 per year. The model assumes that anti-TNF treatment halts disease progression.
- The model adjusts the HAQ score for each cycle to account for the HAQ score of patients transitioning into the health state from previous lines of therapy. Current HAQ score for each health state and cycle number is therefore a function of response status and HAQ decrement from baseline. HAQ decrement from baseline is estimated as a function of time on treatment with anti-TNFs and time on treatment with DMARDs. For each cycle of the model the proportion of patients on treatment 2 this cycle who were receiving treatment 1 in the previous cycle is recorded. HAQ decrement applied to the health state is a weighted average of the total HAQ decrement from baseline for patients who were receiving treatment 1 in the previous cycle and total HAQ decrement for patients who were receiving treatment 2 in the previous cycle plus HAQ decrement incurred this cycle.

6.3.8 Provide a list of all assumptions in the de novo economic model and a justification for each assumption.

The following assumptions were made within the economic analysis:

- Long-term HAQ progression is lower for those patients on anti-TNF therapy.
- Golimumab has a long-term withdrawal rate equal to that of infliximab.
- The model allows patients to cycle through two palliative care settings
- Patients who achieved an ACR20 or ACR50 response after the first 24 weeks may start a different treatment in the next sequence due to intolerable adverse events. This probability occurs only after the first cycle of treatment under the assumption that adverse events are most likely to lead to withdrawal in the early stages of treatment
- Patients rebound equal to gain upon failing anti-TNF treatment

6.4 Measurement and valuation of health effects

Patient experience

6.4.1 Please outline the aspects of the condition that most affect patients' quality of life.

RA has a high burden of morbidity and mortality as shown in a 2002 World Health Organization (WHO) report attributing 1.5 deaths per 100,000 population to RA in the UK (Lundkvist et al 2008). RA is associated with a reduced life expectancy of about 3 to 7 years, most attributable to complications such as cardiovascular disease, renal disease and infection (Symmons 2006). As onset of RA at an earlier age may be associated with poor prognosis, the relative risk of mortality increases with younger age at onset: 35%-50% increased risk of mortality in adults ≥65 years and 60%-75% increased risk in adults 25-64 years (Symmons 2006).

A longitudinal inception cohort study in deceased patients from 1986-1997 for Early RA in England and Wales found that the survival rate of patients in the first 7 years of RA was lower than that expected based on the general population of England and Wales in 1986-2002 (all-cause standardized mortality ratio [SMR] 1.27, 95% CI: 1.04, 1.46) (Young et al 2007). From the examination of death certificates in the 32% of deceased RA patients, cardiovascular disease, and ischemic heart disease specifically, was found to be the most common primary cause of death, with a mean survival time 1-2 years lower for RA patients with ischemic heart disease compared with all other major causes of death (Kvien et al 2004).

6.4.2 Please describe how a patient's HRQL is likely to change over the course of the condition.

Assessment of physical function in RA is commonly evaluated using the Health Assessment Questionnaire Disability Index (HAQ-DI; score range, 0 to 3 where 0 indicates no physical disability). Physical function generally worsens as the number of inflamed joints increases (eg, from 1-5 joints to 6-20 joints) and as disease activity worsens (Disease Activity Score 28).

HRQL data derived from clinical trials

6.4.3 If HRQL data were collected in the clinical trials identified in section 5 (Clinical evidence), please comment on whether the HRQL data are consistent with the reference case. The following are suggested elements for consideration, but the list is not exhaustive.

- Method of elicitation.
- Method of valuation.
- Point when measurements were made.
- Consistency with reference case.
- Appropriateness for cost-effectiveness analysis.
- Results with confidence intervals.

[redacted]

Studies have shown that the HAQ is strongly correlated with measures of health-related quality of life (Hurst 1997). Linear transformations between the HAQ and utility have been widely used in rheumatoid arthritis cost-effectiveness models. In the golimumab economic model patients' HAQ score over time represented an intermediate outcome of the model and is linked to QALYs using a published equation.

A search of the literature identified five studies assessing the relationship between HAQ and health-related quality of life (Hurst 1997; Bansback 2007; Bansback 2005; Ariza-Ariza 2006; Witney 2006). Only four of these looked specifically at the EQ-5D, NICE's preferred healthy utility instrument. Bansback et al. (2004) relates utility measured by the Health Utility Index-III (HUI-3) to HAQ (Bansback et al 2005). This is not used in the model. Of the remaining four only two contained sufficient information to use in the model.

Witney et al (2006) and Ariza-Ariza et al (2006) looked at the relationship between the EQ-5D and HAQ but these analyses focus on the correlation between the

variables using Pearson's correlation (Witney 2006). In Ariza-Ariza et al (2006) a regression model is estimated, but the full results are not reported.

In the base case the equation used is calculated from a study by Hurst et al. (1997). The equation is described in the NICE report addressing the structural issues of the BRAM (Barton 2004). This is chosen to be consistent with the model used in the NICE appraisal of adalimumab, infliximab and etanercept. The model represents a simple linear equation with HAQ as the only independent variable.

$$EQ - 5D = 0.862 - 0.327 * HAQ$$

This equation estimates the following utilities by health state for the first cycle of the model Table 139.

Table 139. Utility scores week 24

Health state	Methotrexate experienced (GO-FORWARD)	Anti-TNF experienced (GO-AFTER)
Baseline	0.401	0.343
Non responder	0.461	0.376
ACR 20	0.581	0.466
ACR 50	0.638	0.572

Mapping

6.4.4 If mapping was used to transform any of the utilities or quality-of-life data in clinical trials, please provide the following information.

- Which tool was mapped from and onto what other tool? For example, SF-36 to EQ-5D.
- Details of the methodology used.
- Details of validation of the mapping technique.

The details of the mapping exercise are available in section 6.4.3.

HRQL studies

6.4.5 Please provide a systematic search of HRQL data. Consider published and unpublished studies, including any original research commissioned for this technology. Provide the rationale for terms used in the search strategy and any inclusion and exclusion criteria used. The search strategy used should be provided in section 9.12, appendix 12.

Studies have shown that the HAQ is strongly correlated with measures of health-related quality of life (Hurst 1997). Linear transformations between the HAQ and utility have been widely used in rheumatoid arthritis cost-effectiveness models. In

the golimumab economic model patients' HAQ score over time represented an intermediate outcome of the model and is linked to QALYs using a published equation.

A search of the literature identified five studies assessing the relationship between HAQ and health-related quality of life (Hurst 1997; Bansback 2007; Bansback 2005; Ariza-Ariza 2006; Witney 2006). Only four of these looked specifically at the EQ-5D, NICE's preferred healthy utility instrument. Bansback et al. (2004) relates utility measured by the Health Utility Index-III (HUI-3) to HAQ (Bansback et al 2005). This is not used in the model. Of the remaining four only two contained sufficient information to use in the model.

Witney et al (2006) and Ariza-Ariza et al (2006) looked at the relationship between the EQ-5D and HAQ but these analyses focus on the correlation between the variables using Pearson's correlation (Witney 2006). In Ariza-Ariza et al (2006) a regression model is estimated, but the full results are not reported.

In the base case the equation used is calculated from a study by Hurst et al. (1997). The equation is described in the NICE report addressing the structural issues of the BRAM (Barton 2004). This is chosen to be consistent with the model used in the NICE appraisal of adalimumab, infliximab and etanercept. The model represents a simple linear equation with HAQ as the only independent variable.

6.4.6 Provide details of the studies in which HRQL is measured. Include the following, but note that the list is not exhaustive.

- Population in which health effects were measured.
- Information on recruitment.
- Interventions and comparators.
- Sample size.
- Response rates.
- Description of health states.
- Adverse events.
- Appropriateness of health states given condition and treatment pathway.
- Method of elicitation.
- Method of valuation.
- Mapping.

- Uncertainty around values.
- Consistency with reference case.
- Appropriateness for cost-effectiveness analysis.
- Results with confidence intervals.
- Appropriateness of the study for cost-effectiveness analysis.

HRQL from GO-FORWARD and GO-AFTER are unavailable at the time of submission.

6.4.7 Please highlight any key differences between the values derived from the literature search and those reported in or mapped from the clinical trials.

Not applicable.

Adverse events

6.4.8 Please describe how adverse events have an impact on HRQL.

Adverse events are incorporated into the model based on the proportion of patients who discontinue treatment due to adverse events in the first 24 weeks. Adverse events are assumed to be class related therefore the costs and utility outcomes are assumed to be equivalent between the anti-TNFs. Furthermore, it is possible that adverse event disutility associated with rheumatoid arthritis treatment was already incorporated into the mapping equation from HAQ to utility.

Quality-of-life data used in cost-effectiveness analysis

6.4.9 Please summarise the values you have chosen for your cost-effectiveness analysis in the following table, referencing values obtained in sections 6.4.3 to 6.4.8. Justify the choice of utility values, giving consideration to the reference case.

Table 140. Summary of quality of life values for cost effectiveness analysis

State	Regression estimate	SE	Ref in submission	Justification
Constant	0.862	0.034	Section 6.4.3	Chen et al 2006
HAQ Coefficient	-0.327	0,0201	Section 6.4.3	Chen et al 2006

6.4.10 If clinical experts assessed the applicability of values available or estimated any values, please provide the following details:

- the criteria for selecting the experts
- the number of experts approached

- the number of experts who participated
- declaration of potential conflict(s) of interest from each expert or medical speciality whose opinion was sought
- the background information provided and its consistency with the totality of the evidence provided in the submission
- the method used to collect the opinions
- the medium used to collect opinions (for example, was information gathered by direct interview, telephone interview or self-administered questionnaire?)
- the questions asked
- whether iteration was used in the collation of opinions and if so, how it was used (for example, the Delphi technique).

Clinical experts were not used in the estimation of HRQL values.

6.4.11 Please define what a patient experiences in the health states in terms of HRQL. Is it constant or does it cover potential variances?

The HRQL in a particular health state is determined by the HAQ and ACR response. The algorithm used in the economic analysis uses both the clinical assessment indicators (HAQ and ACR) to estimate the HRQL of the patient. Therefore, the potential variances in the disease activity and the resultant HRQL are captured by HAQ and ACR response and are reflected in the utility values over the course of the treatment.

6.4.12 Were any health effects identified in the literature or clinical trials excluded from the analysis? If so, why were they excluded?

Based on our literature search, no health effects identified in the literature and the clinical trials have been excluded.

6.4.13 If appropriate, what was the baseline quality of life assumed in the analysis if different from health states? Were quality-of-life events taken from this baseline?

The baseline quality of life is determined by the baseline HAQ which was derived from GO-FORWARD and GO-AFTER clinical trials. On entry into the model the patient cohort is assumed to have a baseline HAQ equivalent to that observed in the placebo and golimumab 50mg arms of the trials.

6.4.14 Please clarify whether HRQL is assumed to be constant over time. If not, provide details of how HRQL changes with time.

HRQL is not assumed to be constant over time and changes based upon HAQ and ACR response as described below.

After 1 cycle of the model patients are assigned to one of three health states: no response; ACR 20 response (excluding ACR 50 responders); and ACR 50 response. The HAQ scores corresponding to these three health states were estimated separately from the golimumab 50mg and placebo clinical trial data. Patients from each arm of the trial were separated into the three health state groups (no response, ACR 20-49, ACR 50+) at week 24 to calculate the mean change in HAQ. The model assumes that change in HAQ is related to treatment administered and response. Table 141 reports the HAQ scores used in all arms of the model. All non-methotrexate arms of the model we assumed to have the same change in HAQ as observed in the golimumab 50mg arm. The change in HAQ for methotrexate was taken from the placebo arm of the GO-FORWARD trial.

Table 141. HAQ scores

Health state	**Methotrexate experienced (GO-FORWARD)**	**Anti-TNF experienced (GO-AFTER)**
Baseline	1.41	1.59

GOL treated Non responder	1.23	1.49
GOL treated ACR 20	0.86	1.21
GOL treated ACR 50	0.69	0.89
GOL treated Non responder	1.44	N/A
GOL treated ACR 20	1.01	N/A
GOL treated ACR 50	0.68	N/A

The model assumes that there is a constant risk of HAQ progression for RA patients. The rate of increase in the HAQ for patients receiving DMARDs is taken from the NICE appraisal model (Chen et al 2006). In this model the HAQ score declines at a rate of 0.045 per year if a patient is receiving normal DMARDs. Patients receiving palliative care have a HAQ progression two times that of patients responding to DMARDs, at 0.09 per year. The model assumes that anti-TNF treatment halts disease progression. This assumption is aligned with comments from the NICE technology appraisal TA130 which states that it is "appropriate to primarily examine the estimates of cost-effectiveness based on the assumption of no HAQ progression while on TNF- α inhibitor therapy, while acknowledging the effects on the estimates of incorporating different assumptions of HAQ progression" (TA130 2007).

The Markov structure of the model applies a single HAQ score to all patients in each health state of the model. This poses problems when patients switch from an anti-TNF health state to a DMARD health state. For example patient A may fail anti-TNF and transition into a DMARD health state in cycle 2, whereas patient B stays on anti-TNF treatment longer and switches to DMARDs in cycle 10. The Markov model dictates that the same HAQ score must be applied to both patients. It is possible to assume that when patients fail anti-TNF they rebound back to natural history. In this scenario HAQ is assumed to return to the level and subsequent trajectory it would have been had they not initially responded to a particular therapy, therefore patient B will have the same HAQ score as patient A after cycle 10. This assumption is likely to underestimate the long term gains of anti-TNF treatment. Consequently, the model adjusts the HAQ score for each cycle to account for the HAQ score of patients transitioning into the health state from previous lines of therapy.

Current HAQ score for each health state and cycle number is therefore a function of response status and HAQ decrement from baseline. HAQ decrement from baseline is estimated as a function of time on treatment with anti-TNFs and time on treatment with DMARDs. For each cycle of the model the proportion of patients on treatment 2 this cycle who were receiving treatment 1 in the previous cycle is recorded. HAQ decrement applied to the health state is a weighted average of the total HAQ decrement from baseline for patients who were receiving treatment 1 in the previous cycle and total HAQ decrement for patients who were receiving treatment 2 in the previous cycle plus HAQ decrement incurred this cycle.

Therefore, if in cycle 10 25% of patients currently on treatment 2 were receiving treatment 1 in cycle 9 the total HAQ decrement from baseline for these patients is estimated as the HAQ decrement accrued up to cycle 9 plus the per cycle decrement

associated with treatment 2. The HAQ decrement for the remaining 75% of the patients on treatment 2 would be the total HAQ decrement accrued for this health state plus the per cycle decrement associated with treatment 2.

6.4.15 Have the values in sections 6.4.3 to 6.4.8 been amended? If so, please describe how and why they have been altered and the methodology.

The values in sections 6.4.3 through 6.4.8 have not been amended.

6.5 *Resource identification, measurement and valuation*

NHS costs

6.5.1 Please describe how the clinical management of the condition is currently costed in the NHS in terms of reference costs and the payment by results (PbR) tariff. Provide the relevant Healthcare Resource Groups (HRG) and PbR codes and justify their selection. Please consider in reference to section 2.

The current clinical management of this condition requires patients to have a regular contact with the specialist rheumatology centres in the UK. This involves regular attendance at an outpatient clinic and face to face consultation with a consultant or non-consultant in rheumatology department. Patients with moderate to severe symptoms may also be hospitalised occasionally.

Resource use was estimated in consultation with two expert clinicians in the UK. Initial resource use estimates were made based on the assumptions made in the BRAM. These were reviewed and validated or changed by the clinical experts. Recent guidelines from the American College of Rheumatology and the British Society for Rheumatology were also reviewed for consistency with our assumptions.

In order to determine the total treatment cost in the model, routine monitoring costs of patients is aggregated. In the UK patient monitoring includes visits to a rheumatologist after 6 months then every 12 months, general practitioner visits every 6 months, and a specialist nurse visit every 6 months.

Resource use costs for the UK were sourced from the NHS reference costs (2006), and the Personal Social Services Research Unit (PSSRU 2008). It is common in the UK for patients to regularly visit a specialist rheumatology nurse more frequently than their rheumatologist. Table 142 presents the unit cost of each health care resource used in the model. Patients incur health care costs at all stages of the model.

Table 142. Unit costs of health care resources UK (£)

Health care resource	Unit cost	Source
Rheumatologist	£109.00	NHS Reference Costs 2006
General practitioner	£31	PSSRU 2008
Specialist nurse	£34.00	PSSRU 2008
Nurse practitioner	£9.00	PSSRU 2008
Full blood count	£2.71	NHS PbR tariff 2008-2009
Erythrocyte Sedimentation rate	£2.71	NHS PbR tariff 2008-2009
Biochemistry profile	£1.42	NHS PbR tariff 2008
C—reactive protein	£2.71	NHS PbR tariff 2008
TB test	£3.48	NHS PbR tariff 2008
Hep B and Hep C	£3.48	NHS PbR tariff 2008
Urinalysis	£1.07	NHS PbR tariff 2008
Chest X-ray	£32.00	NHS Reference Costs 2006

6.5.2 Please describe whether NHS reference costs or PbR tariffs are appropriate for costing the intervention being appraised.

The NHS reference costs cover a wide variety of conditions related to rheumatology and also have wider geographical coverage. Due its generalisability, NHS reference costs are appropriate for costing the biologic treatments within NHS.

Resource identification, measurement and valuation studies

6.5.3 Please provide a systematic search of relevant resource data for the UK. Include a search strategy and inclusion criteria, and consider published and unpublished studies. The search strategy used should be provided as in section 9.13, appendix 13. If the systematic search yields limited UK-specific data, the search strategy may be extended to capture data from non-UK sources. Please give the following details of included studies:

- country of study
- date of study
- applicability to UK clinical practice
- cost valuations used in study
- costs for use in economic analysis
- technology costs.

A number of routine examinations and tests are included in the model. Table 143 presents the test required pre-treatment and frequency of monitoring tests for each treatment included in the model.

Table 143. Monitoring visits and surveillance in the UK

Treatment	Pre-treatment	Monitoring	Frequency
Golimumab + MTX	Full blood count	Full Blood Count	Weeks 2, 4, 8, 12, then every 2 months

Treatment	Pre-treatment	Monitoring	Frequency
	Erythrocyte sedimentation rate Biochemical profile Chest X-ray TB test CRP test	Erythrocyte sedimentation rate Biochemical profile	
Adalimumab + MTX	Full blood count Erythrocyte sedimentation rate Biochemical profile Chest X-ray TB test CRP test	Full Blood Count Erythrocyte sedimentation rate Biochemical profile	Weeks 2, 4, 8, 12, then every 2 months
Infliximab + MTX	Full blood count Erythrocyte sedimentation rate Biochemical profile Chest X-ray TB test CRP test	Full Blood Count Erythrocyte sedimentation rate Biochemical profile	Weeks 2, 4, 8, 12, then every 8 weeks
Etanercept + MTX	Full blood count Erythrocyte sedimentation rate Biochemical profile Chest X-ray TB test CRP test	Full Blood Count Erythrocyte sedimentation rate Biochemical profile	Weeks 2, 4, 8, 12, then every 2 months
Certolizumab + MTX	Full blood count Erythrocyte sedimentation rate Biochemical profile Chest X-ray TB test CRP test	Full Blood Count Erythrocyte sedimentation rate Biochemical profile	Weeks 2, 4 then every 2 months
Rituximab + MTX	Full blood count Erythrocyte sedimentation rate Biochemical profile	Full Blood Count Erythrocyte sedimentation rate Biochemical profile	Weeks 2, 4, 8, 12, then every 8 months

Treatment	Pre-treatment	Monitoring	Frequency
	Chest X-ray TB test CRP test Hep B and Hep C		
Leflunomide	Full blood count Erythrocyte sedimentation rate Biochemical profile Urinalysis CRP test	Full blood count	Every 2 weeks for 6 weeks, every 8 week
		Biochemical profile	Monthly for 6 months bi-monthly thereafter
Gold	Full Blood Count Erythrocyte sedimentation rate Biochemical Profile Urinalysis CRP test	Full Blood Count Biochemical Profile	Weekly for 6 months, monthly thereafter
Azathioprine	Full Blood Count Erythrocyte sedimentation rate Biochemical Profile CRP test	Full Blood Count Biochemical Profile	Weekly for 6 weeks, monthly thereafter
Ciclosporin	Full Blood Count Erythrocyte sedimentation rate Biochemical Profile Urinalysis CRP test	Full Blood Count Biochemical Profile (BCP)	Every 2 weeks for 4 month then BCP monthly

6.5.4 If clinical experts assessed the applicability of values available or estimated any values, please provide the following details:

- the criteria for selecting the experts
- the number of experts approached
- the number of experts who participated
- declaration of potential conflict(s) of interest from each expert or medical speciality whose opinion was sought
- the background information provided and its consistency with the totality of the evidence provided in the submission
- the method used to collect the opinions
- the medium used to collect opinions (for example, was information gathered by direct interview, telephone interview or self-administered questionnaire?)
- the questions asked
- whether iteration was used in the collation of opinions and if so, how it was used (for example, the Delphi technique).

Please see section 6.3.5 for details on clinical expert input regarding resource use.

Intervention and comparators' costs

6.5.5 Please summarise the cost of each treatment in the following table.

The drug costs for all treatments included in the model were extracted from sources relevant to the UK (BNF59). The unit costs for each drug can be found in Table 144. For the purpose of this model the per-cycle cost of golimumab was assumed to be equal to that of adalimumab.

Where drug doses are dependent on patient weight a mean weight of 73kg is assumed in line with the golimumab clinical trials.

Table 144. Unit drug costs

Treatment	Unit	UK cost per unit (£) (31)	Prescription dose	Dosing Frequency
Golimumab	50mg	£774.58	50mg	per month
Adalimumab	40mg pen or syringe	£357.50	40mg	every 2 weeks
Infliximab	100mg vial	£419.62	300mg[2]	every 8 weeks
Etanercept	25mg vial	£89.38	50mg	per week
Certolizumab	200 mg	£357.50	200 mg	every 2 weeks

Rituximab	50mL	£873.15	200mL	every 9 months
Leflunomide	10mg	£1.70	10-20mg	daily
Gold	50mg	£11.23	50mg	weekly
Azathioprine	50mg	£0.17	150mg	daily
Ciclosporin	100mg	£2.12	200mg	daily
Methotrexate	2.5mg	£0.12	7.5mg	weekly

[2]*Wastage assumed; Assumed actual dose based upon 2.67 vials (BSRBR; see Section 6.2.7)*

Total treatment costs were calculated by aggregating the drug cost and the cost of administration. Although many of the anti-TNFs are administered at home, patients are often initially taught how to administer treatment within a hospital, or doctor's office setting. This is calculated as a one-off visit to the specialist nurse.

For intravenous drugs (infliximab and rituximab) administration costs are higher and incurred at every administration of treatment. In the UK the cost of infusion is £34 with an additional £4.81 for every hour spent in the hospital. The cost of infusion is assumed equivalent to a visit to a specialist nurse plus an hourly charge for the care of the patient whilst they are on the ward.

In order to account for the difference in cost between initiation of treatment and maintenance treatment, the cost of the first cycle of treatment is aggregated separately to the cost of subsequent cycles of treatment. Table 145 reports the cost of each treatment included in the model.

Table 145: Acquisition and administration costs associated with the technology in the economic model

	Cost per dose	No. doses per first 6 months	No. doses post 6 months	Treatment cost first 6 months	Treatment cost pot 6 months	Cost per administration first 6 months	Total cost first 6 months	Total cost post 6 months
Golimumab	£774.58	6	6	£4,647.48	£4,647.48	£34.00	£4,681.48	£4,647.48
Adalimumab	£357.50	13	13	£4,647.50	£4,647.50	£34.00	£4,681.50	£4,647.50
Infliximab[A]	£419.62	13.35	8.6775	£5,601.93	£3,641.25	£55.00	£6,336.18	£4,118.52
Etanercept	£89.38	52	52	£4,647.76	£4,647.76	£34.00	£4,681.76	£4,647.76
Rituximab[B]	£873.15	6	4	£5,238.90	£3,492.60	£76.00	£5,694.90	£3,796.60
Certolizumab[C]	£357.50	6	13	£2,145.00	£4,647.50	£34.00	£2,179.00	£4,647.50
Leflunomide	£1.70	194.5	182.5	£331.43	£310.98	£0.00	£331.43	£310.98
Gold	£11.23	26	26	£291.98	£291.98	£0.00	£291.98	£291.98
Azathioprine	£0.17	547.5	547.5	£93.08	£93.08	£0.00	£93.08	£93.08
ciclosporin	£2.12	365	365	£773.80	£773.80	£0.00	£773.80	£773.80
Methotrexate	£0.12	78	78	£9.36	£9.36	£0.00	£9.36	£9.36

(A) Cost per dose based on 73kg patient, 194.91mg IFX (2.67 vials with wastage). No doses per first 6 months based on 2.67 vials (average full vials – BSRBR). Cost per administration based on SPC (1 hr infusions if initial 3 well received); (B) No doses based on 6 month dosing frequency. Cost per administration based on SPC (1st infusion ~3hrs, subsequent infusions ~2hrs); (C) No doses per first 6 months adjusted for PAS.

Health-state costs

6.5.6 Please summarise, if appropriate, the costs included in each health state. Cross-reference to other sections of the submission for the resource costs. Provide a rationale for the choice of values used in the cost-effectiveness model. The health states should refer to the states in section 6.2.4.

Hospitalisation costs

Cost offsets were incorporated into the model through reduced hospitalisations associated with TNFα inhibitor treatment. Data from Brennan (2007) were used to estimate the number of hospitalisations for every cycle of the model dependent on a number of characteristics, including anti-TNF treatment. The coefficients reported in Brennan (2007) are detailed in Table 146.

Table 146. Multivariate regression of number of days of hospital stay

Independent variable	Coefficient
Intercept	0.2351
Utility at baseline	-0.5467
Age (years)	0.0078
Disease duration	0.0075
Previous number of DMARDs	0.0648
Anti-TNF	-0.062

This analysis was incorporated into the model assuming that:

- Utility at baseline was 0.401.
- Age at baseline was 50 years.
- Disease duration at baseline was 7.33 years.
- Patients had received 2 previous DMARDs at baseline.

Example hospitalisation rates for patients receiving TNFα inhibitors and DMARDs in cycle 1 and cycle 10 are detailed in Table 147. A cost of inpatient stay was estimated to be £671 and was sourced from the 2006/7 NHS Reference Costs (PA34B).

Table 147. Hospital days with and without anti-TNF

	Anti-TNF		DMARD	
	No. hospital days per 6 months	Per cycle cost	No. hospital days per 6 months	Per cycle cost
Cycle 1	0.17	£114	0.54	£362
Cycle 10	0.25	£168	0.62	£416

Adverse-event costs

6.5.7 Please summarise the costs for each adverse event listed in section 5.9 (Adverse events). These should include the costs of therapies identified in section 2.7. Cross-reference to other sections of the submission for the resource costs. Provide a rationale for the choice of values used in the cost-effectiveness model discussed in section 6.2.2.

No additional cost for adverse events was included in the analysis. It was assumed that patients suffering from serious adverse events would withdraw from treatment and the cost of minor adverse events was included in the hospitalisation costs.

Miscellaneous costs

6.5.8 Please describe any additional costs that have not been covered anywhere else (for example, PSS costs). If none, please state.

No additional miscellaneous costs were considered.

6.6 Sensitivity analysis

6.6.1 Has the uncertainty around structural assumptions been investigated? Provide details of how this was investigated, including a description of the alternative scenarios in the analysis.

The uncertainty around structural assumptions has been investigated. The following assumptions were changed in scenario analysis.

Rebound equal to natural history HAQ progression:
The base case assumed rebound equal to gain. Therefore, patients withdrawing from treatment were assumed to return to baseline HAQ score and have natural history progression thereafter. In this scenario, it was assumed that patients withdrawing from treatment would return to HAQ score equal to natural history of primary non-responders. This is a pessimistic assumption as it assumes that patients lose all the benefit of TNF-α inhibitor immediately following treatment withdrawal.

6.6.2 Which variables were subject to deterministic sensitivity analysis? How were they varied and what was the rationale for this? If any parameters or variables listed in section 6.3.6 (Summary of selected values) were omitted from sensitivity analysis, please provide the rationale.

One-way sensitivity analysis is conducted to assess the impact of the following key parameters on the outcome of the model.

1. The model is run with a discount rate of 0% and 6%. Differential discount rates of 0% and 3.5% for costs and QALYs (and vice versa) were applied.
2. The model is run without the inclusion of hospitalisation costs.
3. The efficacy of golimumab (transition probability to ACR 20 and ACR50 health states) is increased and decreased by 20% of the base case values.
4. The standardised mortality ratio is varied according to the reported 95% confidence intervals.
5. The assumptions surrounding the rate of HAQ progression over the long-term were tested. The first analysis applied no HAQ decrement to any arms of the model. The second analysis applied an equal HAQ

decrement to anti-TNFs and non anti-TNFs of 0.0225 per cycle. The third analysis applied a HAQ progression rate of 0 for anti-TNFs and 0.0225 for non anti-TNFs. The fourth assumes that there is 0.015 HAQ decrement on anti-TNF per cycle and 0.0225 for DMARDs.

6. Patient age was varied from 50 in the base case to 45 and 60.
7. Long term discontinuation assumed equivalent to etanercept

In the base case the long-term drop out rate for golimumab is assumed to be equivalent to infliximab. However, evidence from the golimumab trials suggests that this may not be appropriate. In the infliximab 1 year trial 21% patients discontinued treatment at week 54. In contrast only 6% of patients in the golimumab 50mg arm discontinued treatment at week 52. These values are similar to the discontinuation rates observed in the etanercept TEMPO trial. The model is run with the assumption that golimumab has a similar long-term rate of discontinuation to etanercept rather than infliximab.

6.6.3 Was PSA undertaken? If not, why not? If it was, the distributions and their sources should be clearly stated if different from those in section 6.3.6, including the derivation and value of 'priors'. If any parameters or variables were omitted from sensitivity analysis, please provide the rationale for the omission(s).

All inputs that could be varied in the model were varied along chosen distributions to characterise the uncertainty in input parameters. All inputs included in the PSA and their measures of precision and distribution used can be found in Section 6.3.6. The Weibull parameters were sampled using the normal distribution. The variance was estimated from the Variance/Covariance matrix. Two thousand replications were generated in the PSA. Correlation between the lambda and gamma values from the Weibull distribution was accounted for. Correlation between the constant and coefficient utility estimates was not accounted for because there was insufficient data available from the literature.

6.7 Results

Clinical outcomes from the model

6.7.1 For the outcomes highlighted in the decision problem (see section 4), please provide the corresponding outcomes from the model and compare them with clinically important outcomes such as those reported in clinical trials. Discuss reasons for any differences between modelled and observed results (for example, adjustment for cross-over). Please use the following table format for each comparator with relevant outcomes included.

Table 148. Summary of model results compared with clinical data

Outcome	Clinical trial result	Model result
ACR20 response: DMARD experienced		QALYs gained for each

• Golimumab 50mg @ 24 weeks • Adalimumab eow @ 24 weeks • Certolizumab 200mg @24 weeks • Etanercept @ 24 weeks • Infliximab @ 24 weeks	60% (Keystone 2008) 75% (Kay 2008) 62% (Kim 2007) 46% (van de Putte 2004) 67% (Weinblatt 2003) 44% (Miyasaka 2008) 53% (Furst 2003 63% (Keystone 2004) 59% (Keystone 2008) 57% (Smolen 2009) 59% (Moreland 1999) 81% (Klareskog 2004) 40% (Maini 1998) 74% (Combe 2006) 55% (Westhovens 2006) 54% (Maini 1999)	individual treatment is in line with the published clinical data as etanercept was found to be the most effective (non-significant) whilst adalimumab was found to be the least (non-significant) for both DMARD experienced and TNFα inhibitor experienced patient populations.
ACR50 response: DMARD experienced • Golimumab 50mg @ 24 weeks • Adalimumab eow @ 24 weeks • Certolizumab 200mg @24 weeks • Etanercept @ 24 weeks • Infliximab @ 24 weeks	 37% (Keystone 2008) 39% (Kay 2008) 43% (Kim 2007) 22% (van de Putte 2004) 55% (Weinblatt 2003) 24% (Miyasaka 2008) 29% (Furst 2003 39% (Keystone 2004) 37% (Keystone 2008) 33% (Smolen 2009) 40% (Moreland 1999) 59% (Klareskog 2004) 52% (Combe 2006) 37% (Maini 1998) 31% (Westhovens 2006) 29% (Maini 1999)	
ACR20 response: TNFα inhibitor experienced • Golimumab 50mg @ 24 weeks • Rituximab @ 24 weeks	 35% (Smolen 2008) 51% (Cohen 2006)	
ACR50 response: TNFα inhibitor experienced • Golimumab 50mg @ 24 weeks	 16% (Smolen 2008)	

• Rituximab @ 24 weeks	27% (Cohen 2006)	

6.7.2 Please provide (if appropriate) the proportion of the cohort in the health state over time (Markov trace) for each state, supplying one for each comparator.

The markov traces are available in the MS Excel model accompanying this submission.

6.7.3 Please provide details of how the model assumes QALYs accrued over time. For example, Markov traces can be used to demonstrate QALYs accrued in each health state over time.

The markov traces are available in the MS Excel model accompanying this submission.

6.7.4 Please indicate the life years and QALYs accrued for each clinical outcome listed for each comparator. For outcomes that are a combination of other states, please present disaggregated results.

The model calculates the QALYs based on two pivotal intermediate outcomes; HAQ and ACR using an algorithm outlined in section 6.4.3. Therefore, it has not been possible to present disaggregated results.

6.7.5 Please provide details of the disaggregated incremental QALYs and costs by health state, and of resource use predicted by the model by category of cost. Suggested formats are presented below.

Not applicable.

Base-case analysis

6.7.6 Please present your results in the following table. List interventions and comparator(s) from least to most expensive and present ICERs in comparison with baseline (usually standard care) and then incremental analysis ranking technologies in terms of dominance and extended dominance.

Table 149. Incremental cost effectiveness results (DMARD experienced RA patient population)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) versus Baseline (Methotrexate)	Incremental analysis
Methotrexate	£35,869	4.569	-	-	-	-
Adalimumab	£66,875	5.792	£31,006	1.223	£25,353	£25,353
Golimumab	£67,747	5.827	£872	0.035	£25,346	£24,914

Infliximab	£69,899	5.651	£2,152	-0.176	£31,464	Dominated
Certolizumab	£73,571	5.768	£3,672	0.117	£31,444	£31,385
Etanercept	£74,208	6.133	£637	0.365	£24,514	£1,745

Table 150. Incremental cost effectiveness results (TNFα inhibitor experienced RA patient population)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) versus Baseline (Methotrexate)	Incremental analysis
Methotrexate	£33,673	3.129	-	-	-	-
Golimumab	£50,175	3.712	£16,502	0.583	£28,286	£28,286
Rituximab	£50,206	3.523	£31	-0.189	£41,935	Dominated

Sensitivity analyses

6.7.7 Please present results of deterministic sensitivity analysis. Consider the use of tornado diagrams.

Table 151. Results of deterministic sensitivity analysis (DMARD experienced)

Variable	Base case	Parameter change	ICER vs Methotrexate
Base Case ICER (Golimumab versus Methotrexate): £25,346			
Time horizon	43 years	5 years 10 years	£95,809 £56,221
Discount rate	3.5%	0% costs & 0% outcomes 0% costs & 3.5% outcomes 3.5% costs & 0% outcomes	£19,247 £31,191 £15,640
Hospitalisation Costs	Included	Excluded	£32,382
Age	50 yrs	45 yrs 60 yrs	£23,272 £32,681
Efficacy of Golimumab (ACR20)	0.213	-20% (0.170) +20% (0.256)	£26,041 £24,786
Efficacy of Golimumab	0.382	-20% (0.306) +20% (0.458)	£27,505 £23,900

(ACR50)			
Efficacy of Golimumab (ACR20 & ACR50)	0.213 / 0.382	-20% (0.170/0.306) +20% (0.256/0.458)	£28,692 £23,532
SMR	1.65	2.5th (1.34) 97.5th (1.98)	£24,382 £26,317
HAQ Progression	0 TNFs, 0.0225 DMARDs, 0.0450 Palliative care	0 for all 0.0225 for all 0 TNFs, 0.0225 non-TNFs 0.015 TNF, 0.0225 DMARDs, 0.0450 palliative care	£132,906 £115,795 £33,219 £39,055
Baseline HAQ score	1.41	- 50% (0.705) + 50% (2.115)	£25,323 £25,366
Golimumab acquisition cost	£774.58	- 20% (£620) + 20% (£929)	£18,797 £31,895
Natural history HAQ progression	0.0719	0.1018	£39,491
Long term withdrawal	Equal to infliximab	Equal to etanercept	£24,965

Table 152. Results of deterministic sensitivity analysis (TNFα inhibitor experienced)

Variable	Base case	Parameter change	ICER vs Methotrexate
Base Case ICER (Golimumab versus Methotrexate): £28,286			
Time horizon	43 years	5 years 10 years	£115,012 £61,537
Discount rate	3.5%	0% costs & 0% outcomes 0% costs & 3.5% outcomes 3.5% costs & 0% outcomes	£21,040 £32,367 £18,387
Hospitalisation Costs	Included	Excluded	£41,254
Age	54 yrs	49 yrs	£25,132

		64 yrs	£39,466
Efficacy of Golimumab (ACR20)	0.157	-20% (0.126) +20% (0.188)	£28,946 £27,760
Efficacy of Golimumab (ACR50)	0.183	-20% (0.146) +20% (0.220)	£30,677 £26,631
Efficacy of Golimumab (ACR20 & ACR50)	0.157 / 0.183	-20% (0.126/0.146) +20% (0.188/0.220)	£31,908 £26,329
SMR	1.65	2.5th (1.34) 97.5th (1.98)	£26,838 £29,740
HAQ Progression	0 TNFs, 0.0225 DMARDs, 0.0450 Palliative care	0 for all 0.0225 for all 0 TNFs, 0.0225 non-TNFs 0.015 TNF, 0.0225 DMARDs, 0.0450 palliative care	£146,172 £126,515 £36,067 £44,245
Baseline HAQ score	1.41	- 50% (0.705) + 50% (2.115)	£28,267 £28,302
Golimumab annual acquisition cost	£774.58	- 20% (£620) + 20% (£929)	£19,966 £36,598
Natural history HAQ progression	Equal to gain	Equal to natural history	£42,237
Long term withdrawal	Equal to infliximab	Equal to etanercept	£27,928

6.7.8 Please present the results of a PSA, and include scatter plots and cost-effectiveness acceptability curves.

Figure 5. CEAC for all TNFα inhibitors (DMARD experienced)

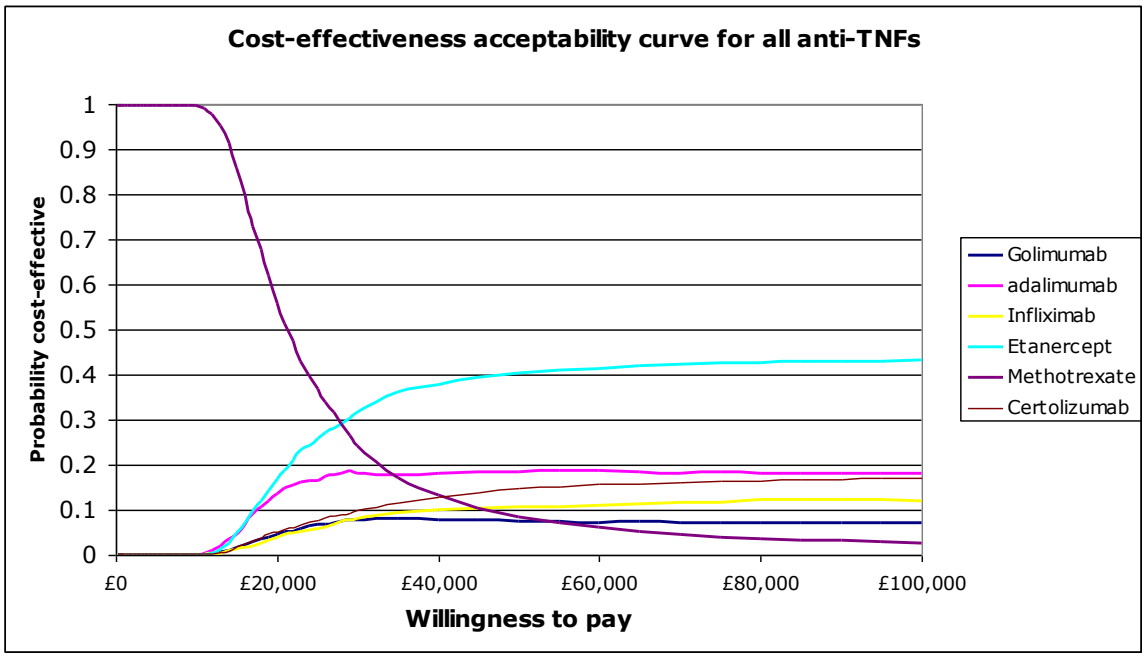
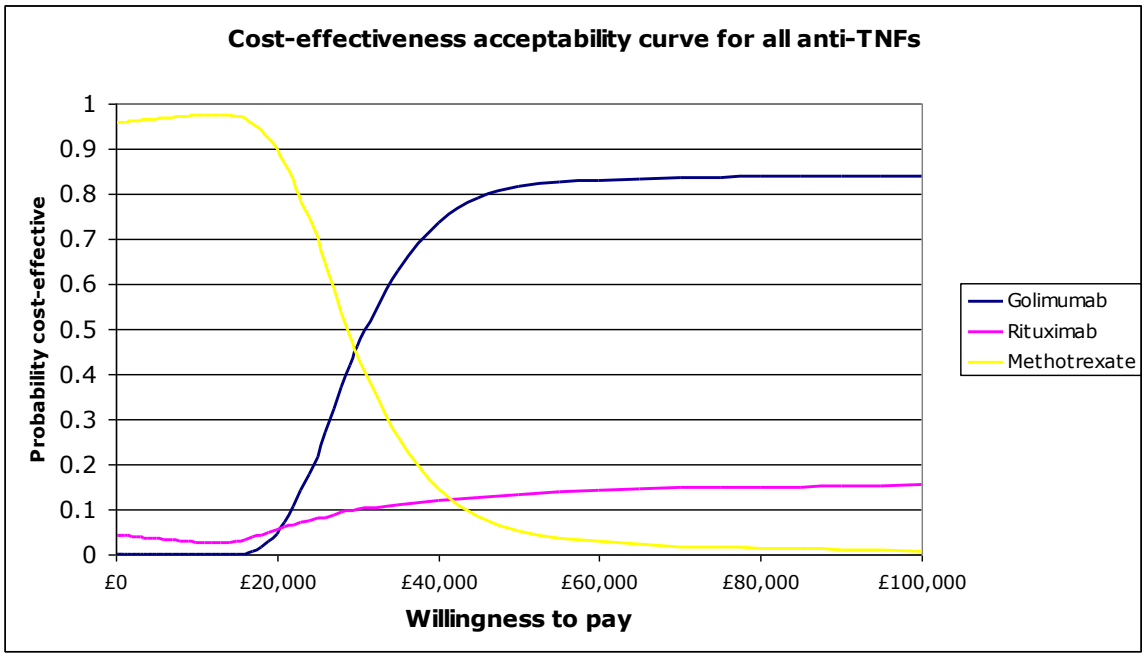


Figure 6. CEAC for all biologics (TNFα inhibitor experienced)



6.7.9 Please present the results of scenario analysis. Include details of structural sensitivity analysis.

Table 153. Results of the structural sensitivity analysis (rebound equal to natural history) – DMARD experienced

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) versus Baseline (Methotrexate)	Incremental analysis
Methotrexate	£35,869	4.489	-	-	-	-

Adalimumab	£66,875	5.276	£31,006	0.787	£39,402	£39,402
Golimumab	£67,747	5.297	£872	0.020	£39,491	£43,600
Infliximab	£69,899	5.174	£2,152	-0.122	£49,696	Dominated
Certolizumab	£73,571	5.675	£3,672	0.500	£31,812	£7,344
Etanercept	£74,208	5.554	£637	-0.121	£36,021	Dominated

Table 154. Results of the structural sensitivity analysis (rebound equal to natural history) – TNFα inhibitor experienced

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£) versus Baseline (Methotrexate)	Incremental analysis
Methotrexate	£33,673	3.129	-	-	-	-
Rituximab	£50,206	3.523	£16,534	0.297	£55,638	£55,638
Golimumab	£52,402	3.712	£2,196	0.178	£39,447	£12,337

6.7.10 What were the main findings of each of the sensitivity analyses?

The structural sensitivity analysis had a significant impact on the results. Changing the assumption from 'rebound equal to gain' to 'rebound equal to natural history' significantly increased the ICERs. In the previous appraisals of TNF-α inhibitors, the committee has acknowledged that the true rebound effect would lie somewhere between gain and natural history and in absence of any evidence have accepted 'rebound equal to gain' as the base case assumption.

One way sensitivity analyses identified the key variables affecting ICERs. Reducing the model time horizon had a significant impact with increased ICERs for shorter time horizons. Changing the HAQ progression for TNFα inhibitors substantially increased the ICERs. Previous NICE RA appraisals have accepted zero progression of HAQ whilst on biologic treatment. Recent clinical data with longer endpoints have suggested that HAQ may actually improve whilst on treatment, thus zero progression may be a conservative assumption. Changing the other parameters such as age, golimumab efficacy, SMR, baseline HAQ score, withdrawal rates and natural history HAQ progression had less significant impact on ICERs.

6.7.11 What are the key drivers of the cost-effectiveness results?

The key drivers for cost effectiveness analyses were ACR response rates and the magnitude of HAQ change for ACR responders. Among the TNF-α inhibitors, both certolizumab and etanercept showed numerically higher ACR response rates. However, it is important to note that none of the MTCs found a statistical difference between the TNFα inhibitors for either ACR20 or ACR50. Furthermore, the

withdrawal rates were strikingly greater in control arms (63-81%) than intervention arms (17-21%) within the certolizumab clinical trials so that any bias resulting from imputation would have a greater effect on control results. Etanercept clinical trials included patients with more active and longer duration of disease (section 5.3.4) compared with golimumab resulting in a larger MTC treatment effect which should therefore be viewed with caution within the cost effectiveness results.

6.8 Validation

6.8.1 Please describe the methods used to validate and quality assure the model. Provide references to the results produced and cross-reference to evidence identified in the clinical, quality of life and resources sections.

Validation of the model structure and assumptions was conducted with two clinicians in the UK. Dr Ostor, Consultant Rheumatologist, Addenbrooke's Hospital, Cambridge, and Dr Andrews, Consultant Rheumatologist, Southampton University Hospitals were interviewed.

Two health economics experts were also consulted at the early stages of the development of the model and to critique the draft model. They were Professor Martin Buxton, Brunel University and Professor Stephen Morris, University College London. A third health economics expert, Stephen Palmer was consulted on the structure of the model and suggested a number of amendments to the model.

Third party validation of the model was conducted at the Quality Control stage. An experienced programmer was asked to check the following aspects of the model:

1. Accuracy of input data. This was checked by comparing the model inputs in Excel against the data sources referenced
2. Top down tests. This involves systematic variation of the model input parameters to establish whether changes in inputs result in predictable changes in the model outputs. These tests are designed to identify failures in model logic or material computation errors.
3. Computation checks of key sensitivities. The following aspects of the spreadsheet were identified as key areas for detailed checking of formulae: translation of drug prices into state costs; derivation of transition rates from clinical inputs; derivation of state distributions from transition rates. Formulae performing these transformations were checked.
4. Submission. The accuracy of the reporting of data inputs and outputs in the modal was checked by reviewing the submission against the model.

The validation identified no major issues with the computational accuracy of the model.

6.9 *Subgroup analysis*

6.9.1 Please specify whether analysis of subgroups was undertaken and how these subgroups were identified. Were they identified on the basis of an a priori expectation of differential clinical or cost effectiveness due to known, biologically plausible, mechanisms, social characteristics or other clearly justified factors?

TNFα inhibitor failed patients

A pre-defined proportion of patients recruited for the GO-AFTER trial had received an inadequate response on a previous TNFα inhibitor. The remaining recruited patients had previously been exposed to at least one TNFα inhibitor. Both of these groups make up the base case analysis presented in the above economic sections. Whilst this combined group is more representative of the current RA patient population within the UK as treatment with a second biologic is not only prescribed to those patients that have had an inadequate response to a first TNFα inhibitor, the TNFα inhibitor failed patient group was explored in a subgroup analysis.

6.9.2 Please clearly define the characteristics of patients in the subgroup.

The patients within the assessed subgroup are comparable to the full population recruited within GO-AFTER with the exception of discontinuing a first TNFα inhibitor due to an inadequate response.

6.9.3 Please describe how the statistical analysis was undertaken.

Data was extracted for GO-AFTER (failed) and for rituximab's REFLEX trial to conduct an indirect comparison between golimumab and the biologic comparator.

Table 155. TNFα inhibitor experienced (failed) data

Drug	Reference	Study name	Sample size	Intervention, dosing	No of patients randomized to intervention	ACR20 6 months	ACR50 6 months
rituximab	Cohen SB *et al. Arth Rheum* 2006; 54:2793–806	REFLEX	499	rituximab i.v. 2 x 1000 mg days 1, 15 + MTX	298	152	80

				placebo i.v. days 1, 15 + MTX	201	36	10
[redacted]							
[redacted]							

A meta-analysis was not appropriate as only one trial was available for both of the biologics within a TNFα inhibitor failed population. RR was derived for both biologics in comparison to placebo for ACR20 and ACR50 response.

[redacted]

[redacted]							
[redacted]							
[redacted]							

[redacted]

[redacted]							
[redacted]							
[redacted]							

Following the known Bucher method an indirect comparison was conducted which found no statistical difference between golimumab and rituximab for ACR20 and ACR50 responders.

[redacted]

[redacted]							
[redacted]							
[redacted]							

6.9.4 What were the results of the subgroup analysis/analyses, if conducted? Please present results in a similar table as in section 6.7.6 (Base-case analysis).

Table 156. Results of the subgroup analysis (TNFα inhibitor failed)

Technologies	[redacted]				**ICER (£) versus Baseline (Methotrexate)**	[redacted]
Methotrexate	[redacted]				-	[redacted]
Golimumab	[redacted]				£23,914	[redacted]
Rituximab	[redacted]				£33,393	[redacted]

6.9.5 Were any obvious subgroups not considered? If so, which ones, and why were they not considered? Please refer to the subgroups identified in the decision problem in section 4.

None of the obvious subgroups were excluded.

6.10 Interpretation of economic evidence

6.10.1 Are the results from this economic evaluation consistent with the published economic literature? If not, why do the results from this evaluation differ, and why should the results in the submission be given more credence than those in the published literature?

Compared to other published studies in literature (Section 6.1.2), our DMARD experienced results indicate similar ICERs for TNFα inhibitors compared to palliation. Our model derives many assumptions from the BRAM model and thus the ICERs are in a similar range of those approved in recent NICE appraisals.

The results from this economic submission are consistent with the results presented in the assessment report for the recent NICE appraisal for the sequential use of biologics (NICE 2010). Due to a previous limitation on available RCT evidence, economic evaluations with a TNFα inhibitor experienced population are limited. However the ICERs within the TNFα inhibitor experienced population are within the range accepted by the Committee for tocilizumab and those submitted for the ongoing sequential use of biologics appraisal.

6.10.2 Is the economic evaluation relevant to all groups of patients who could potentially use the technology as identified in the decision problem in section 4?

The economic evaluation covers all the licensed patient populations for golimumab in RA. Contrary to existing TNFα inhibitors, the RCT evidence for golimumab allows for a thorough assessment of golimumab as both a 1st and 2nd line treatment option for RA patients.

6.10.3 What are the main strengths and weaknesses of the evaluation? How might these affect the interpretation of the results?

The primary strength of this analysis is that it is based on robust RCT evidence in both DMARD and TNFα inhibitor experienced patient populations. In the recent appraisal of the sequential use of biologics, the Committee concluded that the results from GO-AFTER "could be seen as confirming a beneficial effect of TNF inhibitor treatment following failure of a first TNF inhibitor" (FAD RA Sequential, 2010).
Another main strength of this evaluation is that the assessment is based on previous work within this area. Structural and parameter assumptions were informed by the recent tocilizumab, certolizumab and sequential use of biologics NICE appraisals.

The analyses have several limitations. A number of parameters such as QoL algorithm, long term withdrawal rates and resource use estimates were derived from literature and were based on non-randomised evidence. There was no evidence available for some of the structural assumptions such as rebound assumptions. In addition, some of the data were gathered based on expert opinion. This adds significant uncertainly to the findings but can only be attributed to the significant limitations in the available evidence.

6.10.4 What further analyses could be undertaken to enhance the robustness/completeness of the results?

Availability of long term data for biologics in the treatment of RA could enhance the robustness of the results by informing withdrawal rates and HAQ progression.

Section C – Implementation

7 Assessment of factors relevant to the NHS and other parties

7.1 How many patients are eligible for treatment in England and Wales? Present results for the full marketing authorisation/CE marking and for any subgroups considered. Also present results for the subsequent 5 years.

The total numbers of rheumatoid arthritis patients receiving biologic treatment were estimated as 48,755 patients in year 1 and increasing by 2.85% in year 5 to 50,186, as presented in Table 10.

Table 8 presents estimates for the number of existing rheumatoid arthritis patients as 47,869 in year 1 and increasing to 49,273 in year 5. Table 9 presents the estimates for newly diagnosed rheumatoid arthritis patients which ranges from 886 patients in year 1 and increasing to 912 in year 5.

Population estimates for England and Wales (≥ 15 years old) were extracted from the Government Actuary Department 2008 national population projections. Prevalence was extracted from Symmons *et al* 2002 as 0.81%. Incidence was derived from Garcia Rodriguez *et al* 2009 as 0.015%. Percentage of patients diagnosed (60%), treated (75%), biologic eligible (42%) and penetration (69%) were derived from data on file.

Table 157. Existing rheumatoid arthritis patients

PREVALENCE BASED	Percentage	2010	2011	2012	2013	2014
Eng/Wales Pop ≥ 15yrs		45,574,176	45,916,310	46,263,165	46,593,704	46,911,355
Prevalence	0.81%	369,151	371,922	374,732	377,409	379,982
Diagnosed	60%	221,490	223,153	224,839	226,445	227,989
Treated	75%	166,118	167,365	168,629	169,834	170,992
Biologic Eligible	42%	69,770	70,293	70,824	71,330	71,817
Penetration	69%	47,869	48,228	48,593	48,940	49,273

Table 158. Newly diagnosed rheumatoid arthritis patients

INCIDENCE BASED	Percentage	2010	2011	2012	2013	2014
Eng/Wales Pop ≥ 15yrs		45,574,176	45,916,310	46,263,165	46,593,704	46,911,355
Incidence	0.015%	6,836	6,887	6,939	6,989	7,037
Diagnosed	60%	4,102	4,132	4,164	4,193	4,222
Treated	75%	3,076	3,099	3,123	3,145	3,167
Biologic Eligible	42%	1,292	1,302	1,312	1,321	1,330
Penetration	69%	886	893	900	906	912

Table 159. Total number of RA patients in England and Wales estimated to receive biologic treatment

TOTAL		2010	2011	2012	2013	2014
Prevalence		47,869	48,228	48,593	48,940	49,273
Incidence		886	893	900	906	912
Total		**48,755**	**49,121**	**49,493**	**49,846**	**50,186**

7.2 What assumption(s) were made about current treatment options and uptake of technologies?

In line with the EMEA licensed approval received 1 October 2009, golimumab is assumed to be prescribed to DMARD experienced and to TNFα inhibitor experienced RA patients. Comparators within the budget impact model included those biologics which had both of the following:

- Market authorisation within England and Wales for treatment of RA, and
- Positive NICE guidance for use within DMARD experienced or TNFα experienced RA patients.

Based on these criteria, the following comparators were included within the budget impact model:

- Adalimumab
- Certolizumab pegol
- Etanercept
- Infliximab
- Rituximab

Golimumab is assumed to be a replacement therapy for other intravenous (IV) and subcutaneous biologics in a 25:75 split.

7.3 What assumption(s) were made about market share (when relevant)?

Based on MSD forecast estimates, **Error! Reference source not found.** and **Error! Reference source not found.** present the percentage of RA patients who will be treated on each biologic without and with the introduction of golimumab, respectively.

[redacted]

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

7.4 presents absolute patient numbers based on the total patient numbers in Table 10 and the market forecast percentages in

[redacted]

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

7.5 In addition to technology costs, please consider other significant costs associated with treatment that may be of interest to commissioners (for example, procedure codes and programme budget planning).

7.6

Drug dosing schedule

Based on an observed 6 month dosing frequency and in line with the Summary of Product Characteristics (SPC) for rituximab, Table 160 presents the average number of annual 1000 mg doses for rituximab as 4.4.

Table 160. Rituximab 1000 mg dose frequency

Year	Dose schedule (at week number)	Total number of Doses
1	0, 2, 24, 26	4
2	0, 2, 24, 26	4
3	0, 2, 24, 26	4
4	0, 2, 24, 26	4
5	0, 2, 24, 26, 50, 52	6
Average number of annual doses		**4.4**

Table 161 presents the average annual dose frequency for infliximab as 6.8 which was derived from infliximab's SPC.

Table 161. Infliximab 3 mg/kg dose frequency

Year	Dose schedule (at week number)	Total number of Doses
1	0, 2, 6, 14 , 22, 30, 38, 46	8
2	2, 10, 18, 26, 34, 42, 50	7
3	6, 14, 22, 30, 38, 46	6
4	2, 10, 18, 26, 34, 42, 50	7
5	6, 14, 22, 30, 38, 46	6
Average number of annual doses		**6.8**

Drug administration

For drugs administered by IV infusion (infliximab and rituximab), the cost per infusion was inflated from the 2004 Birmingham Rheumatoid Arthritis Model (BRAM) developed during Technology Appraisal 130. The pay and prices index was referenced from the Personal Social Services Research Unit 2009 Unit Costs of Health and Social Care. Table 162 presents the 2009 cost per infusion for infliximab and rituximab as £148.16.

Table 162. Cost per infusion

Year	Cost of Infusion	Pay and Prices Index
2004	£124.00	224.8
2005	£128.14	232.3
2006	£132.88	240.9
2007	£137.79	249.8
2008	£141.76	257.0
2009	**£148.16**	268.6

Weight based doses

Based on the British Society for Rheumatology Biologics Register (BSRBR) weight distribution presented in Table 163, the weighted average vials per infusion for infliximab was estimated to be 2.67 vials per 100 patients. This figure was derived

from the weight distributions of 3,208 patients registered within the BSRBR to conservatively determine the average number of *full* vials that were used per infusion. The weighted average vials per infusion was calculated with the following equation:

(SUMPRODUCT (% patients : vials per infusion)) / (SUM(% patients)).

Table 163. BSRBR weight-based distribution

	Weight (kg)			
Patient weight	<66 kg	66-100 kg	101-133kg	>134kg
% of patients	41%	52%	6%	1%
Vials per infusion	2	3	4	5
Weighted average vials per infusion:		**2.67**		

7.7 What unit costs were assumed? How were these calculated? If unit costs used in health economic modelling were not based on national reference costs or the PbR tariff, which HRGs reflected activity?

British National Formulary prices presented in Table 164 were referenced from BNF59. The monthly drug cost of golimumab is £774.58.

Table 164. Estimated drug unit costs

Comparator	Packaged dose	Recommended dose	Unit cost	Source
Golimumab	50mg	50mg once monthly	£774.58	2010 Data on file.
Adalimumab	40mg	40mg every other week	£357.50	2010 BNF
Certolizumab	200mg	400mg every 4 weeks	£338.25**	2010 BNF
Etanercept	25mg	50mg weekly	£89.38	2010 BNF
Infliximab	100mg	3mg/kg every 8 weeks	£419.62	2010 BNF
Rituximab	500mg	2 1000mg infusions every 6-12 months	£873.15	2010 BNF

*** Based on the BNF59 unit cost of £357.50 for certolizumab pegol and the patient access scheme presented in NICE TA186 (10 syringes free of charge at start of treatment), the adjusted unit cost was estimated as £338.25 as presented in Table 165.*

Table 165. Certolizumab pegol costs

Patient Access Scheme Adjustment	Certolizumab without PAS	Certolizumab with PAS
Year 1 annual cost	£10,367.50	£6,792.50
Year 2 onwards annual cost	£9,295.00	£9,295.00
Average annual cost over 5 year time horizon	£9,509.50	£8,794.50
Unit cost	£365.75	**£338.25**

7.8 Were there any estimates of resource savings? If so, what were they?

The budget impact estimates do not incorporate any assumptions for resource savings.

7.9 What is the estimated annual budget impact for the NHS in England and Wales?

A budget impact model has been used to determine the budget impact upon introducing golimumab as a treatment for the DMARD experienced and the TNFα inhibitor experienced RA population. Table 166 presents the estimates for the 5-year budget implications ranging from no proportional difference in year 1 to 0.27% in year 5.

Table 166. Estimated budget impact for the 5-year treatment and administration costs

Year	Budget Impact (Absolute Difference)	Budget Impact (Proportional Difference)
2010	£0	0.00%
2011	£101,482	0.02%
2012	£278,492	0.06%
2013	£686,267	0.15%
2014	£1,308,426	0.29%

7.10 Are there any other opportunities for resource savings or redirection of resources that it has not been possible to quantify?

Patient-focused aspects of golimumab have been identified as a potential resource savings as they may assist patients in achieving greater compliance, which in turn will improve treatment outcomes, reduce drug wastage, and reduce payer costs.

[redacted]

Convenient, monthly dosing

Golimumab is self-administered by patients via a once-monthly subcutaneous injection. The auto injector has been specifically developed in response to patient needs; its features include an ergonomically designed barrel for easy handling, a large side button for ease of activation that does not require thumb strength, a safety sleeve to avoid accidental firing, a large observation window, audible clicks for initiation and completion of golimumab administration, and a needle which auto-injects and auto-retracts whilst remaining out of sight of patients.

Reduction in injection site reactions

Golimumab differs in its molecular make-up and compound formulation compared to other TNFα inhibitors. Jorgensen *et al* 1996 and Kappelgaard *et al* 2004 found increasing injection volume from 0.5 ml to 1.0 ml and using citrate-acid buffered solutions significantly increased pain levels. Golimumab's buffered solution (without citric acid monohydrate & low injection volume 0.5 ml) correlates with a lower incidence of injection site reactions Keystone *et al* 2008, Smolen *et al* 2008.

Patient Support Programme

MSD will provide a golimumab patient support programme, designed to encourage patients to stay on their treatment as directed, and remind them when their next monthly treatment is due. It will also assist in managing treatment expectations, and provide simple and relevant information and timely practical help so that patients feel comfortable with self-injection.

8 References

Abe T, Takeuchi T, Miyasaka N et al. A multicenter, double-blind, randomized, placebo controlled trial of infliximab combined with low dose methotrexate in Japanese patients with rheumatoid arthritis. J Rheumatol 2006;33:37-44.

Ariza-Ariza R, Hernandez-Cruz B, Carmona L, Dolores Ruiz-Montesinos M, Ballina J, Navarro-Sarabia F. Assessing utility values in rheumatoid arthritis: a comparison between time trade-off and the EuroQol. Arthritis Rheum 2006 Oct 15;55(5):751-6.

Arnett FC, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988; 31: 315-24.

Bansback NJ, Brennan A, Ghatnekar O. Cost effectiveness of adalimumab in the treatment of patients with moderate to severe rheumatoid arthritis in Sweden. Ann Rheum Dis 2005 Jul;64(7):995-1002.

Bansback N, Marra C, Tsuchiya A, Anis A, Guh D, Hammond T, et al. Using the health assessment questionnaire to estimate preference-based single indices in patients with rheumatoid arthritis. Arthritis Rheum 2007 Aug 15;57(6):963-71.

Bansback NJ, Brennan A, Ghatnekar O. Cost effectiveness of adalimumab in the treatment of patients with moderate to severe rheumatoid arthritis in Sweden. Ann Rheum Dis 2005 Jul;64(7):995-1002.

Barrett EM, Scott DG, Wiles NJ, Symmons DP. The impact of rheumatoid arthritis on employment status in the early years of disease: a UK community-based study. Rheumatology (Oxford) 2000;39:1403-9.

Barbieri M, Wong JB, Drummond M. The cost effectiveness of infliximab for severe treatment-resistant rheumatoid arthritis in the UK. Pharmacoeconomics 2005;23(6):607-18.

Barton P, Jobanputra P, Wilson J, Bryan S, Burls A. The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis. Health Technol Assess 2004 Mar;8(11):iii, 1-iii,91.

Bennett AN, Peterson P, Zain A, Grumley J, Panayi G, Kirkham B. Adalimumab in clinical practice. Outcome in 70 rheumatoid arthritis patients, including comparison of patients with and without previous anti-TNF exposure. Rheumatology (Oxford) 2005; 44: 1026-31.

Blumenauer B, Judd M, Wells G et al. Infliximab for the treatment of rheumatoid arthritis. Cochrane Database Syst Rev 2002:Issue 3. Art. No.: CD003785. DOI: 10.1002/14651858.CD003785.

Blumenauer B, Judd M, Cranney A et al. Etanercept for the treatment of rheumatoid arthritis. Cochrane Database Syst Rev 2003:Issue 4. Art. No.: CD004525. DOI: 10.1002/14651858.CD004525.

Bongartz T, Sutton AJ, Sweeting MJ. Anti-TNF anti-TNF antibody therapy in RA and the risk of serious infections and malignancies: Systematic review and meta-analysis of rare harmful effects in randomized controlled trials. JAMA 2006:19:2275-85.

Bongartz T, Warren FC, Mines D et al. Etanercept therapy in rheumatoid arthritis and the risk of malignancies: a systematic review and individual patient data meta-analysis of randomised controlled trials. Ann Rheum Dis 2009;68:1177-1183.

Bucher, H.C, G.H. Guyatt, L.E. Griffith and S.D. Walter, The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials, J Clin Epidemiol 50 (1997), pp. 683–691

Brazier J, Roberts J, Deverill M. The estimation of a
preference-based measure of health from the SF-36. J Health Econ 2002 Mar;21(2):271-92.

Brennan A, Bansback N, Nixon R, Madan J, Harrison M, Watson K, et al. Modelling the cost effectiveness of TNF-alpha antagonists in the management of rheumatoid arthritis: results from the British Society for Rheumatology Biologics Registry. Rheumatology (Oxford) 2007 Aug;46(8):1345-54.

Brennan A, Bansback N, Reynolds A, Conway P. Modelling the cost-effectiveness of etanercept in adults with rheumatoid arthritis in the UK. Rheumatology (Oxford) 2004 Jan;43(1):62-72.

British National Formulary. BNF 59 . 2010.

British National Formulary. BNF 56 . 2008.

Centocor, Inc. C0524T06. Accessed on: 11 May 2010. Available at http://clinicaltrials.gov/ct2/results?term=golimumab

Centocor, Inc. C0524T11. Accessed on: 11 May 2010. Available at:
http://clinicaltrials.gov/ct2/results?term=golimumab

Centocor Inc. C0524T02 Clinical Study Report. A Randomized, Double-blind, Dose-ranging Trial of CNTO 148 Subcutaneous Injection Compared With Placebo in Subjects With Active Rheumatoid Arthritis Despite Treatment With Methotrexate (Phase II Trial).

Chen DY, Chou SJ, Hsieh TY et al. Randomized, double-blind, placebo-controlled, comparative study of human anti-TNF antibody adalimumab in combination with methotrexate and methotrexate alone in Taiwanese patients with active rheumatoid arthritis. J Formos Med Assoc 2009;108:310-9.

Chen YF, Jobanputra P, Barton P, Jowett S, Bryan S, Clark W, et al. A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness. Health Technol Assess 2006 Nov;10(42):iii-xiii, 1.

Chehata JC, Hassell AB, Clarke SA, Mattey DL, Jones MA, Jones PW, et al. Mortality in rheumatoid arthritis: relationship to single and composite measures of disease activity. Rheumatology (Oxford) 2001 Apr;40(4):447-52.

Choi HK, Seeger JD, Kuntz KM. A cost-effectiveness analysis of treatment options for patients with methotrexate-resistant rheumatoid arthritis. Arthritis Rheum 2000 Oct;43(10):2316-27.
Cohen G, Courvoisier N, Cohen JD, Zaltni S, Sany J, Combe B. The efficiency of switching from infliximab to etanercept and vice-versa in patients with rheumatoid arthritis. Clin Exp Rheumatol 2005; 23: 795-800.
Cohen SB, Emery P, Greenwald MW et al. Rituximab for rheumatoid arthritis refractory to anti-tumor necrosis factor therapy: Results of a multicenter, randomized, double-blind, placebo-controlled, phase III trial evaluating primary efficacy and safety at twenty-four weeks. Arthritis Rheum 2006;54:2793-806.
Combe B, Codreanu C, Fiocco U et al. Etanercept and sulfasalazine, alone and combined, in patients with active rheumatoid arthritis despite receiving sulfasalazine: a double-blind comparison. Ann Rheum Dis 2006;65:1357-62.
Deighton C, Hyrich K, Ding T, Ledingham J, Lunt M, Luqmani R, et al. British Society for Rheumatology and British Health Professionals in Rheumatology guidelines on eligibility criteria for the first biological therapyfor the management of rheumatoid arthritis (after the first 2 years). Rheumatology (Oxford) 2010 Mar.
Edwards JC, Szczepanski L, Szechinski J et al. Efficacy of B-cell-targeted therapy with rituximab in patients with rheumatoid arthritis. N Engl J Med 2004;350:2572-81.
Edwards CJ, Arden NK, Fisher D, Saperia JC, Reading I, Van Staa TP, et al. The changing use of disease-modifying anti-rheumatic drugs in individuals with rheumatoid arthritis from the United Kingdom General Practice Research Database. Rheumatology (Oxford) 2005 Nov;44(11):1394-8.
Emery P, Kosinski M, Li T et al. Treatment of rheumatoid arthritis patients with abatacept and methotrexate significantly improved health-related quality of life. J Rheumatol 2006;33:681-9.
Emery P, Keystone E, Tony HP et al. IL-6 receptor inhibition with tocilizumab improves treatment outcomes in patients with rheumatoid arthritis refractory to anti-tumour necrosis factor biologicals: results from a 24-week multicentre randomised placebo-controlled trial. Ann Rheum Dis 2008;67:1516-23.
Emery P, Fleischmann RM, Moreland LW et al. Golimumab, a human anti-tumor necrosis factor alpha monoclonal antibody, injected subcutaneously every four weeks in methotrexate-naive patients with active rheumatoid arthritis: twenty-four-week results of a phase III, multicenter, randomized, double-blind, placebo-controlled study of golimumab before methotrexate as first-line therapy for early-onset rheumatoid arthritis. Arthritis Rheum 2009;60:2272-83.
Emery P, Fleischmann R, Filipowicz-Sosnowska A, Schechtman J, Szczepanski L, Kavanaugh A, et al. The efficacy and safety of rituximab in patients with active rheumatoid arthritis despite methotrexate treatment: Results of a phase IIb randomized, double-blind, placebo-controlled, dose-ranging trial. Arthritis Rheum 2006;54(5):1390-400.
European Medicines Agency. Assessment report for Simponi. International nonproprietary name: golimumab. London, UK: European Medicines Agency, 2009. Available from: www.ema.europa.eu/humandocs/PDFs/EPAR/simponi/H-992-en6.pdf (Accessed 16 Mar 2010).
Fleischmann R, Vencovsky J, van Vollenhoven R et al. Efficacy and safety of certolizumab pegol monotherapy every 4 weeks in patients with rheumatoid arthritis failing previous disease-modifying antirheumatic therapy: the FAST4WARD study. Ann Rheum Dis 2009;68:805-11.
Furst DE, Gaylis N, Bray V, et al. Open-label, pilot protocol of patients with rheumatoid arthritis who switch to infliximab after an incomplete response to etanercept: the opposite study. Ann Rheum Dis 2007; 66: 893-99.
Furst DE, Schiff MH, Fleischmann RM et al. Adalimumab, a fully human anti tumor necrosis factor-alpha monoclonal antibody, and concomitant standard antirheumatic therapy for the treatment of rheumatoid arthritis: results of STAR (Safety Trial of Adalimumab in Rheumatoid Arthritis). J Rheumatol 2003;30:2563-71.
Fransen J, Creemers MCW, van Riel PLCM. Remission in rheumatoid arthritis: agreement of the disease activity score (DAS28) with the ARA preliminary remission criteria. Rheumatology 2004: 43(10):1252-5.
Fries JF, Sptiz P, Kraines RG, Holman HR. Measurement of patient outcomes in arthritis. Arthritis Rheum. 1980;23(2):137-145.
Garcia Rodriguez et al. Rheumatoid arthritis in UK primary care: incidence and prior morbidity. Scand J Rheumat 2009;38:173-177.
Geborek P, Crnkic M, Petersson IF, Saxne T. Etanercept, infliximab, and leflunomide in established rheumatoid arthritis: clinical experience using a structured follow up programme in southern Sweden. Ann Rheum Dis 2002 Sep;61(9):793-8.
Genovese MC, Becker JC, Schiff M et al. Abatacept for rheumatoid arthritis refractory to tumor necrosis factor alpha inhibition. N Engl J Med 2005;353:1114-23.
Genovese MC, McKay JD, Nasonov EL et al. Interleukin-6 receptor inhibition with tocilizumab reduces disease activity in rheumatoid arthritis with inadequate response to disease-modifying antirheumatic drugs: the tocilizumab in combination with traditional disease-modifying antirheumatic drug therapy study. Arthritis Rheum 2008;58:2968-80.
Gartlehner G, Hansen RA, Jonas BL, Thieda P, Lohr KN. The comparative efficacy and safety of biologics for the treatment of rheumatoid arthritis: a systematic review and metaanalysis. J Rheum 2006;33(12)2398-2408.
Goodson N, Marks J, Lunt M, Symmons D. Cardiovascular admissions and mortality in an inception cohort of patients with rheumatoid arthritis with onset in the 1980s and 1990s. Ann Rheum Dis 2005;64:1595-601.

Government Actuary's Department. Projections database. Accessed on 12 April 2010. Available at: www.gad.gov.uk/demographyData/Population/index.aspx
Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). Br J Rheumatol 1997 May;36(5):551-9.
Hyrich KL, Lunt M, Watson KD, Symmons DP, Silman AJ. Outcomes after switching from one anti-tumor necrosis factor alpha agent to a second anti-tumor necrosis factor alpha agent in patients with rheumatoid arthritis: results from a large UK national cohort study. Arthritis Rheum 2007; 56: 13-20.
IMS Health. Systematic literature review and indirect comparison analysis of golimumab against adalimumab, etanercept, infliximab, and rituximab in the treatment of rheumatoid arthritis. 2008.
Interim Life Tables: United Kingdom. Office of National Statistics 2006 Available from http://www.statistics.gov.uk/downloads/theme_population/Interim_Life/ILTUK0406Reg.xls#Notation!A4
Karlsson JA, Kristensen LE, Kapetanovic MC, Gulfe A, Saxne T, Geborek P. Treatment response to a second or third TNF-inhibitor in RA: results from the South Swedish Arthritis Treatment Group Register. Rheumatology (Oxford) 2008 Apr;47(4):507-13.
Kay J, Matteson EL, Dasgupta B et al. Golimumab in patients with active rheumatoid arthritis despite treatment with methotrexate: a randomized, double-blind, placebo-controlled, dose-ranging study. Arthritis Rheum 2008;58:964-75.
Kennedy T, McCabe C, Struthers G, Sinclair H, Chakravaty K, Bax D, et al. BSR guidelines on standards of care for persons with rheumatoid arthritis. Rheumatology (Oxford) 2005 Apr;44(4):553-6.
Keystone E, Burmester GR, Furie R et al. Improvement in patient-reported outcomes in a rituximab trial in patients with severe rheumatoid arthritis refractory to anti-tumor necrosis factor therapy. Arthritis Rheum 2008a;59:785-93.
Keystone E, Heijde D, Mason D, Jr. et al. Certolizumab pegol plus methotrexate is significantly more effective than placebo plus methotrexate in active rheumatoid arthritis: findings of a fifty-two-week, phase III, multicenter, randomized, double-blind, placebo-controlled, parallel-group study. Arthritis Rheum 2008b;58:3319-29.
Keystone E, Fleischmann R, Emery P, Furst DE, van VR, Bathon J, et al. Safety and efficacy of additional courses of rituximab in patients with active rheumatoid arthritis: an open-label extension analysis. Arthritis Rheum 2007 Dec;56(12):3896-908.
Keystone E, Freundlich B, Schiff M, Li J, Hooper M. Patients with moderate rheumatoid arthritis (RA) achieve better disease activity states with etanercept treatment than patients with severe RA. J Rheumatol 2009a;36:522-31.
Keystone EC, Kavanaugh AF, Sharp JT et al. Radiographic, clinical, and functional outcomes of treatment with adalimumab (a human anti-tumor necrosis factor monoclonal antibody) in patients with active rheumatoid arthritis receiving concomitant methotrexate therapy: a randomized, placebo-controlled, 52-week trial. Arthritis Rheum 2004;50:1400-11.
Keystone EC, Genovese MC, Klareskog L et al. Golimumab, a human antibody to tumour necrosis factor {alpha} given by monthly subcutaneous injections, in active rheumatoid arthritis despite methotrexate therapy: the GO-FORWARD Study. Ann Rheum Dis 2009b;68:789-96.
Kielhorn A, Porter D, Diamantopoulos A, Lewis G. UK cost-utility analysis of rituximab in patients with rheumatoid arthritis that failed to respond adequately to a biologic disease-modifying antirheumatic drug. Curr Med Res Opin 2008 Sep;24(9):2639-50.
Kim H-Y, Lee S-K, Song Y-W et al. A randomized, double-blind, placebo-controlled, phase III study of the human anti-tumor necrosis factor antibody adalimumab administered as subcutaneous injections in Korean rheumatoid arthritis patients treated with methotrexate. International Journal of Rheumatic Disease 2007;10:9-16.
Klareskog L, van der Heijde D, de Jager JP et al. Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: double-blind randomised controlled trial. Lancet 2004;363:675-81.
Kobelt G, Jonsson L, Lindgren P, Young A, Eberhardt K. Modeling the progression of rheumatoid arthritis: a two-country model to estimate costs and consequences of rheumatoid arthritis. Arthritis Rheum 2002 Sep;46(9):2310-9.
Kobelt G, Lindgren P, Lindroth Y, Jacobson L, Eberhardt K. Modelling the effect of function and disease activity on costs and quality of life in rheumatoid arthritis. Rheumatology (Oxford) 2005 Sep;44(9):1169-75.
Kobelt G, Lindgren P, Singh A, Klareskog L. Cost effectiveness of etanercept (Enbrel) in combination with methotrexate in the treatment of active rheumatoid arthritis based on the TEMPO trial. Ann Rheum Dis 2005 Aug;64(8):1174-9.
Kremer JM, Genovese MC, Cannon GW, Caldwell JR, Cush JJ, Furst DE, et al. Concomitant leflunomide therapy in patients with active rheumatoid arthritis despite stable doses of methotrexate. A randomized, double-blind, placebo-controlled trial. Ann Intern Med 2002 Nov 5;137(9):726-33.
Kremer JM, Westhovens R, Leon M et al. Treatment of rheumatoid arthritis by selective inhibition of T-cell activation with fusion protein CTLA4Ig. N Engl J Med 2003;349:1907-15.

Kremer JM, Dougados M, Emery P et al. Treatment of rheumatoid arthritis with the selective costimulation modulator abatacept: twelve-month results of a phase iib, double-blind, randomized, placebo-controlled trial. Arthritis Rheum 2005;52:2263-71.
Kremer JM, Genant HK, Moreland LW et al. Effects of abatacept in patients with methotrexate-resistant active rheumatoid arthritis: a randomized trial. Ann Intern Med 2006;144:865-76.
Kristensen LE, Saxne T, Geborek P. The LUNDEX, a new index of drug efficacy in clinical practice: results of a five-year observational study of treatment with infliximab and etanercept among rheumatoid arthritis patients in southern Sweden. Arthritis Rheum 2006 Feb;54(2):600-6.
Kwoh CK, Anderson LG, Greene JM et al. Guidelines for the management of rheumatoid arthritis. Arthritis Rheum. 2002;46(2):328-346.
Kvien TK. Epidemiology and burden of illness of rheumatoid arthritis. Pharmacoeconomics. 2004;22(Suppl 1):1-12.
Lipsky PE, van der Heijde DM, St Clair EW et al. Infliximab and methotrexate in the treatment of rheumatoid arthritis. Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy Study Group. N Engl J Med 2000;343:1594-602.
Luqmani R, Hennell S, Estrach C et al. British Society for Rheumatology and British Health Professionals in Rheumatology guideline for the management of rheumatoid arthritis (after the first 2 years). Rheumatology (Oxford) 2009;48:436-9.
Luqmani R, Hennell S, Estrach C, Birrell F, Bosworth A, Davenport G, et al. British Society for Rheumatology and british health professionals in Rheumatology guideline for the management of rheumatoid arthritis (the first two years). Rheumatology (Oxford) 2006 Sep;45(9):1167-9.
Lundkvist J, Kastang F, Kobelt G. The burden of rheumatoid arthritis and access to treatment: health burden and costs. Eur J Heart Econ. 2008;8(Suppl 2):S49-S60.
Maini R, St Clair EW, Breedveld F et al. Infliximab (chimeric anti-tumour necrosis factor alpha monoclonal antibody) versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: a randomised phase III trial. ATTRACT Study Group. Lancet 1999;354:1932-9.
Maini RN, Breedveld FC, Kalden JR et al. Therapeutic efficacy of multiple intravenous infusions of anti-tumor necrosis factor alpha monoclonal antibody combined with low-dose weekly methotrexate in rheumatoid arthritis. Arthritis Rheum 1998;41:1552-63.
Maxwell L, Singh JA. Abatacept for rheumatoid arthritis. Cochrane Database Syst Rev 2009:Issue 4. Art. No.: CD007277. DOI: 10.1002/14651858.CD007277.pub2.
Meads C, Jit M, Tsourapas A et al. Tocilizumab for the treatment of rheumatoid arthritis. A single technology appraisal. Birmingham, UK: West Midlands, Health Technology Assessment Collaboration., 2009. Available from: www.nice.org.uk/guidance/index.jsp?action=download&o=45575 (Accessed 9 Mar 2010).
Meune C, Touze E, Trinquart L, Allanore Y. Trends in cardiovascular mortality in patients with rheumatoid arthritis over 50 years: a systematic review and meta-analysis of cohort studies. Rheumatology (Oxford) 2009;48:1309-13.
Miyasaka N. Clinical investigation in highly disease-affected rheumatoid arthritis patients in Japan with adalimumab applying standard and general evaluation: the CHANGE study. Mod Rheumatol 2008;18:252-62.
Moreland LW, Schiff MH, Baumgartner SW et al. Etanercept therapy in rheumatoid arthritis. A randomized, controlled trial. Ann Intern Med 1999;130:478-86.
MSD. Market research data on file. 2010.
MSD. Simponi® 50mg solution for injection. Summary of Product Characteristics. 2010.
National Institute for Health and Clinical Excellence. Adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis (TA130) London, UK: National Institute for Health and Clinical Excellence, 2007. Available from: www.nice.org.uk/TA130 (Accessed 9 Mar 2010).
National Institute for Health and Clinical Excellence. Rheumatoid Arthritis - National clinical guideline for management and treatment in adults (CG79) London, UK: Royal College of Physicians of London, 2009a. Available from: www.nice.org.uk/nicemedia/pdf/CG79FullGuideline.pdf (Accessed 9 Mar 2010).
National Institute for Health and Clinical Excellence. Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a TNF inhibitor (part review of TA36, review of TA126 and 141): final scope London, UK: National Institute for Health and Clinical Excellence, 2009b. Available from: http://guidance.nice.org.uk/TA/WaveR/61 (Accessed 8 Apr 2010).
National Institute for Health and Clinical Excellence. Certolizumab pegol for the treatment of rheumatoid arthritis (TA186) London, UK: National Institute for Health and Clinical Excellence, 2010a. Available from: www.nice.org.uk/guidance/TA186 (Accessed 9 Mar 2010).
National Institute for Health and Clinical Excellence. Tocilizumab for the treatment of rheumatoid arthritis London, UK: National Institute for Health and Clinical Excellence, 2010b. Available from: http://guidance.nice.org.uk/TA/Wave18/5 (Accessed 8 Apr 2010).
NHS PbR tariff 2008-2009. Department of Health 2007 Available from http://www.dh.gov.uk/en/Managingyourorganisation/Financeandplanning/NHSFinancialReforms/DH_077279

NHS reference costs 2006-07. Department of Health 2008 Available from http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_082571

Navarro-Sarabia F, Ariza-Ariza R, Hernandez-Cruz B, Villanueva I. Adalimumab for treating rheumatoid arthritis. Cochrane Database Syst Rev 2005:Issue 3. Art. No.: CD005113. DOI: 10.1002/14651858.CD005113.pub2.

Nikas SN, Voulgari PV, Alamanos Y, et al. Effi cacy and safety of switching from infl iximab to adalimumab: a comparative controlled study. Ann Rheum Dis 2006; 65: 257-60.

Nishimoto N, Miyasaka N, Yamamoto K et al. Study of active controlled tocilizumab monotherapy for rheumatoid arthritis patients with an inadequate response to methotrexate (SATORI): significant reduction in disease activity and serum vascular endothelial growth factor by IL-6 receptor inhibition therapy. Mod Rheumatol 2009;19:12-9.

Personal Social Services Research Unit. Unit costs of health and social care. 2008.

Pincus T, Yazici Y, Sokka T, Aletaha D, Smolen JS. Methotrexate as the "anchor drug" for the treatment of early rheumatoid arthritis. Clin Exp Rheumatol 2003;21:S179-85.

Pugner KM, Scott DI, Holmes JW, Hieke K. The costs of rheumatoid arthritis: an international long-term view. Semin Arthritis Rheum 2000;29:305-20.

Radovits BJ, Fransen J, Shamma SA, Eijsbouts AM, Riel PLCMv, Laan RFJM. Excess mortality emerges after 10 years in an inception cohort of early rheumatoid arthritis. Arthritis Care & Research 2010;62:362-370.

Roche. RoActemra (tocilizumab) NICE STA Submission. Achieving clinical excellence in the treatment of moderate to severe rheumatoid arthritis. Roche Submission to the National Institute for Health and Clinical Excellence.: Roche, 2009. Available from: www.nice.org.uk/guidance/index.jsp?action=folder&o=45559 (Accessed 9 Mar 2010).

Rosenberg AE. Bones, Joints, and Soft Tissue Tumors. In: Kumar V, Fausto N, Abbas A, editors. Robbins and Cotran Pathologic Basis of Disease. 7th ed. Philadelphia, PA: Elsevier Saunders; 2005.

Russell AS, Wallenstein GV, Li T et al. Abatacept improves both the physical and mental health of patients with rheumatoid arthritis who have inadequate response to methotrexate treatment. Ann Rheum Dis 2007;66:189-94.

Saag KG, Teng GG, Patkar NM, Anuntiyo J, Finney C, Curtis JR, et al. American College of Rheumatology 2008 recommendations for the use of nonbiologic and biologic disease-modifying antirheumatic drugs in rheumatoid arthritis. Arthritis Rheum 2008 Jun 15;59(6):762-84.

Schiff M, Keiserman M, Codding C et al. Efficacy and safety of abatacept or infliximab vs placebo in ATTEST: a phase III, multi-centre, randomised, double-blind, placebo-controlled study in patients with rheumatoid arthritis and an inadequate response to methotrexate. Ann Rheum Dis 2008;67:1096-103.

Scott DL, Strand V. The effects of disease-modifying anti-rheumatic drugs on the Health Assessment Questionnaire score. Lessons from the leflunomide clinical trials database. Rheumatology (Oxford) 2002 Aug;41(8):899-909.

Scott DL, Smolen JS, Kalden JR, van de Putte LB, Larsen A, Kvien TK, et al. Treatment of active rheumatoid arthritis with leflunomide: two year follow up of a double blind, placebo controlled trial versus sulfasalazine. Ann Rheum Dis 2001 Oct;60(10):913-23.

Smolen JS, Beaulieu A, Rubbert-Roth A et al. Effect of interleukin-6 receptor inhibition with tocilizumab in patients with rheumatoid arthritis (OPTION study): a double-blind, placebo-controlled, randomised trial. Lancet 2008;371:987-97.

Smolen JS, Kay J, Doyle MK et al. Golimumab in patients with active rheumatoid arthritis after treatment with tumour necrosis factor alpha inhibitors (GO-AFTER study): a multicentre, randomised, double-blind, placebo-controlled, phase III trial. Lancet 2009a;374:210-21.

Smolen JS, Landewe RB, Mease PJ et al. Efficacy and safety of certolizumab pegol plus methotrexate in active rheumatoid arthritis: the RAPID 2 study. Ann Rheum Dis 2009b;68:797-804.

Smolen JS, Landewe R, Breedveld C, Dougados M, Emery P, Gaujoux-Viala C, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs. Ann Rheum Dis. [online]. Accessed 11 May 2010. Available at: www.eular.org

Strand V, Cohen S, Schiff M, Weaver A, Fleischmann R, Cannon G, et al. Treatment of active rheumatoid arthritis with leflunomide compared with placebo and methotrexate. Leflunomide Rheumatoid Arthritis Investigators Group. Arch Intern Med 1999 Nov 22;159(21):2542-50.

Strand V, Balbir-Gurman A, Pavelka K et al. Sustained benefit in rheumatoid arthritis following one course of rituximab: improvements in physical function over 2 years. Rheumatology (Oxford) 2006;45:1505-13.

Strand V, Mease P, Burmester GR et al. Rapid and sustained improvements in health-related quality of life, fatigue, and other patient-reported outcomes in rheumatoid arthritis patients treated with certolizumab pegol plus methotrexate over 1 year: results from the RAPID 1 randomized controlled trial. Arthritis Res Ther 2009;11:R170.

Symmons DP, Barrett EM, Bankhead CR, Scott DG, Silman AJ. The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. Br J Rheumatol 1994;33:735-9.

Symmons D, Mathers C, and Pfleger B. The global burden of rheumatoid arthritis in the year 2000. World Health Organisation. 2006.

Symmons D, Turner G, Webb R et al. The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. Rheumatology (Oxford) 2002;41:793-800.
UCB Pharma. Certolizumab pegol (CIMZIA®) for the treatment of rheumatoid arthritis. Single technology appraisal (STA) manufacturer submission to nice.: UCB Pharma, 2009. Available from: www.nice.org.uk/guidance/index.jsp?action=folder&o=45791 (Accessed 9 Mar 2010).
Van der Linden MA. Disease Activity Scores using C-reactive protein. April 5, 2004. Available at: www.umcn.nl/scientist/afdelingen/reumatologie/the_das/the_das_crp. Accessed 19 June 2009.
van de Putte LB, Atkins C, Malaise M et al. Efficacy and safety of adalimumab as monotherapy in patients with rheumatoid arthritis for whom previous disease modifying antirheumatic drug treatment has failed. Ann Rheum Dis 2004;63:508-16.
van der Heijde D, Klareskog L, Rodriguez-Valverde V et al. Comparison of etanercept and methotrexate, alone and combined, in the treatment of rheumatoid arthritis: two-year clinical and radiographic results from the TEMPO study, a double-blind, randomized trial. Arthritis Rheum 2006a;54:1063-74.
van der Heijde D, Klareskog L, Singh A et al. Patient reported outcomes in a trial of combination therapy with etanercept and methotrexate for rheumatoid arthritis: the TEMPO trial. Ann Rheum Dis 2006b;65:328-34.
Vera-Llonch M, Massarotti E, Wolfe F, Shadick N, Westhovens R, Sofrygin O, et al. Cost-effectiveness of abatacept in patients with moderately to severely active rheumatoid arthritis and inadequate response to methotrexate. Rheumatology (Oxford) 2008 Apr;47(4):535-41.
Weinblatt M, Combe B, Covucci A, Aranda R, Becker JC, Keystone E. Safety of the selective costimulation modulator abatacept in rheumatoid arthritis patients receiving background biologic and nonbiologic disease-modifying antirheumatic drugs: A one-year randomized, placebo-controlled study. Arthritis Rheum 2006;54:2807-16.
Weinblatt ME, Kremer JM, Bankhurst AD et al. A trial of etanercept, a recombinant tumor necrosis factor receptor:Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate. N Engl J Med 1999;340:253-9.
Weinblatt ME, Keystone EC, Furst DE et al. Adalimumab, a fully human anti-tumor necrosis factor alpha monoclonal antibody, for the treatment of rheumatoid arthritis in patients taking concomitant methotrexate: the ARMADA trial. Arthritis Rheum 2003;48:35-45.
Welsing PM, Severens JL, Hartman M, van Riel PL, Laan RF. Modeling the 5-year cost effectiveness of treatment strategies including tumor necrosis factor-blocking agents and leflunomide for treating rheumatoid arthritis in the Netherlands. Arthritis Rheum 2004 Dec 15;51(6):964-73.
Welsing PM, Severens JL, Hartman M, van Gestel AM, van Riel PL, Laan RF. The initial validation of a Markov model for the economic evaluation of (new) treatments for rheumatoid arthritis. Pharmacoeconomics 2006;24(10):1011-20.
Westhovens R, Cole JC, Li T et al. Improved health-related quality of life for rheumatoid arthritis patients treated with abatacept who have inadequate response to anti-TNF therapy in a double-blind, placebo-controlled, multicentre randomized clinical trial. Rheumatology (Oxford) 2006a;45:1238-46.
Westhovens R, Yocum D, Han J et al. The safety of infliximab, combined with background treatments, among patients with rheumatoid arthritis and various comorbidities: a large, randomized, placebo-controlled trial. Arthritis Rheum 2006b;54:1075-86.
Wilson K. Anti-TNF Therapies for Rheumatoid Arthritis: A systematic review of Historical Trends. 2008. [Unpublished].
Witney AG, Treharne GJ, Tavakoli M, Lyons AC, Vincent K, Scott DL, et al. The relationship of medical, demographic and psychosocial factors to direct and indirect health utility instruments in rheumatoid arthritis. Rheumatology (Oxford) 2006 Aug;45(8):975-81.
Young A, Dixey J, Kulinskaya E et al. Which patients stop working because of rheumatoid arthritis? Results of five years' follow up in 732 patients from the Early RA Study (ERAS). Ann Rheum Dis 2002;61:335-40.

Appendices

8.1 Appendix 1

8.1.1 SPC/IFU, scientific discussion or drafts.

Please see attached references (SPC included electronically).

8.2 *Appendix 2: Search strategy for section 5.1 (Identification of studies)*

The following information should be provided.

8.2.1 The specific databases searched and the service provider used

To identify relevant studies, searches were performed in the following databases:

- Medline via PubMed
- Embase via Embase.com
- Medline (R) In-Process and other non-indexed citations via PubMed
- Cochrane Library Central Trials Register via Wiley Interscience

8.2.2 The date on which the search was conducted.

Searches were performed on 23 March 2010.

8.2.3 The date span of the search.

Not restricted by date.

8.2.4 The complete search strategies used, including all the search terms: textwords (free text), subject index headings (for example, MeSH) and the relationship between the search terms (for example, Boolean).

EMBASE search strategy (3rd week March 2010)

1 rheumatoid arthritis
2 abatacept OR orencia
3 adalimumab OR humira
4 'certolizumab pegol' OR cimzia
5 etanercept OR enbrel
6 golimumab OR simponi
7 infliximab OR remicade
8 rituximab OR mabthera OR rituxan
9 tocilizumab OR atlizumab OR actemra OR roactemra
10 'tumor necrosis factor'
11 'tumor necrosis factor inhibitor'
12 'anti TNF'
13 'anti tumor necrosis factor'
14 'TNFR-Fc fusion protein'
15 'interleukin 6 antibody'
16 'anti interleukin'

17 'cd20 antibody'
18 'anti CD20'
19 'biological response modifier' AND 'disease modifying antirheumatic drug'
20 'biologic dmard'
21 'biologic agent'
22 #2–21/OR
23 'randomized controlled trial'
24 'controlled clinical trial'
25 randomisation
26 'random allocation'
27 'randomly allocated'
28 'double blind procedure'
29 'clinical trial'
30 'placebo controlled'
31 #23–30/OR
32 #1 AND #22 AND #31
33 #32 NOT review:it
34 #33 NOT animals (1712 references in total)

Medline (PubMed) search strategy (3rd week March 2010)

1 rheumatoid arthritis
2 abatacept OR orencia
3 adalimumab OR humira
4 certolizumab OR cimzia
5 etanercept OR enbrel
6 golimumab OR simponi
7 infliximab OR remicade
8 rituximab OR mabthera OR rituxan
9 tocilizumab OR atlizumab OR actemra OR roactemra
10 anti tnf
11 tumour necrosis factor
12 anti tumour necrosis factor
13 TNFR-Fc fusion protein
14 TNF receptor fusion protein
15 anti interleukin
16 anti CD20
17 biologic DMARD
18 biologic agent
19 #2–18/OR
20 randomized controlled trial
21 controlled clinical trial
22 random* allocate*
23 double blind method
24 clinical trial
25 placebo controlled
26 #20–25/OR

27 #1 AND #19 AND #26
28 #27 NOT review[PT]
29 #28 NOT animals (972 references in total)

Cochrane Central Register of Controlled Trials search strategy (3rd week March 2010)

1 rheumatoid NEXT arthritis
2 MeSH descriptor Arthritis, Rheumatoid, explode all trees
3 #1 OR #2
4 abatacept OR orencia
5 adalimumab OR humira
6 certolizumab OR cimzia
7 etanercept OR enbrel
8 golimumab OR simponi
9 infliximab OR remicade
10 rituximab OR mabthera OR rituxan
11 tocilizumab OR atlizumab OR actemra OR roactemra
12 tumor necrosis factor
13 tumour necrosis factor
14 anti TNF
15 anti tumor necrosis factor
16 anti interleukin
17 anti CD20
18 TNFR-Fc fusion protein
19 biologic DMARD
20 biologic agent
21 Mesh descriptor Receptor, Tumor Necrosis Factor, explode all trees
22 #4–21/OR
23 #3 AND #22 (408 references in total)

8.2.5 Details of any additional searches, such as searches of company databases (include a description of each database).

No additional searches were conducted.

8.2.6 The inclusion and exclusion criteria.

Studies in the systematic review were included or excluded according to the eligibility criteria described in Table 167.

Table 167. Inclusion and exclusion criteria

Inclusion criteria	
Populations	1. Adult patients (≥ 18 years) with active RA despite treatment with at least one conventional DMARD for ≥ 3 months; no previous use of anti-TNF-α agents or other biologic agents. 2. Adult patients (≥ 18 years) with active RA despite treatment with at least one anti-TNF-α agent.

Interventions	• Abatacept, adalimumab, certolizumab, etanercept, golimumab, infliximab, rituximab or tocilizumab compared with any other agent including placebo.
Study design	• Double-blind, randomized controlled trials.
Outcomes	Any of the following outcomes of interest: • measures of treatment efficacy: ACR responses, mean DAS or DAS28, number of patients achieving low DAS (< 3.2) or DAS remission (< 2.6), HAQ-DI. • measures of safety and tolerability: adverse events, treatment discontinuations.
Report characteristics	• Articles for which the full text was available in English. • No publication date restrictions were imposed.
Exclusion criteria	
Populations	• Conventional DMARD-naïve patients. • Mixed populations of both conventional DMARD-experienced and anti-TNF-α-experienced patients (> 10% from each group), unless analysed separately
Study design	• Studies with no appropriate comparisons between biologic agents and other active comparators or placebo (e.g. open-label extensions and observational studies). • Studies in which the drug of interest is not administered at the EMEA-approved dose or details of dosing are not given. If a study includes more than one treatment arm of the intervention of interest, one of them must be at the approved dose.
Report characteristics	• Reviews, systematic reviews and meta-analyses.

Table 168 presents the rationale for the 74 excluded trials.

Table 168. Excluded references

Reference	*Study drug*	*Reason(s) for exclusion*
Emery, Kosinski *et al.* (2006) *J Rheumatol* 33:681-9	abatacept	secondary study of Kremer, Westhovens *et al.* (2003), reports SF-36 only
Weisman, Durez *et al.* (2006) *J Rheumatol* 33:2162-6	abatacept	secondary study of Kremer, Westhovens *et al.* (2003), does not report relevant outcomes
Weinblatt, Schiff *et al.* (2007) *Ann Rheum Dis* 66:228-34	abatacept	used non-approved dosing
Schiff and Bessette (2010) *Clin Rheumatol* DOI: 10.1007/s10067-009-1363-0	abatacept	review article
Fernandez-Lopez and Blanco (2006) *Reumatologia Clinica Suplementos* 1:34-43	abatacept	foreign-language duplicate of Genovese, Becker *et al.* (2005)
Wells, Li *et al.* (2008) *Ann Rheum Dis* 67:260-5	abatacept	secondary study of ATTAIN, does not report change from baseline data
Westhovens, Cole *et al.* (2006) *Rheumatology* 45:1238-46	abatacept	secondary study of ATTAIN, does not report relevant outcomes
Moreland, Alten *et al.* (2002) *Arthritis Rheum* 46:1470-9	abatacept	unclear what proportion of patients were biologic-experienced
Rau, Simianer *et al.* (2004) *Scand J Rheumatol* 33:145-53	adalimumab	used non-approved dosing
Van De Putte, Rau *et al.* (2003) *Ann Rheum Dis* 62:1168-77	adalimumab	used non-approved dosing
Keystone, Haraoui *et al.* (2003) *Clin Exp Rheumatol* 21:S198-S9	adalimumab	review article
Breedveld, Weisman *et al.* (2006) *Arthritis Rheum* 54:26-37	adalimumab	PREMIER study, only one third of patients DMARD-experienced
Emery, Genovese *et al.* (2009) *J Rheumatol* 36:1429-41	adalimumab	secondary study of PREMIER
Kimel, Cifaldi *et al.* (2008) *J Rheumatol* 35:206-15	adalimumab	secondary study of PREMIER
Bejarano, Quinn *et al.* (2008) *Arthritis Rheum* 59:1467-74	adalimumab	majority of patients DMARD-naive
Huang, Zhang *et al.* (2009) *Zhonghua Nei Ke Za Zhi* 48:916-21	adalimumab	English language version of the full article not available
Yount, Sorensen *et al.* (2007) *Clin Exp Rheumatol* 25:838-46	adalimumab	does not report relevant outcomes
Fleischmann *et al.* (2009) *Ann Rheum Dis* 68: 805-811	certolizumab	FAST4WARD study includes dose outside of licence
Kavanaugh, Smolen *et al.* (2009) *Arthritis Care Res* 61:1592-600	certolizumab	secondary study of RAPID 1 and RAPID 2, does not report relevant outcomes

Reference	*Study drug*	*Reason(s) for exclusion*
Emery, Breedveld *et al.* (2008) *Lancet* 372:375-82	etanercept	COMET study, only 20% patients DMARD-experienced
Bathon, Martin *et al.* (2000) *N Engl J Med* 343:1586-93	etanercept	ERA study, majority of patients DMARD-naive
Bathon and Genovese (2003) *Clin Exp Rheumatol* 21:S195-S7	etanercept	secondary study of ERA
Moreland, Genovese *et al.* (2006) *Arthritis Rheum* 55:287-93	etanercept	does not report relevant outcomes
Kavanaugh, Klareskog *et al.* (2008) *Ann Rheum Dis* 67:1444-7	etanercept	secondary study of TEMPO, compares non-relevant patient groups
Landewe, Van Der Heijde *et al.* (2006) *Arthritis Rheum* 54:3119-25	etanercept	secondary study of TEMPO, does not report relevant outcomes
Bankhurst (1999) *Clin Exp Rheumatol* 17:S69-S72	etanercept	review article
Chen, Lin *et al.* (2006) *Ann Rheum Dis* 65:35-9	etanercept	unclear if placebo-controlled and blinded
Combe, Codreanu *et al.* (2009) *Ann Rheum Dis* 68:1146-52	etanercept	reports 2-year data, not within systematic review scope
Hu, Bao *et al.* (2009) *Rheumatol Int* 29:297-303	etanercept	unclear what proportion of patients were DMARD-experienced
Keystone, Schiff *et al.* (2004) *Arthritis Rheum* 50:353-63	etanercept	placebo group switched to etanercept at week 8
Kosinski, Kujawski *et al.* (2002) *Am J Manag Care*8:231-40	etanercept	unclear what proportion of patients were DMARD-experienced
Lan, Chou *et al.* (2004) *J Formos Med Assoc* 103:618-23	etanercept	unclear what proportion of patients were DMARD-experienced
Mathias, Colwell *et al.* (2000) *Clin Ther* 22:128-39	etanercept	secondary study of Moreland, Schiff *et al.* (1999), no relevant outcomes reported
Moreland, Baumgartner *et al.* (1997) *N Engl J Med* 337:141-7	etanercept	used non-approved dosing
Weinblatt, Schiff *et al.* (2008) *Arthritis Rheum* 58:1921-30	etanercept	no placebo control
Weisman, Paulus *et al.* (2007) *Rheumatology* 46:1122-5	etanercept	unclear what proportion of patients were DMARD-experienced
De Filippis, Caliri *et al.* (2006) *Panminerva Medica* 48:129-35	etanercept, infliximab	open-label, observational study
Emery, Fleischmann *et al.* (2009) *Arthritis Rheum* 60:2272-83	golimumab	GO-BEFORE study, only 50–60% of patients were DMARD-experienced
Kremer, Ritchlin *et al.* (2010) *Arthritis Rheum* 62:917-28	golimumab	used non-approved dosing, unclear what proportion of patients were TNF inhibitor-experienced
Zhou, Jang *et al.* (2007) *J Clin Pharmacol* 47:383-96	infliximab	no relevant outcomes reported
Smolen, Han *et al.* (2006) *Arthritis Rheum* 54:716-22	infliximab	secondary study of ASPIRE
Smolen, Han *et al.* (2009) *Ann Rheum Dis* 68:823-7	infliximab	secondary study of ASPIRE
St Clair, van der Heijde *et al.* (2004) *Arthritis Rheum* 50:3432-43	infliximab	ASPIRE study, majority of patients DMARD-naive

Reference	*Study drug*	*Reason(s) for exclusion*
Visvanathan, Marini *et al.* (2007) *J Rheumatol* 34:1465-74	infliximab	secondary study of ASPIRE
Breedveld, Emery *et al.* (2004) *Ann Rheum Dis* 63:149-55	infliximab	secondary study of ATTRACT, examines non-relevant subgroup of patients
Maini, Breedveld *et al.* (2004) *Arthritis Rheum* 50:1051-65	infliximab	secondary study of ATTRACT, reports 2-year data, not within systematic review scope
Smolen, Han *et al.* (2005) *Arthritis Rheum* 52:1020-30	infliximab	secondary study of ATTRACT, examines non-relevant subgroup of patients
St.clair, Wagner *et al.* (2002) *Arthritis Rheum* 46:1451-9	infliximab	no relevant outcomes reported
Han, Smolen *et al.* (2008) *Arthritis Rheum* 59:510-4	infliximab	secondary study of ATTRACT and ASPIRE, examines non-relevant subgroup of patients
Allaart, Geokoop-Ruiterman *et al.* (2006) *Clin Exp Rheumatol* 24:S77-S82	infliximab	secondary study of BeSt
Allaart, Breedveld *et al.* (2007) *J Rheumatol Suppl* 80:25-33	infliximab	secondary study of BeSt
Goekoop-Ruiterman, de Vries-Bouwstra *et al.* (2005) *Arthritis Rheum* 52:3381-90	infliximab	BeSt study, majority of patients were DMARD-naive
Goekoop-Ruiterman, de Vries-Bouwstra *et al.* (2007) *Ann Intern Med* 146:406-15	infliximab	secondary study of BeSt
Goekoop-Ruiterman, de Vries-Bouwstra *et al.* (2008) *Arthritis Rheum* 58:S126-35	infliximab	secondary study of BeSt
Van Der Kooij, Goekoop-Ruiterman *et al.* (2008) *Ann Rheum Dis* 67:266-9	infliximab	secondary study of BeSt
Takeuchi, Miyasaka *et al.* (2009) *Modern Rheumatology* 19:478-87	infliximab	RISING study, no control
Rahman, Strusberg *et al.* (2007) *Ann Rheum Dis* 66:1233-8	infliximab	secondary study of START, no placebo comparison
van Vollenhoven, Ernestam *et al.* (2009) *Lancet* 374:459-66	infliximab	Swefot study, patients were DMARD-naive
Durez, Nzeusseu Toukap *et al.* (2004) *Ann Rheum Dis* 63:1069-74	infliximab	unclear if blinded
Durez, Malghem *et al.* (2007) *Arthritis Rheum* 56:3919-27	infliximab	unclear what proportion of patients were DMARD-experienced
Kavanaugh, St Clair *et al.* (2000) *J Rheumatol* 27:841-50	infliximab	only first infusion blinded, then open-label study
Montecucco (2005) *Clin Exp Rheumatol* 23:289-91	infliximab	no relevant outcomes reported
Taylor, Steuer *et al.* (2004) *Arthritis Rheum* 50:1107-16	infliximab	used non-approved dosing

Reference	*Study drug*	*Reason(s) for exclusion*
Taylor, Steuer *et al.* (2006) *Arthritis Rheum* 54:47-53	infliximab	secondary study of Taylor, Steuer *et al.* (2004)
Zhang, Hou *et al.* (2006) *International Journal of Rheumatic Disease* 9:127-30	infliximab	data not provided at timepoints relevant to systematic review
Emery, Fleischmann *et al.* (2006) *Arthritis Rheum* 54:1390-400	rituximab	DANCER study, 25–30% patients were TNF inhibitor-experienced
Mease, Revicki *et al.* (2008) *J Rheumatol* 35:20-30	rituximab	secondary study of DANCER
Keystone, Emery *et al.* (2009) *Ann Rheum Dis* 68:216-21	rituximab	REFLEX study, no relevant outcomes reported
Keystone (2005) *Rheumatology* 44:ii8-ii12	rituximab	review article
Owczarczyk, Hellmann *et al.* (2008) *Ann Rheum Dis* 67:1648-9	rituximab	unclear if randomized, controlled trial
Maini, Taylor *et al.* (2006) *Arthritis Rheum* 54:2817-29	tocilizumab	CHARISMA study, approximately 1/6 of patients were anti-TNF experienced
Jones, Sebba *et al.* (2010) *Ann Rheum Dis* 69:88-96	tocilizumab	AMBITION study, unclear what proportion of patients were DMARD-experienced
Garnero, Thompson *et al.* (2010) *Arthritis Rheum* 62:33-43	tocilizumab	OPTION study, no relevant outcomes reported
Nishimoto, Hashimoto *et al.* (2007) *Ann Rheum Dis* 66:1162-7	tocilizumab	SAMURAI study, 20% patients appeared to be biologic-experienced
Nishimoto, Yoshizaki *et al.* (2004) *Arthritis Rheum* 50:1761-9	tocilizumab	unclear what proportion of patients were biologic-experienced

8.2.7 The data abstraction strategy.

Data were extracted from the identified studies by one reviewer and recorded on a data extraction form. The extraction form was checked by a second reviewer against the original articles. Study design information, patient baseline characteristics and the outcomes listed in Table 167 were sought. For outcomes not reported in the published golimumab studies, data were obtained from the clinical study reports. For all studies, if data for a relevant outcome were available only in graphical form, the values were estimated from the relevant figure.

8.3 *Appendix 3: Quality assessment of RCT(s)*

Table 169 Quality Assessment of GO FORWARD

GO FORWARD		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Adequate sample size Number randomised was stated True randomisation carried out Subjects were randomized in a 3:3:2:2 ratios to 1 of 4 treatment groups: placebo plus MTX, golimumab 100 mg plus placebo, golimumab 50 mg plus MTX , and golimumab 100 mg plus MTX. Relatively even treatment balance within sites was ensured, within baseline MTX usage and within the study overall, using an adaptive stratified randomisation design.	Yes
Was the concealment of treatment allocation adequate?	Randomised treatment allocation was done using a centralised interactive voice response system	Yes

Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Demographic characteristics of subjects at baseline were generally well balanced across treatment groups: • majority of subjects were women (82.0% • majority subjects were Caucasian (75.9%) • mean age was 51.2 years • mean duration of disease (8.62 years)	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Randomisation files containing treatment assignments for individual subjects were maintained in limited-access directories within the electronic data filing system at the central randomization centre. Personnel having contact with study sites, including the medical monitor, remained blinded to the treatment assignment of individual subjects until the 24-week database lock. All site monitors, site personnel, and subjects remained blinded to treatment assignment until the last subject completes Week 52 evaluations and the database is locked.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	> 90% patients were part of follow-up assessment	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the publication	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes

Table 170. Quality Assessment of GO AFTER

GO AFTER		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Adequate sample size Number randomised was stated True randomisation carried out Subjects were randomised in a 1:1.1 ratio to 1 of 3 treatment groups: placebo, golimumab 50 mg, and golimumab 100 mg. Relatively even treatment balance within sites was ensured, within baseline MTX usage and within the study overall, using an adaptive stratified randomisation design.	Yes
Was the concealment of treatment allocation adequate?	Randomised treatment allocation was done using a centralised interactive voice response system	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Demographic characteristics of subjects at baseline were generally well balanced across treatment groups: • majority of subjects were women (85.0%) • most subjects were Caucasian (88.2%) • mean age was 54.8 years • mean duration of disease (12.40 years)	Yes

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Randomisation files containing treatment assignments for individual subjects were maintained in limited-access directories within the electronic data filing system at the central randomisation centre. Both patients and investigators were masked to treatment assignment. Personnel having contact with study sites, including the medical monitor, remained blinded to the treatment assignment of individual subjects until the 24-week database lock. Furthermore, all site monitors, site personnel, and subjects remained blinded to treatment assignment until the last subject completes Week 52 evaluations and the database is locked.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	> 80% patients were part of follow-up assessment	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the publication	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes

Table 171 Quality assessment results for RCTs

Clinical Trial	Combe 2006	Moreland 1999	Weinblatt 1999	GO-AFTER	GO-FORWARD	Kay 2008	ATTEST Schiff	ATTRACT	ATTRAT	START	Abe 2006	Maini 1998	Edwards 2004	Strand 2006	REFLEX (Cohen)	REFLEX	OPTION	RADIATE	SATORI	TOWARD
Was randomisation carried out appropriately?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	NC	Y	Y	Y	Y
Was the concealment of treatment allocation adequate?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	NC	Y	Y	Y	Y
Were the groups similar at the outset of the study in terms of prognostic factors?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Were the care providers, participants and outcome assessors blind to treatment allocation?	Y	Y	Y	Y	Y	Y	Y	Y	Y	NC	Y	Y	Y	Y	Y	Y	Y	NC	Y	Y
Were there any unexpected imbalances in drop-outs between groups?	N	N	N	N	N	Y	Y	N	N	N	N	N	N	N	Y	N	Y	N	Y	N
Is there any evidence to suggest that the authors measured more outcomes than they reported?	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Clinical Trial	Kremer 2003	Kremer 2005	Kremer 2006	AIM Russel et al	ASSURE	ATTAIN	Kim 2007	van de Putte 2004	ARMADA	CHANGE	DE019	STAR	Chen 2009	RAPID 1	RAPID 1 (Strand)	RAPID 2	TEMPO	TEMPO (van der	TEMPO & ERA
Was randomisation carried out appropriately?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Was the concealment of treatment allocation adequate?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Were the groups similar at the outset of the study in terms of prognostic factors?	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Were the care providers, participants and outcome assessors blind to treatment allocation?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Were there any unexpected imbalances in drop-outs between groups?	N	N	N	N	N	N	N	Y	N	Y	Y	N	N	Y	Y	Y	N	N	N
Is there any evidence to suggest that the authors measured more outcomes than they reported?	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

8.4 *Appendix 4: Search strategy for section 5.7 (Indirect and mixed treatment comparisons)*

The following information should be provided.

8.4.1 The specific databases searched and the service provider used (for example, Dialog, DataStar, OVID, Silver Platter), including at least:

- Medline
- Embase
- Medline (R) In-Process
- The Cochrane Library.

No additional data searches were conducted. This information was expected to be available from the search strategy in Appendix 2 (section 9.2).

8.4.2 The date on which the search was conducted.

Please see section 9.2.2.

8.4.3 The date span of the search.

Please see section 9.2.3.

8.4.4 The complete search strategies used, including all the search terms: textwords (free text), subject index headings (for example, MeSH) and the relationship between the search terms (for example, Boolean).

Please see section 9.2.4.

8.4.5 Details of any additional searches (for example, searches of company databases [include a description of each database]).

Please see section 9.2.5.

8.4.6 The inclusion and exclusion criteria.

Please see section 9.2.6.

8.4.7 The data abstraction strategy.

Please see section 9.2.7.

8.5 *Appendix 5: Quality assessment of comparator RCT(s) in section 5.7 (Indirect and mixed treatment comparisons)*

Table 172. Quality assessment of comparator RCTs

Kim et al 2007		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Double blind 1:1 randomisation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	Not reported in trial publication	Not clear
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Baseline demographic and disease characteristics were comparable between the two treatment groups.	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind so presumably participants and outcome assessors were unaware of treatment allocation, though this is not specifically stated in the methods.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No. Equal number of withdrawals in the treatment groups	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the publication	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT analysis. Patients with missing data at 24 week follow up counted as non-responders	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

Van De Putte 2004		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**

Was randomisation carried out appropriately?	Computer generated double blind randomisation Power calculation for sample size Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	Not reported in trial publication	Not clear
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	There were no statistically significant differences in the demographic characteristics and baseline disease activity between the treatment groups	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind so presumably participants and outcome assessors were unaware of treatment allocation, though this is not specifically stated in the methods. Further, blinding of study drug was achieved through packaging procedure for the study drug.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Withdrawals occurred in 118/434 (27.2%) adalimumab treated patients and 62/110 (56.4%) placebo treated patients. All withdrawals from the treatment groups explained.	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the publications	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT analysis. Patients not completing the trial (that is, who were withdrawn or required rescue) despite fulfilling ACR criteria were considered non-responders.	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

ARMADA (Weinblatt et al 2003)		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Double blind block randomisation Power calculation for sample size Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	Patients were instructed in self-injection techniques. It is not clear how treatment concealment from the subjects was established.	Not clear

Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	There were no statistically significant differences in the demographic and baseline characteristics among the treatment groups.	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind so presumably participants and outcome assessors were unaware of treatment allocation, though this is not specifically stated in the methods.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No unexpected imbalances between groups. 18 patients withdrew from the study prematurely and all of these were explained.	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications.	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT analysis. Patients who dropped out before week 24 and patients who did not achieve an ACR20 response were classified as non-responders.	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

CHANGE (Miyasaka et al 2008)		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**

Was randomisation carried out appropriately?	Double blind randomisation. Patients were randomly assigned in a 1:1:1:1 ratio to four treatment groups: Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	Study drug was administered by a physician or nurse supervised by an investigator. As this study was double blind presumably nurse/investigator was blind to treatment allocation.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Baseline demographics were similar between groups. Baseline disease characteristics were consistent with what is generally observed in patients with RA with moderate to severe disease and were comparable between treatment groups with the following exceptions: the mean of the patient's global assessment of disease activity for the placebo group (64.6 mm) was lower than those for the adalimumab groups (20 mg: 73.1 mm; 40 mg: 71.2 mm; 80 mg: 75.7 mm; P = 0.003); the mean of HAQ DI for the placebo group (1.39) was lower than those for the adalimumab groups (20 mg: 1.57; 40 mg: 1.64; 80 mg: 1.77; P = 0.010); and the mean CRP for the 20 mg group (4.97 mg/dl) was lower than those for other groups (placebo: 5.86 mg/dl; 40 mg: 6.48 mg/dl; 80 mg: 6.56 mg/dl; P = 0.020).	No
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind but not clear whether outcomes assessment was blinded.	Not clear
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No unexpected imbalances. Reasons for discontinuations explained.	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications.	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes, ITT analysis. LOCF for missing data.	Yes

Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination

DE019 (Keystone 2004)		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Double blind RCT Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	As this study was a double blind study presumably both subjects and investigators were blind to treatment allocation.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Patients were well matched across the treatment groups. There were no statistically significant differences in baseline and disease characteristics between the treatment groups.	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind but not clear whether outcomes assessment was blinded.	Not clear
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No unexpected imbalances in drop outs. Reasons for discontinuations explained.	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	None such reported in the trial publication	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT and missing values imputed	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

STAR (Furst et al 2003)		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**

Was randomisation carried out appropriately?	Double blind RCT Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	As this study was a double blind study presumably both subjects and investigators were blind to treatment allocation.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Yes. Demographic and baseline disease characteristics were balanced between the groups at baseline.	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind but not clear whether outcomes assessment was blinded.	Not clear
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No unexpected imbalances in drop outs. Reasons for discontinuations explained.	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT and missing values imputed	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

(Chen et al 2009)		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Double blind RCT No Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	As this study was a double blind study presumably both subjects and investigators were blind to treatment allocation though this is not specifically made clear in the methods.	Yes

Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	There was no significant difference in baseline characteristic between the groups.	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind but not clear whether outcomes assessment was blinded.	Not clear
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Not stated	Not clear
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT and missing values imputed	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

RAPID 1 (Keystone et al 2008)		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Double blind RCT Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	As this study was a double blind study presumably both subjects and investigators were blind to treatment allocation though methods of concealment not specified.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	The baseline demographic features and disease activity status of the study patients were similar among the treatment groups.	Yes

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind and the outcomes that were assessed by radiographs were assessed by blinded radiographers.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	63% of placebo patients and 21%, 17% of 200mg and 400mg Certoluzimab patients withdrew from the study.	Yes
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT and missing values imputed	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

RAPID 2 (Smolen 2009)		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Double blind RCT Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	As this study was a double blind study presumably both subjects and investigators were blind to treatment allocation though there is no detail on steps taken to ensure concealment of treatment allocation.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Patient demographics and baseline characteristics were similar in the three treatment groups and indicated high baseline disease activity	Yes

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind and the outcomes that were measured by radiographs were assessed by blinded radiographers.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	More placebo-treated patients (79.5%) discontinued treatment owing to lack of ACR20 response at week 16 versus certolizumab pegol 200 mg (19.9%) and 400 mg (18.7%).	Yes
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications.	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT and missing values imputed	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

TEMPO (Klareskog et al 2004)		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Centralised double blind telephone randomisation Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	As this study was a double blind study presumably both subjects and investigators were blind to treatment allocation. Injections identical in appearance used in different treatment groups.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Demographics or baseline disease characteristics did not differ between treatment groups.	Yes

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind and the outcomes that were measured by radiographs were assessed by blinded radiographers.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	30%, 23% and 16% from the Methotrexate, Etanercept and Methotrexate & Etanercept combination groups respectively discontinued. Reasons for discontinuations explained.	Not clear
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT and missing values imputed	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

Combe et al 2006.		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Double blind RCT Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	As this study was a double blind study presumably both subjects and investigators were blind to treatment allocation. Also all patients received identical-appearing injectible and oral test articles.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	No major differences among the groups in baseline characteristics other than the number of patients with a history of corticosteroid use.	Yes

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind so presumably both participants and outcome assessors blind to treatment allocation.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Unsatisfactory response to treatment, the most common primary reason for discontinuation, was reported by more patients receiving sulfasalazine alone (24%) than by those receiving etanercept alone (1%) or etanercept and sulfasalazine (4%). Reasons for discontinuations explained.	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT and missing values imputed using LOCF	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

Moreland et al 1999		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Double blind block randomised RCT No Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	As this study was a double blind study presumably both subjects and investigators were blind to treatment allocation. Placebo and treatment was supplied in an identical format and was formulated in the same way.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Aside from differences in concurrent medications (more patients in the 25-mg group were receiving corticosteroids and more placebo recipients were receiving (NSAIDs), no baseline imbalances were detected.	Yes

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind so presumably both participants and outcome assessors blind to treatment allocation.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	More patients from the placebo group than the active treatment groups discontinued due to lack of efficacy. No unexpected imbalances in drop outs. Reasons for discontinuations explained.	Yes
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT and missing values imputed using LOCF	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

Weinblatt et al 1999		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Double blind RCT with 2:1 randomisation to treatment arm Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	As this study was a double blind study presumably both subjects and investigators were blind to treatment allocation. No additional information given about methods used for treatment concealment.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	The treatment groups were generally well matched.	Yes

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind so presumably both participants and outcome assessors blind to treatment allocation.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	20% placebo patients and 3% etanercept plus methotrexate group patients discontinued. Reasons for discontinuations explained.	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT and missing values imputed using LOCF	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

Kay et al 2008		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Double blind RCT Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	As this study was a double blind study presumably both subjects and investigators were blind to treatment allocation. Patients in the golimumab groups remained blinded to their dose assignment through the end of the study.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	The population was somewhat heterogeneous because of the small number of patients in each treatment group, but none of the baseline characteristics of the combined. Golimumab groups was significantly different from those of the placebo group (P > 0.05).	Yes

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind so presumably both participants and outcome assessors blind to treatment allocation.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Through week 52, a significantly greater proportion of patients in the placebo/infliximab plus MTX group (40.0%) discontinued treatment compared with the proportion of patients in the combined golimumab plus MTX groups (21.2%) (P <0.0217).	Yes
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT and missing values imputed using LOCF.	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

ATTEST (Schiff et al 2008)		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Double blind, double dummy RCT Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	As this study was a double blind study presumably both subjects and investigators were blind to treatment allocation. Also 2 intravenous bags were infused simultaneously to ensure blinding to treatment group assignment.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Baseline demographics and clinical characteristics were similar between groups.	Yes

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind, double dummy so presumably both participants and outcome assessors blind to treatment allocation.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	During the first 6 months, discontinuations occurred in 5.8%, 2.7% and 7.9% of the abatacept, placebo and infliximab groups, respectively. Reasons for discontinuations explained.	Not clear
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT and missing values imputed using LOCF.	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

ATTRACT (Maini et al 1999)		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Double blind RCT Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	As this study was a double blind study presumably both subjects and investigators were blind to treatment allocation.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	The baseline characteristics of the five treatment groups were well matched and consisted of a predominantly white, female, rheumatoid factor-positive population, with a median age range of 51–56 years, and disease duration of 7·2 to 9·0 years.	Yes

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind, so presumably both participants and outcome assessors blind to treatment allocation. Further an independent assessor, unaware of the patient assignment or other clinical response indices and not involved in the administration of the infusions, assessed the joint scores.	Not clear
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No unexpected imbalances in drop outs 36% of placebo and 9-18% of Infliximab treated patients discontinued, mostly for lack of efficacy.	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT and missing values imputed for non responders.	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

START (Westhovens et al 1999).		
Study question	How is the question addressed in the study?	Grade (yes/no/not clear/N/A)
Was randomisation carried out appropriately?	Double blind RCT Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	As this study was a double blind study presumably both subjects and investigators were blind to treatment allocation. Both infliximab and placebo were supplied as sterile, white, lyophilized powders in single use 20-ml vials Patients, investigators, and other study personnel, except for pharmacists, were blinded to the study treatment assignments	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	The treatment groups were well matched with regard to the patients' baseline characteristics. Patients receiving placebo + MTX had longer disease duration than those receiving 3 mg/kg infliximab and group 3 (receiving 10 mg/kg infliximab). However, this difference between the groups was not statistically significant ($P < 0.083$).	Yes

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind, so presumably both participants and outcome assessors blind to treatment allocation. Patients, investigators, and other study personnel, except for pharmacists, were blinded to the study treatment assignments	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	About 9% of patients from each treatment group discontinued by week 54. No unexpected imbalances in drop outs. Reasons for discontinuations explained.	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT and missing values imputed using LOCF.	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

Abe et al 2006		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Double blind RCT Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	As this study was a double blind study presumably both subjects and investigators were blind to treatment allocation.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Baseline demographics were comparable among the 3 groups, with the exception of body weight. The difference had no influence on the result of the primary endpoint using covariance adjustment.	Yes

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind, so presumably both participants and outcome assessors blind to treatment allocation.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No unexpected imbalances in drop outs. 11% of placebo patients and 5% of infliximab patients discontinued . Reasons for discontinuations explained.	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT and missing values imputed using LOCF	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

Maini et al 1999		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Double blind RCT Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	As this study was a double blind study presumably both subjects and investigators were blind to treatment allocation. Randomization was performed centrally, and the nature of the codcd study medications was not revealed to the patients or the assessors.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	In these relatively small patient groups, there were some between-group differences in demographic characteristics, baseline medication, or baseline disease activity, but these did not reach statistical significance, with the exception of the baseline HAQ Disability Index (P = 0.026 in an overall treatment effect). In the analysis of efficacy, adjustment for baseline HAQ there was no change in P values on the primary end-point.	Yes

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind, so presumably both participants and outcome assessors blind to treatment allocation. The nature of the codcd study medications was not revealed to the patients or the assessors.	Not clear
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No unexpected imbalances in drop outs. The proportion of patients who discontinued treatment because of lack of efficacy was highest in the placebo infusion plus MTX and 1 mg/kg cA2 without MTX groups (57% and 33%, respectively), and was notably lower in the groups receiving 3 and 10 mg/kg of cA2 alone (7% and 13%). Reasons for discontinuations explained.	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT.	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

ASSURE (Weinblatt al 200)		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Double blind RCT Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	As this study was a double blind study presumably both subjects and investigators were blind to treatment allocation.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	The treatment groups were similar with respect to demographic characteristics and most baseline disease characteristics, with the exception of C-reactive protein level, which was slightly higher in the subgroup receiving non-biologic background therapy	Yes

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind, so presumably both participants and outcome assessors blind to treatment allocation.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	13% in the Abatacept group and 18% in the placebo group discontinued treatment. treatment. Reasons for discontinuations explained.	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT though not clear how missing values dealt with	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

ATTAIN (Genovese et al 2003)		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Double blind RCT Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	There was a central randomization system and the randomization schedule was generated by a drug-management group.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Baseline demographic and clinical characteristics were similar in the two groups. Patients had active disease at baseline as evidenced by high mean counts of tender and swollen joints, scores for the HAQ disability index, C-reactive protein levels, and the DAS28. The percentages of current and former anti–TNFa users were similar in the abatacept and placebo groups (38% and 41%, respectively; and 62% and 59%, respectively).	Yes

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	The drug was prepared by pharmacists or other qualified personnel who had no interaction with the patients. Medication was administered intravenously in a blinded fashion by qualified personnel. All clinical assessments of response were performed in a blinded fashion by the same trained assessors throughout the study	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No unexpected imbalances in drop outs. 14% of the Abatacept group and 26% of the placebo group discontinued. Lack of efficacy was the main reason for discontinuation in both groups.	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT and missing values imputed using LOCF	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

REFLEX (Cohen et al 2006)		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Double blind RCT Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	This was a blinded study, with the study sponsor, investigators, and patients unaware of the treatment assignment of each patient.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Both study groups were balanced for age, sex, disease activity, and previous and concomitant treatments for RA	Yes

Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	This was a blinded study, with the study sponsor, investigators, and patients unaware of the treatment assignment of each patient.	Yes
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	46% in the placebo group and 18% in the Rituximab group withdrew from the study Reasons for discontinuations explained.	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT Patients who withdrew prematurely from the study or who started rescue therapy were included in the ITT population as non-responders. Not clear how missing values treated.	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

RADIATE (Emery et al 2008)		
Study question	**How is the question addressed in the study?**	**Grade (yes/no/not clear/N/A)**
Was randomisation carried out appropriately?	Double blind RCT Power calculation Number randomised was stated	Yes
Was the concealment of treatment allocation adequate?	As this study was a double blind study presumably both subjects and investigators were blind to treatment allocation.	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	The three groups were reasonably well balanced for demographics and RA characteristics at baseline except on CRP or ESR values.	Not clear
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Study was double blind, so presumably both participants and outcome assessors blind to treatment allocation. Joint assessors were blinded as to other data including CRP, ESR and treatment assignment.	Yes

Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	13% 8mg /kg Tociluzamab, 15% 4mg /kg Tociluzamab & 18% placebo withdrew from the study. No unexpected imbalances in drop outs. Reasons for discontinuations explained.	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No such reference in the study publications	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes ITT and missing values imputed using LOCF	Yes
Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination		

8.6 Appendix 6: Search strategy for section 5.8 (Non-RCT evidence)

Only RCT evidence was considered within searches.

8.6.1 The specific databases searched and the service provider used (for example, Dialog, DataStar, OVID, Silver Platter), including at least:

- Medline
- Embase
- Medline (R) In-Process
- The Cochrane Library.

Not applicable.

8.6.2 The date on which the search was conducted.

Not applicable.

8.6.3 The date span of the search.

Not applicable.

8.6.4 The complete search strategies used, including all the search terms: textwords (free text), subject index headings (for example, MeSH) and the relationship between the search terms (for example, Boolean).

Not applicable.

8.6.5 Details of any additional searches (for example, searches of company databases [include a description of each database]).

Not applicable.

8.6.6 The inclusion and exclusion criteria.

Not applicable.

8.6.7 The data abstraction strategy.

Not applicable.

8.7 Appendix 7: Quality assessment of non-RCT(s) in section 5.8 (Non-RCT evidence)

8.7.1 Please tabulate the quality assessment of each of the non-RCTs identified.

Only RCT evidence was considered within searches.

8.8 *Appendix 8: Search strategy for section 5.9 (Adverse events)*

8.8.1 The specific databases searched and the service provider used

- Medline
- Embase
- Medline (R) In-Process
- The Cochrane Library

8.8.2 The date on which the search was conducted.

25 January 2010.

8.8.3 The date span of the search.

No date restrictions.

8.8.4 The complete search strategies used, including all the search terms: textwords (free text), subject index headings (for example, MeSH) and the relationship between the search terms (for example, Boolean).

The search strategies used to identify adverse events are outlined below.

Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R)

1 (etanercept or enbrel).ab,ti. (2439)
2 (infliximab or remicade).ab,ti. (4327)
3 (adalimumab or humira).ab,ti. (1135)
4 (golimumab or simponi).ab,ti. (2)
5 (certolizumab or cimzia).ab,ti. (6)
6 (rituximab or mabthera).ab,ti (2316)
7 (abatacept or orencia).ab,ti (652)
8 OR/1-7 (7281)
9 Safety/ (27578)
10 (safe or safety).ab,ti. (296803)
11 (side effect or side effects).ab,ti. (136384)
12 emergency treatment.ab,ti. (2773)
13 undesirable effect$.ab,ti. (1538)
14 tolerability.ab,ti. (21310)

15 Drug Toxicity/ (3176)
16 toxicity.ab,ti. (183499)
17 Adverse Drug Reaction Reporting Systems/ (4034)
18 adrs.ab,ti. (1095)
19 (adverse adj3 (effect or effects or reaction or reactions or event or events or outcome or outcomes)).ab,ti. (160529)
20 (undesire$ adj2 (effect or effects or reaction or reactions or event or events or outcome or outcomes)).ab,ti. (1034)
21 Drug Hypersensitivity/ (17890)
22 (hypersensit$ or hyper sensit$).ab,ti. (46875)
23 harm$.ab,ti. (64460)
24 OR/9-23 (919647)
25 8 and 24 (2134)
26 exp infection/ci [Chemically induced] (2951)
27 exp urinary tract infections/ci [Chemically induced] (62)
28 exp respiratory tract infections/ci [Chemically induced] (3720)
29 exp skin diseases, infectious/ci [Chemically induced] (458)
30 exp bone diseases, infectious/ci [Chemically induced] (137)
31 exp arthritis, infectious/ci [Chemically induced] (57)
32 exp neoplasms/ci [Chemically induced] (50719)
33 exp tuberculosis/ci [Chemically induced] (323)
34 OR/26-33 (57199)
35 25 and 34 (37)
36 from 35 keep 2,11 (2)

EMBASE <1988 to 2010 Week 03>

1 (etanercept or enbrel).ab,ti. (2443)
2 (infliximab or remicade).ab,ti. (4406)
3 (adalimumab or humira).ab,ti. (1142)
4 (golimumab or simponi).ti,ab. (39)
5 (certolizumab or cimzia).ab,ti. (51)
6 (rituximab or mabthera).ab,ti (2934)
7 (abatacept or orencia).ab,ti (947)
8 OR/1-7 (6303)
9 (safe or safety).ti,ab. (242470)
10 side effect$.ti,ab. (106920)
11 emergency treatment.ti,ab. (1439)
12 undesirable effect$.ti,ab. (1269)
13 toxicity.ti,ab. (147110)

14 adrs.ti,ab. (1223)
15 (adverse adj3 (effect or effects or reaction or reactions or event or events or outcome or outcomes)).ti,ab. (142027)
16 safety/ or drug safety/ (190216)
17 side effect/ (99926)
18 adverse drug reaction/ (7122)
19 drug tolerability/ (58872)
20 toxicity/ or drug toxicity/ (21491)
21 drug surveillance program$/ (7548)
22 adverse outcome/ (2097)
23 hypersensit$.ti,ab. (29136)
24 harm$.ti,ab. (44607)
25 drug hypersensitivity/ (21245)
26 OR/9-25 (771311)
27 8 and 26 (2793)
28 etanercept/ae, to [adverse drug reaction, drug toxicity] (2375)
29 infliximab/ae, to [adverse drug reaction, drug toxicity] (3859)
30 adalimumab/ae, to [adverse drug reaction, drug toxicity] (1283)
31 golimumab/ae, to [adverse drug reaction, drug toxicity] (40)
32 certolizumab/ae, to [adverse drug reaction, drug toxicity] (67)
33 rituximab/ae, to [adverse drug reaction, drug toxicity] (2981)
34 OR/28-33 (5104)
35 26 or 34 (5888)
36 Urinary tract infection/si [side effects] (2320)
37 Lower respiratory tract infection/si [side effects] (172)
38 skin infection/si [side effects] (547)
39 bone infection/si [side effects] (30)
40 infectious arthritis/si [side effects] (64)
41 neoplasm/si [side effects] (549)
42 tuberculosis/si [side effects] (1406)
43 OR/36-42 (4644)
44 35 and 43 (36)

8.8.5 Details of any additional searches (for example, searches of company databases [include a description of each database]).

No additional searches were conducted.

8.8.6 The inclusion and exclusion criteria.

Following inclusion and exclusion criteria were applied.

Study design

- Randomised controlled trials (RCTs) (including any open-label extensions of these RCTs)
- Non randomised trials only when the information was not available in RCTs

Interventions

- Etanercept
- Infliximab
- Adalimumab
- Golimumab
- Certolizumab
- Rituximab
- Abatacept
- Palliative care which included NSAIDs and DMARDs

Participants

Active RA with an inadequate response to previous standard therapy (including at least two DMARDs).

Outcomes

- Malignancies
- Severe infections (i.e. those that require IV antibiotic therapy and/or hospitalisation or cause death)
- Reactivation of latent tuberculosis.

8.8.7 The data abstraction strategy.

Please see section 9.2.7.

8.9 Appendix 9: Quality assessment of adverse event data in section 5.9 (Adverse events)

8.9.1 Please tabulate the quality assessment of each of the non-RCTs identified.

No studies specific to the adverse events of TNF-α inhibitors were identified.

8.10 Appendix 10: Search strategy for cost-effectiveness studies (section 6.1)

8.10.1 The specific databases searched and the service provider used

The following databases were searched to retrieve economic evaluations, data on costs and cost-effectiveness of biologics for the treatment of rheumatoid arthritis:

Table 173. Databases, date span & search strategy location for cost-effectiveness review

Database	Date Span	Search Strategy
OVID MEDLINE (R) In-Process & Other Non-indexed Citations	2000 – 2010 Week 3	Table 175
OVID EMBASE	2000 – 2010 Week 3	Table 176
NHS Economic Evaluation Database, NHS EED (Centre for Reviews & Dissemination, CRDWeb)	Unrestricted	Section 8.10.4
CRD Database of Abstracts of Reviews of Effectiveness (DARE)	Unrestricted	Section 8.10.4
CRD Health Technology Assessment (HTA) Database	Unrestricted	Section 8.10.4

Reasons for exclusion for the publications not assessed are presented in Table 174.

Table 174. Justification for excluded publications

First Author	Source	Year	Title
Does not include one of the comparators			
Alldred A	EMBASE	2001	Leflunomide: A novel DMARD for the treatment of rheumatoid arthritis
Clark W	HTA	2004	The clinical and cost-effectiveness of anakinra for the treatment of rheumatoid arthritis in adults: a systematic review and economic analysis
Moreland L	EMBASE	2006	Abatacept
Vera-Llonch M	NHS EED / EMBASE	2008	Cost-effectiveness of abatacept in patients with moderately to severely active rheumatoid arthritis and inadequate response to methotrexate
Non-investigational, no interest, not relevant			
Adis International	EMBASE	2005	Adalimumab is an effective option in the treatment of patients with refractory rheumatoid arthritis
Alldred A	EMBASE	2001	Etanercept in rheumatoid arthritis
Augustsson J	EMBASE	2006	Infliximab in the treatment of rheumatoid arthritis
Baker T	EMBASE	2003	A case study on rheumatoid arthritis
Bansback NJ	EMBASE	2005	NICE reappraisal of biologics in 2005: What rheumatologists need to know
Bombardier MK	EMBASE	2007	Quality of life in patients with rheumatoid arthritis: Does abatacept make a difference?
Brandt J	EMBASE	2005	Biologics: TNF alpha antagonists as therapeutic expansion in inflammatory rheumatic diseases
Braun J	EMBASE	2003	Overview of the use of the anti-TNF agent infliximab in chronic inflammatory diseases
Brennan A	MEDLINE	2008	Modelling the cost-effectiveness of TNF-alpha antagonists in the management of rheumatoid arthritis: results from the British Society for Rheumatology Biologics Registry (author reply)
Calabrese L	EMBASE	2006	The yin and yang of tumor necrosis factor inhibitors

Cole P	EMBASE	2004	The soluble tumor necrosis factor receptor etanercept: A new strategy for the treatment of autoimmune rheumatic disease
Cush J	EMBASE	2003	Rheumatoid Arthritis: Current Perspectives on DMARD Therapy
Gobelet CLF	EMBASE	2007	Work in inflammatory and degenerative joint diseases
Honeywell M	EMBASE	2007	Infliximab: A chimeric monoclonal antibody against tumor necrosis factor
Isenberg DA	EMBASE	2007	30 million, around 10 000[midline ellipsis] and 18 [midline ellipsis] Figuring out the optimal treatment for musculoskeletal conditions in the National Health Service
Kalden JR	EMBASE	2002	Expanding role of biologic agents in rheumatoid arthritis
Kavanaugh A	EMBASE	2005	Health economics: Implications for novel antirheumatic therapies
Keating GM	EMBASE	2002	Management of rheumatoid arthritis: Defining the role of etanercept
Le Loet X	EMBASE	2005	Rheumatoid arthritis in 2005: Which patients should receive TNFalpha antagonists and when? A point of view
Maetzel A	EMBASE	2005	Cost-effectiveness estimates reported for tumor necrosis factor blocking agents in rheumatoid arthritis refractory to methotrexate - A brief summary
Maini RN	EMBASE	2005	The 2005 International Symposium on Advances in Targeted Therapies: What have we learned in the 2000s and where are we going?
Martin Alcalde E	MEDLINE / EMBASE	2004	New perspectives in the treatment of rheumatoid arthritis
Meisler JG	EMBASE	2001	Toward optimal health: The experts discuss rheumatoid arthritis
Mittendorf T	EMBASE	2004	The meaning of quality of life with Rheumatoid Arthritis (RA).
Mody GM	EMBASE	2008	Challenges in the management of rheumatoid arthritis in developing countries
Moreland L	EMBASE	2005	Unmet needs in rheumatoid arthritis
Moreland LW	EMBASE	2004	Drugs That Block Tumour Necrosis Factor: Experience in Patients with Rheumatoid Arthritis
Moreland LW	EMBASE	2003	Early rheumatoid arthritis: A medical emergency?
Mousa SA	EMBASE	2007	Recent advances of TNF-alpha antagonists in rheumatoid arthritis and chronic heart failure

Nash PT	EMBASE	2005	Tumour necrosis factor inhibitors
O'Dell JR	EMBASE	2008	The Best way to treat early rheumatoid arthritis?
Pego-Reigosa	EMBASE	2008	Autoimmune diseases
Van De Putte L	EMBASE	2004	Adalimumab for Rheumatoid Arthritis: Considerations for Reimbursement by Third-Party Payors
Walsh CA	MEDLINE / NHS EED	2007	Quality of life and economic impact of switching from established infliximab therapy to adalimumab in patients with rheumatoid arthritis
Westhovens R	EMBASE	2008	Translating co-stimulation blockade into clinical practice
Winning A	EMBASE	2002	Infliximab in the treatment of rheumatoid arthritis
Wolfe F	EMBASE	2004	Measuring the efficacy and effectiveness of rheumatoid arthritis therapy: Time to change our thinking and adopt a new model
Wolfe F	EMBASE	2004	Do rheumatology cost-effectiveness analyses make sense?
Systematic review or meta-analysis			
Bansback N	EMBASE	2008	Economic evaluations in rheumatoid arthritis: A critical review of measures used to define health states
Bansback N	EMBASE	2005	A pharmacoeconomic review of adalimumab in the treatment of rheumatoid arthritis
Bansback NJ	EMBASE	2005	An overview of economic evaluations for drugs used in rheumatoid arthritis: Focus on tumour necrosis factor-alpha antagonists
Cole P	MEDLINE	2004	The soluble tumor necrosis factor receptor etanercept: a new strategy for the treatment of autoimmune rheumatic disease
Coyle D	HTA	2006	Infliximab and etanercept in patients with rheumatoid arthritis: a systematic review and economic evaluation
Cvetkovic RS	MEDLINE / EMBASE	2006	Adalimumab: A review of its use in adult patients with rheumatoid arthritis
Dhillon S	MEDLINE / EMBASE	2007	Etanercept: A review of its use in the management of rheumatoid arthritis

Doan QV	MEDLINE / EMBASE	2007	Review of eight pharmacoeconomic studies of the value of biologic DMARDs (adalimumab, etanercept, and infliximab) in the management of rheumatoid arthritis
Emery P	EMBASE	2004	Review of Health Economics Modelling in Rheumatoid Arthr
Gladman DD	EMBASE	2008	Pharmacoeconomics of adalimumab for rheumatoid arthritis, psoriatic arthritis, ankylosing spondylitis and Crohn's disease
Homik JE	EMBASE	2004	An economic approach to health care
Keating GM	EMBASE	2002	Infliximab: An updated review of its use in Crohn's disease and rheumatoid arthritis
Lyseng-Williamson KA	EMBASE	2004	Infliximab: A Pharmacoeconomic Review of its Use in Rheumatoid Arthritis
Pichon Riviere A	HTA	2006	Etanercept, infliximab and adalimumab for the treatment of rheumatoid arthritis
Regier DA	EMBASE	2007	Cost-effectiveness of tumor necrosis factor-alpha antagonist in rheumatoid arthritis, psoriatic arthritis and ankylosing spondylitis
Sharma PK	MEDLINE	2004	Biologics in rheumatoid arthritis
Welsing PMJ	EMBASE	2005	Quality of life and costs for different treatment strategies for rheumatoid arthritis
Different indication, disease			
Ancowitz B	EMBASE	2006	Infusion Services in the Gastroenterology Practice
Bansback NJ	NHS EED	2006	Estimating the cost and health status consequences of treatment with TNF antagonists in patients with psoriatic arthritis
Barnes T	EMBASE	2007	Targeting nanomedicines in the treatment of rheumatoid arthritis: Focus on certolizumab pegol
Bartalena L	EMBASE	2007	Immunotherapy for Grave's orbitopathy: Easy enthusiasm, but let's keep trying
Bell S	EMBASE	2000	Antibodies to tumour necrosis factor alpha as treatment for Crohn's disease
Brandt J	EMBASE	2002	Infliximab in the treatment of active and severe ankylosing spondylitis
Brown R	NHS EED	2007	Cost-effectiveness of omalizumab in patients with severe persistent allergic asthma
D'Haens G	EMBASE	2007	Risk and benefits of biologic therapy for inflammatory bowel diseas
Ghosh S	EMBASE	2006	Biological therapies in inflammatory bowel disease

Hanauer SB	EMBASE	2007	Turning traditional treatment strategies on their heads: Current evidence for "step-up" versus "top-down".
Kobelt G	NHS EED	2004	The burden of ankylosing spondylitis and the cost-effectiveness of treatment with infliximab (Remicade)
Marshall JK	DARE	2002	Clinical and economic assessment: infliximab for the treatment of Crohn's disease
Mease P	EMBASE	2006	Psoriatic arthritis update
Mossner R	EMBASE	2008	Tumor necrosis factor antagonists in the therapy of psoriasis
Reed MR	EMBASE	2008	Tumour necrosis factor inhibitors in ankylosing spondylitis
Yang C	EMBASE	2008	Management of spondyloarthropathy in Asian countries
Zochling J	EMBASE	2005	Management and treatment of ankylosing spondylitis
No economic outcomes			
Barton P	HTA / EMBASE	2004	The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis
Bejarno V	MEDLINE / EMBASE	2008	Effect of the early use of the anti-tumor necrosis factor adalimumab on the prevention of job loss in patients with early rheumatoid arthritis
Chou CT	EMBASE	2006	The clinical application of etanercept in Chinese patients with rheumatic diseases
Koberlein J	EMBASE	2008	Insufficient treatment with innovative therapies in rheumatoid arthritis
Kosinski M	MEDLINE	2002	Health-related quality of life in early rheumatoid arthritis: impact of disease and treatment response
Lubeck DP	EMBASE	2002	Health-related quality of life measurements and studies in rheumatoid arthritis
Marshall NJ	MEDLINE	2004	Patients' perceptioins of treatment with anti-TNF therapy for rheumatoid arthritis: a qualitative study
Rothschild BM	EMBASE	2008	Individual DMARDs have similar efficacy for RA, but combination therapy improves response
Juvenile population			
Culy CR	MEDLINE / EMBASE	2002	Etanercept: An updated review of its use in rheumatoid arthritis, psoriatic arthritis and juvenile rheumatoid arthritis

Cummins C	DARE / EMBASE	2002	A systematic review of effectiveness and economic evaluation of new drug treatments for juvenile idiopathic arthritis: etanercept
Reiff AO	EMBASE	2004	Developments in the treatment of juvenile arthritis
Cost study external to UK			
Brady B	HTA	2007	Long-term clinical and cost-effectiveness of infliximab and etanercept for rheumatoid arthritis
Sorensen J	EMBASE	2005	The case of tumour necrosis factor-alpha inhibitors in the treatment of rheumatoid arthritis: A budget impact analysis
Suarez-Almazor M	HTA	2007	Infliximab and etanercept in rheumatoid arthritis: systematic review of long-term clinical effectiveness, safety, and cost-effectiveness
Preliminary analysis			
Choi HK	NHS EED	2000	A cost-effectiveness analysis of treatment options for patients with methotrexate-resistant rheumatoid arthritis
Tanno M	NHS EED	2006	Modeling and cost-effectiveness analysis of etanercept in adults with rheumatoid arthritis in Japan: a preliminary analysis

8.10.2 The date on which the search was conducted.

21 January 2010.

8.10.3 The date span of the search.

Table 173 provides date spans for the conducted searches.

8.10.4 The complete search strategies used, including all the search terms: textwords (free text), subject index headings (for example, MeSH) and the relationship between the search terms (for example, Boolean).

The following search strategies were used for the individual bibliographic databases:

Table 175. MEDLINE (OVID) search strategy

No.	**Search term**	***Results***
1	RHEUMATOID ARTHRITIS.mp.	61476
2	TUMOR NECROSIS FACTOR.mp.	97824
3	ANTI-TNF.mp.	3916
4	ANTI-TNF-ALPHA.mp.	2181
5	ANTI-TUMOR NECROSIS FACTOR.mp.	1214
6	INFLIXIMAB.mp.	5102
7	REMICADE.mp.	173
8	ETANERCEPT.mp.	2342
9	ENBREL.mp.	142
10	ADALIMUMAB.mp.	1236
11	HUMIRA.mp.	71
12	GOLIMUMAB.mp.	19
13	or/2-5	98455
14	or/6-12	6919
15	1 and 13 and 14	1586
16	COST.mp.	229622
17	COST-EFFECTIVENESS.mp.	22988
18	COST-BENEFIT ANALYSIS.mp.	45914
19	COST UTILITY ANALYSIS.mp.	734
20	COST ESTIMATE.mp.	108
21	ECONOMIC EVALUATION.mp.	3284
22	HEALTH ECONOMIC.mp.	1050
23	ECONOMIC MODEL.mp.	595
24	ECONOMIC.mp.	87194
25	or/16-24	293501
26	15 and 25	86

27	QUALITY OF LIFE.mp.	116648
28	HEALTH STATUS.mp.	68448
29	HEALTH STATUS INDICATORS.mp.	13768
30	VALUE OF LIFE.mp.	5204
31	or/27-30	176565
32	26 and 31	20
33	limit 32 to yr=2000-2010	20
34	limit 33 to human	16

Of the 16 results, 6 fit within the inclusion and exclusion criteria.

Table 176. EMBASE (OVID) search strategy

No.	**Search term**	***Results***
1	RHEUMATOID ARTHRITIS.mp.	41488
2	TUMOR NECROSIS FACTOR.mp.	95255
3	ANTI-TNF.mp.	2477
4	ANTI-TNF-ALPHA.mp.	1376
5	ANTI-TUMOR NECROSIS FACTOR.mp.	883
6	INFLIXIMAB.mp.	10948
7	REMICADE.mp.	2191
8	ETANERCEPT.mp.	7510
9	ENBREL.mp.	1699
10	ADALIMUMAB.mp.	3482
11	HUMIRA.mp.	942
12	GOLIMUMAB.mp.	64
13	or/2-5	95476
14	or/6-12	13927
15	1 and 13 and 14	3232
16	COST.mp.	178054
17	COST-EFFECTIVENESS.mp.	51101
18	COST-BENEFIT ANALYSIS.mp.	24142
19	COST UTILITY ANALYSIS.mp.	2570
20	COST ESTIMATE.mp.	77
21	ECONOMIC EVALUATION.mp.	5754
22	HEALTH ECONOMIC.mp.	903
23	ECONOMIC MODEL.mp.	402
24	ECONOMIC.mp.	54730
25	or/16-24	212923
26	15 and 25	377
27	QUALITY OF LIFE.mp.	94623
28	HEALTH STATUS.mp.	34736
29	HEALTH STATUS INDICATORS.mp.	85
30	VALUE OF LIFE.mp.	65
31	or/27-30	122878

32	26 and 31	86
33	limit 32 to yr=2000-2010	85
34	limit 33 to human	84

Of the 84 results, 8 fit within the inclusion and exclusion criteria.

The Centre for Reviews & Dissemination, CRDWeb Databases (NHS EED, DARE, HTA) were searched with no date span restrictions using the following search strategy:

Rheumatoid Arthritis AND cost AND (quality of life OR health status OR health status indicators OR value of life) AND (tumour necrosis factor OR Anti-TNF OR infliximab OR Remicade OR Etanercept OR Enbrel OR Adalimumab OR Humira OR Golimumab)

The CRD NHS EED Database resulted in 16 hits, with 10 fitting within the search criteria.

The CRD DARE Database resulted in 4 hits, with 2 fitting within the search criteria.

The CRD HTA Database resulted in 9 hits, with 2 fitting with the search criteria.

8.10.5 Details of any additional searches (for example, searches of company databases [include a description of each database]).

No additional searched were conducted.

8.11 *Appendix 11: Quality assessment of cost-effectiveness studies (section 6.1)*

Economic Study	Bansback 2005	Barbieri 2005	Brennan 2007	Brennan 2004	Chen 2006	Chiou 2004	Choi 2002	Jobanputra 2002	Kobelt 2003	Kobelt 2004	Spalding 2006	Wailoo 2008	Wong 2002
1. Was the research question stated?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
2. Was the economic importance of the research question stated?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3. Was/were the viewpoint(s) of the analysis clearly stated and justified?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
4. Was a rationale reported for the choice of the alternative programmes or interventions compared?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	NC	Y	Y
5. Were the alternatives being compared clearly described?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
6. Was the form of economic evaluation stated?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
7. Was the choice of form of economic evaluation justified in relation to the questions addressed?	Y	N	N	N	N	Y	N	Y	Y	N	N	Y	N
8. Was/were the source(s) of effectiveness estimates used stated?	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
9. Were details of the design and results of the effectiveness study given (if based on a single study)?	Y	Y	Y	Y	NA	Y	Y	Y	Y	Y	Y	Y	Y
10. Were details of the methods of synthesis or meta-analysis of estimates given (if based on an overview of a number of effectiveness studies)?	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	N	Y
11. Were the primary outcome measure(s) for the economic evaluation clearly stated?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
12. Were the methods used to value health states and other benefits stated?	N	N	N	N	Y	Y	Y	Y	NC	Y	Y	Y	Y
13. Were the details of the subjects from whom valuations were obtained given?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
14. Were productivity changes (if included) reported separately?	Y	N	N	N	N	Y	N	Y	Y	N	N	Y	N
15. Was the relevance of productivity changes to the study question discussed?	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y

16. Were quantities of resources reported separately from their unit cost?	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
17. Were the methods for the estimation of quantities and unit costs described?	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	N	Y
18. Were currency and price data recorded?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
19. Were details of price adjustments for inflation or currency conversion given?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
20. Were details of any model used given?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
21. Was there a justification for the choice of model used and the key parameters on which it was based?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
22. Was the time horizon of cost and benefits stated?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
23. Was the discount rate stated?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
24. Was the choice of rate justified?	N	N	N	N	N	N	N	N	N	N	N	N	N
25. Was an explanation given if cost or benefits were not discounted?	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
26. Were the details of statistical test(s) and confidence intervals given for stochastic data?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
27. Was the approach to sensitivity analysis described?	Y	Y	Y	Y	N	N	Y	N	Y	Y	N	N	Y
28. Was the choice of variables for sensitivity analysis justified?	N	N	N	N	N	N	Y	N	N	N	N	N	N
29. Were the ranges over which the parameters were varied stated?	Y	Y	Y	Y	N	N	Y	N	Y	Y	N	N	Y
30. Were relevant alternatives compared? (That is, were appropriate comparisons made when conducting the incremental analysis?)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
31. Was an incremental analysis reported?	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	N	Y
32. Were major outcomes presented in a disaggregated as well as aggregated form?	N	N	N	N	N	N	Y	Y	N	N	N	Y	Y
33. Was the answer to the study question given?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
34. Did conclusions follow from the data reported?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
35. Were conclusions accompanied by the appropriate caveats?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
36. Were generalisability issues addressed?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

8.12 *Appendix 12: Search strategy for section 6.4 (Measurement and valuation of health effects)*

The algorithm from Hurst et al 1997 was applied within the modelling as it is one of the most widely referenced mapping equations and is described in a NICE report as addressing the structural issues of the BRAM (Barton 2004). For consistency with TA130, Hurst et al 1997 is applied within the base case.

8.12.1 The specific databases searched and the service provider used (for example, Dialog, DataStar, OVID, Silver Platter), including at least:

- Medline
- Embase
- Medline (R) In-Process
- NHS Economic Evaluation Database (NHS EED)
- EconLIT.

Not applicable.

8.12.2 The date on which the search was conducted.

Not applicable.

8.12.3 The date span of the search.

Not applicable.

8.12.4 The complete search strategies used, including all the search terms: textwords (free text), subject index headings (for example, MeSH) and the relationship between the search terms (for example, Boolean).

Not applicable.

8.12.5 Details of any additional searches (for example, searches of company databases [include a description of each database]).

Not applicable.

8.12.6 The inclusion and exclusion criteria.

Not applicable.

8.12.7 The data abstraction strategy.

Not applicable.

8.13 Appendix 13: Resource identification, measurement and valuation (section 6.5)

As discussed in Section 6.5 consultation with expert clinicians informed the resource use so that the model inputs were specifically tailored to current UK clinical practice.

8.13.1 The specific databases searched and the service provider used (for example, Dialog, DataStar, OVID, Silver Platter), including at least:

- Medline
- Embase
- Medline (R) In-Process
- NHS EED
- EconLIT.

Not applicable.

8.13.2 The date on which the search was conducted.

Not applicable.

8.13.3 The date span of the search.

Not applicable.

8.13.4 The complete search strategies used, including all the search terms: textwords (free text), subject index headings (for example, MeSH) and the relationship between the search terms (for example, Boolean).

Not applicable.

8.13.5 Details of any additional searches (for example, searches of company databases [include a description of each database]).

Not applicable.

8.13.6 The inclusion and exclusion criteria.

Not applicable.

8.13.7 The data abstraction strategy.

Not applicable.

8.14 Appendix 14: Eligibility criteria for RCTs (section 5.3)

Table 177. Eligibility criteria in the RCTs

Trial no. (acronym)	Inclusion criteria	Exclusion criteria
Kremer et al, 2003	-patients 18 to 65 years of age who met the ACR criteria for rheumatoid arthritis and were in functional class I, II, or III. -active disease, characterized by 10 or more swollen joints, 12 or more tender joints, and C-reactive protein levels of at least 1 mg per deciliter (upper limit of the normal range, 0.4). - Patients had to have been treated with methotrexate (10 to 30 mg weekly) for at least 6 months and to have received a stable dose for 28 days before enrollment. - All patients continued to receive methotrexate. - All other disease-modifying antirheumatic drugs were discontinued. - Leflunomide and infliximab were discontinued at least 60 days before enrollment, and other disease-modifying antirheumatic drugs were discontinued at least 28 days before enrollment. - Stable low-dose corticosteroids (≤10 mg per day) and nonsteroidal anti-inflammatory drugs were permitted.	- Women who were nursing or pregnant
Kremer et al, 2005	- Patients enrolled in this study met the ACR (formerly, the American Rheumatism Association) criteria for the diagnosis of RA (functional classes I, II, or III), had active disease defined by ≥10 swollen and ≥12 tender joints, had a C-reactive protein (CRP) level >1 mg/dl, and had been treated with MTX (10–30 mg/week) for at least 6 months with a stable dosage for 28 days prior to enrollment. - Patients were required to undergo a washout of all other DMARDs excluding MTX. Levels of corticosteroids were reduced to the equivalent of ≤10 mg/day prednisone and stabilized for ≥28 days prior to day 1.	- Pregnant or nursing women were excluded from the trial.
AIM Kremer et al, 2006	- Eligible patients were at least 18 years of age, had had rheumatoid arthritis for at least 1 year, and met the American Rheumatism Association criteria for rheumatoid arthritis. Rheumatoid arthritis was persistent and active despite methotrexate treatment. - All patients must have been treated with methotrexate (≤15 mg/wk) for 3 months or longer, with a stable dose for 28 days before enrollment. - Patients under went a washout of all other disease-modifying antirheumatic drugs at least 28 days before randomization. - Corticosteroid use was allowed, with dosages equal to 10 mg of prednisone or less per day, stabilized for 25 days before randomization. - At randomization, patients were required to have 10 or more swollen joints, 12 or more tender joints, and C-reactive protein levels of 10.0 mg/L or greater (normal range, 1.0 mg/L to 4.0 mg/L) while receiving methotrexate.	

	- Tuberculin skin testing was required before randomization. - Patients with a positive tuberculin skin test result were excluded unless they had completed treatment for latent tuberculosis before enrollment.	

ASSURE (Weinblatt et al, 2006)	- at least 18 years of age meeting the 1987 American College of Rheumatology (ACR; formerly, the American Rheumatism Association) criteria for the diagnosis of RA (7) and the 1991 ACR criteria for RA functional classes I, II, III, or IV. -Patients had to have active disease despite receiving background DMARDs and/or biologic therapy, warranting additional therapy at the discretion of the investigator. -To qualify for this study, the average score for the patient's global assessment of disease activity, as assessed by visual analog scale (VAS) measurements obtained at the time of screening and randomization (day 1), was required to be _20 mm - Patients were required to have been receiving ≥1 biologic and/or nonbiologic DMARD approved for RA for at least 3 months, and at a stable dose for at least 28 days prior to day 1 of the trial. -Patients with stable medical conditions such as congestive heart failure (CHF), asthma, chronic obstructive pulmonary disease (COPD), and diabetes mellitus were included.	- Patients were excluded if they had unstable or uncontrolled renal, endocrine, hepatic, hematologic, gastrointestinal, pulmonary, cardiac, or neurologic diseases, or any autoimmune disorder other than RA as the main diagnosis. - Other exclusion criteria were active or chronic recurrent bacterial infections unless treated and resolved, active herpes zoster infection within the previous 2 months, hepatitis B or hepatitis C virus infection, and active or latent tuberculosis (as assessed via chest radiography and tuberculin testing) unless appropriately treated. - Pregnant or nursing women were also excluded.
ATTAIN Genovese et al 2005	- Eligible patients met the American College of Rheumatology (ACR) criteria for rheumatoid arthritis, were at least 18 years of age, had had rheumatoid arthritis for at least one year, and had an inadequate response to anti–TNF*a* therapy with etanercept, infliximab, or both at the approved dose after at least three months of treatment. This study was initiated before the use of adalimumab became widespread. - Patients who had adverse events while receiving anti–TNF*a* therapy but who discontinued treatment primarily because of a lack of efficacy were also eligible. -Two groups of patients were enrolled: those receiving anti–TNF*a* therapy at the time of screening (current users) and those who had previously received such therapy (former users). - All users were required to stop taking etanercept or infliximab for at least 28 or 60 days, respectively, before undergoing randomization. -At randomization, patients had to have at least 10 swollen joints, at least 12 tender joints, and C-reactive protein levels of at least 1 mg per deciliter (upper limit of the normal range, 0.5). Patients had to have been taking an oral DMARD or anakinra for at least 3 months, and the dose had to have been stable for at least 28 days. - Use of oral corticosteroids (no more than 10 mg of prednisone or its equivalent per day) was allowed if the dose had been stable for at least 28 days.	

	-Changes in the doses of background DMARDs were not permitted except to avoid adverse effects.	

Kim et al 2007	-Patients were 18 years of age or older, met American College of Rheumatology (ACR) criteria for diagnosis of active RA, and had ≥6 swollen joints and ≥9 tender joints at both screening and baseline visits. -Patients had to have received at least one prior DMARD other than MTX but could have had efficacy failures to no more than four standard DMARDs other than MTX. - Patients had to have been treated with MTX for at least 6 months and been receiving a stable dosage for at least 4 weeks prior to screening. -Patients with acute inflammatory joint diseases other than RA were excluded, as were patients with active *Listeria* or tuberculosis infection; positive serology for human immunodeficiency virus antibody, hepatitis B surface antigen, or hepatitis C antibody; calcified granuloma and/or pleural scarring on chest radiograph. -Patients with positive RT23 2TU skin test (= 5 mm of induration) could be enrolled if receiving prophylactic isoniazid 300 mg daily at least 3 weeks prior to baseline.	
van de Putta et al 2004	- Patients aged 18 years or older -Patients met the diagnostic criteria for RA established by the American College of Rheumatology (ACR), - treatment with at least one DMARD had previously failed, - patients had active disease defined as >12 tender joints based on a 68 joint assessment, >10 swollen joints based on a 66 joint evaluation, and either an erythrocyte sedimentation rate (ESR) >28 mm/1st h or a serum C reactive protein (CRP) concentration >20 mg/l. -A negative pregnancy test and the use of a reliable contraceptive method were mandatory in women of childbearing potential.	- Exclusion criteria included joint surgery within 2 months before screening or infection requiring admission to hospital or treatment with intravenous (iv) antibiotics within 1 month before screening. - Patients were excluded if they had received treatment with either an intra-articular or intramuscular corticosteroid within 1 month before the study or an investigational small molecule drug or biological agent within 2 months or 6 months before screening, respectively. -Patients with impaired renal or hepatic function, or a history of tuberculosis as shown by radiographs, were excluded from the study. consent.
ARMADA Weinblatt et al 2003	-Eligible patients were 18 years of age or older and had RA that was diagnosed according to the 1987 revised criteria of the American College of Rheumatology (ACR; formerly, the American Rheumatism Association). -Active disease was defined as the presence of at least 9 tender joints (of 68 joints evaluated) and 6 swollen joints (of 66 joints evaluated). - participants must have been treated with MTX for a minimum of 6 months and must have been taking a stable weekly dose (12.5–25 mg, or 10 mg if intolerant to higher doses) for at least 4 weeks before entering the study. - All participants must have failed treatment with at least 1 DMARD besides MTX, but no more than 4 DMARDs.	- Exclusion criteria consisted of standard exclusion criteria used in trials of other biologics in patients with RA. - patients who had received treatment with anti-CD4 therapy or TNF_ antagonists, had a history of active listeriosis or mycobacterial infection, and had a major episode of infection requiring hospitalization or treatment with intravenous antibiotics within 30 days or oral antibiotics within 14 days prior to screening were also excluded.

CHANGE Miyaska 2008	- patients aged 20 years or older - Eligible patients met the American College of Rheumatology (ACR) criteria for active RA, had failed treatment with at least one prior disease-modifying antirheumatic drug (DMARD), and had C10 swollen joints and C12 tender joints (excluding distal interphalangeal joints) at both the screening visit and baseline visit. - Patients also had a C-reactive protein (CRP) concentration C2 mg/dl. - Patients taking a DMARD, including MTX, must have discontinued DMARDs at least 28 days prior to study drug administration and returned for baseline visit within 42 days. - Use of a live vaccine within 3 months; treatment with an investigational biologic, including anti-CD4 antibody, within 6 months; or prior treatment with any TNF antagonist or an alkylating agent was not permitted. - Patients with a positive (C5 mm of induration), but not strongly positive, tuberculin skin test could be enrolled if receiving prophylactic isoniazid 300 mg daily at least three weeks prior to baseline. - A negative pregnancy test and use of reliable contraception were mandatory for women of childbearing potential.	- Exclusion criteria included patients with acute inflammatory joint diseases other than RA, active Listeria or tuberculosis, lymphoma, or leukemia, or any malignancy except for successfully treated nonstatic basal-cell carcinoma of the skin. - Patients with positive serology for anti-human immunodeficiency virus antibody, hepatitis B virus surface antigen, or anti-hepatitis C virus antibody, ongoing or active infection, advanced or poorly controlled diabetes, or central nervous system demyelinating disorders were also excluded. - Patients with positive chest X-ray or strongly positive tuberculin skin test (C10 mm diameter of erythema and double redness/bullae/necrosis) could not be enrolled.
Keystone et al 2004 DE019	-18 years of age or older, had active RA diagnosed according to the 1987 revised American College of Rheumatology (ACR; formerly, American Rheumatism Association) criteria (10), and had ≥9 tender joints (of 68 evaluated), ≥6 swollen joints (of 66 evaluated), a C-reactive protein concentration ≥1 mg/dl, and either rheumatoid factor positivity or at least 1 joint erosion on radiographs of the hands and feet. -been on MTX therapy for ≥3 months at a stable dose of 12.5–25 mg/week (or ≥10 mg/week in patients intolerant to MTX) for ≥4 weeks.	Major exclusion criteria consisted of -prior use of anti-CD4 antibody therapy or TNF antagonists, a history of an active inflammatory arthritide other than RA, -a history of active listeriosis or mycobacterial infection, a history of lymphoma or leukemia or other malignancy besides nonmelanoma skin cancer within 5 years. -a major episode of infection (i.e., infections requiring hospitalization, treatment with intravenous antibiotics within 30 days prior to screening, or oral antibiotics within 14 days prior to screening), any uncontrolled medical condition, and pregnancy or breastfeeding.
STAR Furst et al 2003	- Eligible patients were 18 years of age or older, had active RA at both screening and baseline visits defined by at least 6 swollen joints and at least 9 tender joints (excluding distal interphalangeal joints), and met the 1987 revised American College of Rheumatology (ACR) criteria9 for diagnosis of RA for at least 3 months.	-Exclusion criteria consisted of those used in trials of other biologic DMARD in RA. - patients treated with anti-CD4 therapy or biologic DMARD (e.g., TNF antagonists, interleukin-1 receptor antagonists) and/or with a history of an active inflammatory arthritide other than RA, a history of active listeriosis or mycobacterial infection, a major episode of infection (i.e., infections requiring hospitalization, treatment with intravenous antibiotics within 30 days prior to screening, or oral antibiotics within 14 days prior to screening), and any uncontrolled medical condition. Patients experienced a variety of comorbid diseases.
Chen et al 2009	-Patients had to fulfill the American College of Rheumatology (ACR)	- Patients were excluded if they had received any of the following: TNF-α

	1987 revised criteria for RA and with a disease duration of more than 1 year were enrolled. -Active disease was defined as having more than six swollen joints and nine tender joints for at least 3 months. Disease activity for each RA patient was also assessed by the 28-joint disease activity score (DAS28) - Those who had been receiving MTX at a stable dose of 10–15 mg weekly and had been washed clean from other DMARDs for at least 4 weeks prior to entry were considered eligible for participation in this study.	inhibitors including adalimumab; alkylating agents such as chlorambucil or cyclophosphamide; investigational biological agents including anti-CD4 antibody; other investigational agents within 30 days; or live vaccine within 3 months prior to study. - Other criteria for exclusion were as follows: patients with clinically active tuberculosis (TB) or radiographic evidence of old pulmonary TB; patients with renal (serum creatinine > 1.5 mg/dL) and hepatic impairment (alanine aminotransferase, aspartate transaminase values > 2 times the upper limit of normal range); total bilirubin level >3mg/dL; hemoglobin < 9.5 mg/dL for men and < 9.0 mg/dL for women; platelet count < 150,000/mm3; white blood cell count < 3000 cells/mm3; pregnant and nursing mothers; patients with a history of alcohol and drug abuse; patients with positive serology for human immunodeficiency virus antibody, hepatitis B surface antigen or hepatitis C antibody; history of another collagen–vascular disease; preexisting or recent onset of central nervous system demyelinating disorders; patients with significant medical diseases including uncompensated congestive heart failure, severe myocardial infarction within 1 year, uncontrolled hypertension, poorly controlled diabetes mellitus, and chronic or active infection; and patients with any condition that might cause their participation in this study to be detrimental, as judged by a physician. In addition, concomitant use of hydroxychloroquine, sulfasalazine, azathioprine, cyclophosphamide, minocycline, mycophenolate mofetil, other DMARDs or any other investigational drug was prohibited during this study. However, nonsteroidal anti-inflammatory drugs, oral corticosteroids, MTX, and aspirin concomitantly used during the study were allowed as long as the dose was maintained throughout the study.

RAPID 1 Keystone et al 2008b	-patients had to be at least 18 years of age and had a diagnosis of RA, as defined by the American College of Rheumatology (ACR; formerly, the American Rheumatism Association) 1987 criteria for ≥6 months prior to screening but for <15 years. - Active disease was defined as ≥9 tender and 9 swollen joints at screening and at baseline, with either an erythrocyte sedimentation rate (ESR; Westergren) ≥30 mm/hour or a C-reactive protein (CRP) level≥15 mg/liter. -Patients were required to have received MTX for ≥6 months, with a stable dosage of ≥10 mg/week for ≥2 months prior to baseline.	- Exclusion criteria consisted of diagnoses of any other inflammatory arthritis or a secondary noninflammatory arthritis that could have interfered with our evaluation of the effects of certolizumab pegol on RA. - Patients with a history of tuberculosis or a chest radiograph showing active or latent tuberculosis were also excluded. - Patients with positive findings on a purified protein derivative (PPD) skin test were excluded, unless the PPD positivity was associated with previous vaccination with BCG (PPD positive by local standard). - If there was no clinical or radiographic suspicion of tuberculosis in these latter patients, they were enrolled at the discretion of the investigator. - Patients who, in the investigator's opinion, were at a high risk of infection were excluded, as were patients who had a history of malignancy, demyelinating disease, blood dyscrasias, or severe, progressive, and/or uncontrolled renal, hepatic, hematologic, gastrointestinal, endocrine, pulmonary,cardiac, neurologic, or cerebral disease. Patients who had received any biologic therapy within 6 months (or had received etanercept and/or anakinra within 3 months) of baseline and/or any previous biologic therapy that resulted in a severe hypersensitivity or anaphylactic reaction were excluded, as were patients who had previously failed to respond to treatment with an anti-TNF agent.
RAPID 1 Strand et al 2009	As above RAPID 1	As above RAPID 1
RAPID 2 Smolen et al 2009b	- The full inclusion and exclusion criteria are available online as supplementary material. - Eligible patients were aged >18 years with a diagnosis of RA, defined by American College of Rheumatology (ACR) 1987 criteria,19 of >6 months' duration but not longer than 15 years, with active disease at screening and baseline. - Patients had to have received prior MTX for >6 months (stable dose >10 mg/week for >2 months before baseline).	- Patients were excluded if they had received any biological agent for RA within 6 months before enrolment (3 months for etanercept and anakinra), had received previous treatment with a biological agent resulting in a severe hypersensitivity or anaphylactic reaction, or had not initially responded to previous anti-TNF therapy. Oral corticosteroids ((10 mg/day prednisone equivalent) and non-steroidal anti-inflammatory drugs and cyclo-oxygenase-2 inhibitors were permitted provided that the doses were stable within 28 and 14 days of baseline, respectively and remained stable during the study. -Patients with history of, or positive chest x-ray findings for, tuberculosis, or a positive purified protein derivative (PPD) skin test (defined as positive indurations per local medical practice) were excluded. As per protocol, if a positive PPD skin test was assumed by the local investigators to be related to previous bacille Calmette–Gue´rin (BCG) vaccination and was not associated with clinical or radiographic suspicion of tuberculosis,

		patients could be enrolled at the discretion of the investigator. In total, 101 patients (16%) were enrolled with a PPD test >5 mm at baseline.

TEMPO Kleskog et al 2004	- Eligible patients were those aged 18 years or older with disease duration of 6 months to 20 years who had active, adult-onset rheumatoid arthritis (American College of Rheumatology [ACR] functional class I–III), defined as ten or more swollen and 12 or more painful joints and at least one of the following: erythrocyte sedimentation rate 28 mm/h or greater; plasma C-reactive protein 20 mg/L or greater; or morning stiffness for 45 min or more. - Eligible patients should also have had a less than satisfactory response at the discretion of the investigator to at least one disease-modifying antirheumatic drug other than methotrexate. - Individuals previously treated with methotrexate could be enrolled provided they had not had clinically important toxic effects or lack of response, at the discretion of the investigator, and had not been treated with methotrexate within 6 months of enrolment.	- Patients were ineligible if they had previously received etanercept or other TNF antagonists. - Other exclusion criteria included previous treatment with immunosuppressive drugs within 6 months of screening; use of any investigational drug or biological agent within 3 months of screening; any other disease-modifying antirheumatic drug or corticosteroid injection within 4 weeks of baseline visit; and presence of relevant comorbidity, including active infections.
TEMPO van der Heijde et al 2006b	The TEMPO study enrolled patients of at least 18 years of age who had active RA (ACR functional class I–III) with disease duration of 6 months to 20 years. Active disease was defined as having 10 or more swollen joints, 12 or more painful joints, and at least one of the following: erythrocyte sedimentation rate >28 mm/1st h, C reactive protein >20 mg/l, or morning stiffness for >45 minutes. - Patients were also required to have experienced an unsatisfactory response to at least one DMARD other than MTX. All enrolees were considered suitable candidates for MTX treatment, had never had an unsatisfactory response to MTX, and had not received MTX in the 6 months before enrolment.	- Patients were ineligible if they had been treated with any DMARD within 4 weeks before the study baseline, or if they had ever received etanercept or another tumour necrosis factor antagonist. Patients were also excluded if they had received recent treatment with investigational, immunosuppressive, or corticosteroid drugs, or had significant concurrent disease.
Combe et al 2006	- Eligible patients were >18 years of age with disease duration (20 years who had active adult-onset rheumatoid arthritis (functional class I–III),16 defined as >6 swollen and >10 painful joints, and at least one of the following: erythrocyte sedimentation rate (ESR) >28 mm at the end of the first hour, serum C reactive protein (CRP) >20 mg/l or morning stiffness >45 min. -Patients must have received stable doses of sulfasalazine (2–3 g daily) for >4 months before screening, without signs of toxicity.	Patients were ineligible if -they had received etanercept or other TNF antagonists or - if they had received a DMARD other than sulfasalazine within 3 months before baseline. -Other exclusion criteria included the use of any immunosuppressive biological agents or cyclophosphamide within 6 months before screening, parenteral corticosteroids within 4 weeks before screening, and the presence of relevant comorbidity, including active infections. Patients were permitted stable doses of oral corticosteroids ((10 mg/day of prednisone or equivalent), one non-steroidal anti-inflammatory drug, simple analgesics with no anti-inflammatory action or daily doses of aspirin ((300 mg)

		during the study. -Patients with diseases that included cancer, congestive heart failure, uncontrolled hypertension, severe pulmonary disease, leucopenia, renal disease, thrombocytopenia or connective tissue disorders other than rheumatoid arthritis were not included. -Pregnant or breastfeeding women were also not included.

Moreland et al 999	- Eligible patients were adults who were at least 18 years of age, met the American Rheumatism Association's diagnostic criteria for rheumatoid arthritis, and were in functional class I, II, or III. - Patients were required to have had an inadequate response to one to four DMARDs (such as azathioprine, methotrexate, sulfasalazine, penicillamine, hydroxychloroquine, or oral or injectable gold); an inadequate response was defined as discontinuation of therapy because of lack of effect. - If a patients were receiving DMARDs, they were required to complete a DMARD washout period that lasted at least 1 month before starting study drug treatment; no DMARDs were permitted during the study. -Patients had to have active disease at enrollment (before the DMARD washout period), defined as 12 or more tender joints, 10 or more swollen joints, and at least one of the following: erythrocyte sedimentation rate of at least 28 mm/h, C-reactive protein level greater than 20 mg/L, or morning stiffness for at least 45 minutes. -All patients were required to have aminotransferase levels no greater than twice the upper limit of normal, a hemoglobin level of 85 g/dL or greater, a platelet count of at least 125 000 cells/ mm3, a leukocyte count of 3500 cells/mm3 or higher, and a serum creatinine level of 176.8 mmol/L (2 mg/dL) or less. -Concomitant therapy with stable doses of oral corticosteroids and nonsteroidal anti-inflammatory drugs (NSAIDs) was permitted. -Corticosteroid doses could not exceed the equivalent of 10 mg of prednisone per day, and NSAID doses could not exceed the maximum dose recommended by the manufacturer. -Patients could receive analgesics during the study except for the 24 hours before scheduled joint examinations. Intra-articular corticosteroids were not permitted during the study or beginning 4 weeks before enrollment. Because the inclusion criteria for this study were similar to the inclusion criteria for the previous 3-month trial in rheumatoid arthritis, 8 patients who received placebo in the 3-month trial were enrolled in the current study.	
Weinblatt et al 999	-Eligible patients were at least 18 years of age and fulfilled the 1987 criteria for rheumatoid arthritis of the American Rheumatism Association ; were in functional class I, II, or III according to the revised criteria of the American College of Rheumatology (ACR); and had active disease, as manifested by at least six joints that were swollen and six that were tender at the time of enrollment. - Before receiving the study drugs, all the patients had been	

	taking methotrexate for at least six months, and at a stable dose of 15 to 25 mg per week for the last four weeks (weekly doses as low as 10 mg were acceptable for patients who could not tolerate higher doses).	

GO AFTER Smolen et al 2009a	Eligible patients were aged 18 years or older, and had been diagnosed with active rheumatoid arthritis (persistent disease activity with at least four swollen and four tender joints), according to the criteria of the American College of Rheumatology (ACR) at least 3 months before screening. Patients must have been treated with at least one dose of a TNFα inhibitor (etanercept, adalimumab, or infl iximab), the last dose of which must have been given at least 8 weeks (adalimumabor etanercept) or 12 weeks (infliximab) before the fi rstdose of the study drug. Previous treatment with the TNFα inhibitor could have been discontinued for any reason. Investigators were asked to categorise the reasons for discontinuation as lack of eff ectiveness, intolerance, orother. A text box was available for the Investigator to specify if other was selected. Inconvenience and accessibility issues were the most common entries in the text box. Concomitant disease-modifying anti-rheumatic drug (DMARD) treatment with methotrexate, sulfasalazine, and hydroxychloroquine (alone or in combination) was permitted but not required. Patients receiving such drugs must have tolerated the dose for at least 12 weeks, and the dose must have been stable for 4 weeks before the fi rst dose of study drug. Patients who were receiving methotrexate, sulfasalazine, or hydroxychloroquine at baseline were allowed to discontinue these drugs before starting the study. However, if they continued these drugs, the dose had to be maintained throughout the study. Oral corticosteroids (not exceeding the equivalent of 10 mg of prednisone per day) or non-steroidal anti-infl amatory drugs were also allowed if the doses had been stable for at least 2 weeks before the first dose of study drug.	Patients were ineligible if they had inflammatory diseases other than rheumatoid arthritis; had a serious adverse reaction to a previous TNFα inhibitor (judged by the investigator); had ever received natalizumab or rituximab; had received anakinra less than 4 weeks, or alefacept or efalizumab less than 3 months before the first dose of study drug; had ever received cytotoxic drugs; had a history of latent or active granulomatous infection, except latent tuberculosis, that was treated prophylactically in the past 3 years; had a BCG vaccination less than 12 months before screening; had an opportunistic infection less than 6 months before screening; had a serious infection (judged by the investigator) less than 2 months before screening; had a history of chronic infection, demyelinating disease, congestive heart failure, or severe, progressive, uncontrolled renal, hepatic, haematological, gastro intestinal, endocrine, pulmonary, cardiac, neurological, psychiatric, or cerebral disease; or had a transplanted organ or a malignancy in the past 5 years.
GO FORWARD Keystone et al 2009b	-18 years of age or older, had a diagnosis of RA according to the revised 1987 criteria of the American College of Rheumatology (ACR) for at least 3 months before screening, and were to have been on a stable methotrexate dose of 15 mg/week or greater but 25 mg/week or less during the 4-week period immediately preceding screening. -Patients were to have tolerated 15 mg/ week or greater of methotrexate for at least3 months before screening. -Patients were required to have active RA, defined as four of more swollen joints (out of 66 total) and four or more tender joints (out of 68 total) and at least two of the following: (1) C-reactive protein (CRP) of 1.5 mg/dl or greater (normal range 0–0.6 mg/dl) or erythrocyte sedimentation rate (ESR) by the Westergren method of 28 mm/h or greater; (2) at least 30 minutes of morning stiffness; (3) bone erosion determined by x ray and/or magnetic resonance imaging; or (4) anticyclic citrullinated peptide antibody	- Patients were excluded from study participation if they had a known hypersensitivity to human immunoglobulin proteins or other components of golimumab. -Any previous use of any anti-TNF agent, rituximab, natalizumab or cytotoxic agents excluded patients from study participation. -Patients should not have received anakinra; disease-modifying antirheumatic drugs other than methotrexate; or intravenous, intramuscular, or intra-articular corticosteroids within 4 weeks before the first dose of study agent or alefacept or efalizumab within 3 months before the first dose of the study agent. A complete list of inclusion and exclusion criteria is provided in supplemental material 1 (available online only).

	or rheumatoid factor positive test results. -Eligible patients had to have met the tuberculosis screening criteria (supplemental material 1, available online only). Patients who were using non-steroidal anti-inflammatory drugs or other analgesics for RA had to be taking a stable dose for at least 2 weeks before the first dose of study agent. Patients who were taking oral corticosteroids had to have been receiving a stable dose equivalent to 10 mg/day or less of prednisone for at least 2 weeks before the first dose of study agent.	

Kay et al 2008	-Adult patients who had active RA for at least 3 months before screening were eligible for the study. -Active RA was defined by the 1987 revised criteria of the American College of Rheumatology (ACR; formerly, the American Rheumatism Association). -Patients were considered to have active RA if they demonstrated persistent disease activity despite receiving MTX at a stable dosage of at least 10 mg/week. -Persistent disease activity was defined as _6 swollen joints and _6 tender joints and at least 2 of the following 3 criteria: C-reactive protein (CRP) level _1.5 mg/dl, erythrocyte sedimentation rate (ESR) of _28 mm in the first hour according to the Westergren method, and morning stiffness of _30 minutes. - Patients had to have been treated with MTX at a dosage of at least 10 mg/week for _3 months and at a stable dosage for _4 weeks before receiving their first dose of study medication. -Patients were allowed to receive oral corticosteroids at a dosage not exceeding the equivalent of 10 mg of prednisone per day and could also take commercially available nonsteroidal antiinflammatory drugs (NSAIDs). -The dose of each was required to be stable for 4 weeks before the patient entered the study and was required to be held stable during the study. -If patients were not receiving corticosteroids or NSAIDs at the start of the study, they were not allowed to receive them during the study. All patients were required to receive folic acid supplementation at a stable dosage of at least 5 mg every week for at least 4 weeks before the first dose of study medication. -Patients were required to have laboratory test results within the following ranges at screening: hemoglobin _8.5 gm/dl, white blood cells _3,000/_l, neutrophils _1,500/_l, platelets _100,000/_l, serum transaminase levels not exceeding 1.5 times the upper limit of normal, and serum creatinine not exceeding 1.5 mg/dl.	-All patients underwent screening for tuberculosis. Patients were excluded from the study if they had evidence of prior or currently active tuberculosis by chest radiography, a history or current evidence of latent or active tuberculosis, recent close contact with an individual with active tuberculosis, or a positive tuberculin skin test result (induration of _5 mm or according to local guidelines). -Chest radiographs for tuberculosis screening were required to have been performed within 3 months before the first dose of study medication, and tuberculin skin tests had to have been performed within 1 month before the first dose. mrw.interscience.wiley.com/suppmat/0004-3591/suppmat/). Additional exclusion criteria are provided elsewhere (see Table 1, available on the *Arthritis & Rheumatism* Web site at http://www.
ATTEST Schiff et al 2008	-Eligible patients met the American College of Rheumatology (ACR) criteria for RA, were at least 18 years of age, had RA for at least 1 year,4 and had an inadequate response to MTX, as demonstrated by ongoing active disease (at randomisation >10 swollen joints, >12 tender joints, and C-reactive protein (CRP) levels >1 mg/dl using a high sensitivity assay (upper limit of the normal range, 0.5)).	

	-All patients had received MTX >15 mg/week for >3 months prior to randomisation (stable for at least 28 days) and washed out all DMARDs (>28 days prior) except for MTX. -No prior experience of abatacept or anti-TNF therapy was permitted. -All patients were screened for tuberculosis (TB) by purified protein derivative (PPD) testing and chest x ray. The protocol used for TB screening was the same as that employed in the 'Anti-TNF Trial in rheumatoid arthritis with Concomitant Therapy'' (ATTRACT) trial.5 Extended report 1096 Ann Rheum Dis 2008;67:1096–1103. doi:10.1136/ard.2007.080002	

ATTRACT Lipsky et al 2000	The eligibility criteria and the design of the study have been described in detail elsewhere (see below: ATTRACT Maini et al 1999).	
ATTRACT Maini et al 1999	-Patients were eligible if they had been diagnosed with rheumatoid arthritis according to the 1987 American College of Rheumatology criteria and had evidence of active disease despite treatment with methotrexate (six or more swollen and tender joints plus two of: morning stiffness greater than or equal to 45 min, erythrocyte sedimentation rate greater than 28 mm/h, C-reactive protein greater than 2 mg/dL. -The patients were classified into a functional class (American College of Rheumatology criteria). -Patients must also have been receiving oral or parenteral methotrexate for at least 3 months with no break in treatment of more than 2 weeks during this period. -The methotrexate dose must have been stable at 12·5 mg/week or more, for at least 4 weeks before screening and the patient must have been on a stable dose of folic acid for the same period. -Patients using oral corticosteroids (10 mg/kg or less prednisone equivalent) or non-steroidal anti-inflammatory drugs (NSAIDs) must have been on a stable dose for at least 4 weeks before screening: if a patient was not using such drugs, the patient must not have received either drug for at least 4 weeks before creening. The screening laboratory tests must have met the following criteria: haemoglobin 5·3 mmol/L or more, white blood cells 3·5_109/L or more, neutrophils 1·5_109/L, platelets 100_109/L or more, serum aminotransferase and alkaline phosphatase concentration 2 times or less the upper limit of normal, and serum creatinine 150 _mol/L or less.	- Patients were excluded if they had little or no ability for self-care; any current inflammatory condition with signs and symptoms that might confound the diagnosis (eg, connective tissue disease or Lyme disease); used a DMARD other than methotrexate or received intraarticular, intramuscular, or intravenous corticosteroids in the 4 weeks before screening; received any other agent to reduce tumour necrosis factor or had any previous use of cyclophosphamide, nitrogen mustard, chlorambucil, or other alkylating agents; or a history of known allergies to murine proteins. - Patients were also excluded if they had had infected joint prosthesis during the previous 5 years; serious infections, such as hepatitis, pneumonia, pyelonephritis in the previous 3 months; any chronic infectious disease such as renal infection, chest infection with bronchiectasis or sinusitis; active tuberculosis requiring treatment within the previous 3 years; opportunistic infections such as herpes zoster within the previous 2 months; any evidence of active cytomegalovirus; active *Pneumocystis carinii*; or drug-resistant atypical mycobacterial nfection. - Other contraindications for inclusion were: current signs or symptoms of severe, progressive, or uncontrolled renal, hepatic, haematological, gastrointestinal, endocrine, pulmonary, cardiac, neurological, or cerebral disease; a history of lymphoproliferative disease including lymphoma or signs suggestive of disease, such as lymphadenopathy of unusual size or location (ie, lymph nodes in the posterior triangle of the neck, infraclavicular epitrochlear, or periaortic areas); splenomegaly; any known malignant disease except basal cell carcinoma currently or in the past 5 years.
START Westhovens et al 2006b	- Adults were considered eligible for the study if they had a diagnosis of RA according to the revised criteria of the American College of Rheumatology (ACR; formerly, the American Rheumatism Association), and had active disease despite receiving MTX; patients may or may not have been treated with other concomitant DMARDs. - Active RA was defined as the presence of 6 swollen joints and 6 tender joints. At screening, patients were required to have a chest radiograph that showed no evidence of malignancy, infection, fibrosis, or active tuberculosis.	- Patients were excluded from the study if they had been treated with an investigational drug (within 3 months or 5 half-lives from the time of screening, whichever was greater), with cyclophosphamide, nitrogen mustard, chlorambucil, or other alkylating agents, with more than 5 mg/kg of cyclosporine, or with any approved or investigational biologic agent (including infliximab) at any time prior to the study, with the exception of approved vaccines for the purpose of immunization.

	- All patients must have been receiving MTX for at least 3 months prior to randomization. The MTX dose must have been stable for at least 4 weeks prior to randomization. -Patients were permitted to continue receiving stable doses (for at least 4 weeks prior to randomization) of ongoing antirheumatic therapy, including the following medications: chloroquine, azathioprine, penicillamine, oral or intramuscular gold, hydroxychloroquine, sulfasalazine, leflunomide, cyclosporine, oral corticosteroids, or nonsteroidal antiinflammatory drugs.	- Patients were excluded from the study if they had opportunistic infections, serious infections during the 2 months prior to screening, known human immunodeficiency virus infection, active tuberculosis or history of active tuberculosis with inadequate documentation of treatment, evidence of latent tuberculosis (according to a positive finding on the purified protein derivative of tuberculin [PPD] test in the US, and as defined by local guidelines outside the US) and an inability to receive prophylaxis with isoniazid, a history of lymphoproliferative disease or malignancy, or a diagnosis of congestive heart failure.

Abe et al 2006	- Eligible patients were 20–75 years of age and fulfilled the diagnostic criteria for RA of the American Rheumatism Association16 at least 6 months prior to enrollment. Patients were eligible for the DBT if they had ≥ 6 tender joints (of 68 counted) and ≥ 6 swollen joints (of 66 counted), plus at least 2 of the following: morning stiffness ≥ 45 min, erythrocyte sedimentation rate ≥ 28 mm/h, or C-reactive protein (CRP) ≥ 2 mg/dl, despite treatment with MTX for more than 3 months. The MTX dosage must have been stable 6 mg/week or more during the last 4 weeks. - Patients receiving oral or suppository nonsteroidal antiinflammatory drugs (NSAID), folic acid, oral or suppository corticosteroid (10 mg/day or less prednisolone equivalent) must have been taking a stable dose for 4 weeks prior to entry. -Patients were not allowed to use DMARD, immunosuppressive drugs other than MTX, intraarticular, intramuscular, intravenous or epidural corticosteroids, to have arthrocentesis and plasma exchange (for 4 wks prior to entry), or use alkylating agents (for 5 yrs prior to entry). -Patients were excluded if they had functional class IV using Steinbrocker's criteria, any other systemic rheumatic diseases except Sjögren's syndrome, serious infections, opportunistic infections (within the previous 3 mo), tuberculosis (within the previous 3 yrs), infections of artificial joints (within the previous 5 yrs), human immunodeficiency virus infection, malignancies (within the previous 5 yrs), a history of known allergies to human/murine chimeric antibodies, or pregnancy. Laboratory exclusion criteria were: hemoglobin < 8.5 g/dl; leukocyte count < 3500 ⊚ 106/l; neutrophil count < 1500 ⊚⊚106/l; platelet count < 10 ⊚⊚104/μl; serum creatinine level > 1.5 mg/dl; and alanine aminotransferase (ALT) levels, aspartate aminotransferase (AST) levels, and alkaline phosphatase (ALP) levels greater than twice the normal upper limit. Patients who completed all scheduled infusions and evaluations in the DBT and desired extended treatment with infliximab were enrolled in the OLT. infusions	
Maini et al 1998	-Patients who met the criteria for RA had taken MTX once a week at a dosage of 7.5-15 mgiweek for a minimum of 6 months.	- DMARLIs other than MTX, if any, were withdrawn at least 4 weeks before screening for trial eligibility and were not permitted during the 26 weeks of study

	- If the disease was still active, either due to an incomplete response to MTX or a disease flare, patients were considered for the trial. -No attempt was made to establish whether patients were nonrespondersto a "protocol-defined" higher dosage of MTX or were limited to a lower dosage because of toxicity.. -For 4 weeks prior to screening for study entry, patients taking oral corticosteroids (not exceeding the equivalent of 7.5 mg/day of prednisolone) were maintained on a stable dosage of these drugs, and all patients also took a fixed dosage of 7.5 mg of MTX weekly. After a 4-week stabilization period, patients were screened and were eligible for entry into the trial if on the day of screening, they had 26 swollen joints (of 66 counted) and at least 2 of the following: *2 6* tenderipainful joints (of 68 counted), >45 minutes of morning stiffness, and an ESR of >28 mmihour (Westergren) or a CRP level of >15 mgidl.	- Patients were excluded if they were pregnant, severely physically incapacitated (Steinbrocker class IV) (18), had previously been exposed to murine or chimeric **MAb,** or had a history of chronic infection, a recent serious infection, or a history of malignancy. -Laboratory exclusions were a hemoglobin level 4 . 5 gmidl, a white blood cell count <3.5 X 10"/liter, a platelet count <I00 x 1OY/Iiter, a serum creatinine level >150 p,molesiliter, serum transaminase levels 1.25 times the upper limit of normal, or alkaline phosphatase levels >2 times the upper limit of normal.

Edwards et al 2004	- Eligible patients were at least 21 years of age, fulfilled the revised 1987 American Rheumatism Association criteria, and had active disease despite treatment with at least 10 mg of methotrexate per week. -Active disease was defined by the presence of at least eight swollen and eight tender joints and at least two of the following: a serum C-reactive protein level of at least 15 mg per liter, an erythrocyte sedimentation rate of at least 28 mm per hour, or morning stiffness lasting longer than 45 minutes. - In addition, eligible patients were seropositive for rheumatoid factor, as defined by a plasma rheumatoid factor level of at least 20 IU per milliliter. -Patients were allowed to receive nonsteroidal antiinflammatory drugs at stable doses or corticosteroids at doses that did not exceed 12.5 mg per day of prednisolone (or the equivalent).	- Patients were excluded if they had an autoimmune disease other than rheumatoid arthritis (except concurrent Sjögren's syndrome), American Rheumatism Association functional class IV disease, active rheumatoid vasculitis, a history of systemic diseases associated with arthritis, chronic fatigue syndrome, serious and uncontrolled coexisting diseases, active infection, a history of recurrent clinically significant infection or of recurrent bacterial infections with encapsulated organisms, primary or secondary immunodeficiency, or a history of cancer (except basal-cell carcinoma of the skin that had been excised). -Concurrent treatment with any disease-modifying antirheumatic drug or any anti–tumor necrosis factor *a* therapy during the trial was prohibited.
Strand et al 2006	- Adult rheumatoid factor seropositive patients, with RA diagnosed by revised 1987 American Rheumatism Association criteria, who had failed 1–5 disease-modifying anti-rheumatic drugs (DMARDs) and had active disease despite ongoing treatment with methotrexate (_10 mg/week) for _16 weeks, were enrolled - Patients had to have a swollen joint count (SJC) _8, a tender joint count (TJC) _8 and at least two of the following: an elevated C-reactive protein (CRP) level (_1.5 mg/dl) and/or erythrocyte sedimentation rate (ESR) (_30 mm/h), and/or morning stiffness _45 min. -Patients were randomized in a double-blind fashion to one of the four treatment groups: continuing oral methotrexate (_10 mg/week)þplacebo rituximab; rituximab alone [1000 mg intravenous (iv) infusion days 1 and 15]; rituximabþcyclophosphamide (750mg iv on days 3 and 17); or rituximabþcontinuing methotrexate. -All patients received methylprednisolone 100 mg iv before infusions (rituximab or placebo) and oral prednisone for 2 weeks after the first infusion (total prednisone dose 510 mg).	
REFLEX Cohen et al 2006	-Patients had RA for at least 6 months, according to the ACR 1987 revised criteria, and had active disease, which was defined as _8 swollen joints (of 66 joints assessed) and _8 tender joints (of 68 joints assessed), a C-reactive protein (CRP) level _1.5 mg/dl or an erythrocyte sedimentation rate (ESR) _28 mm/hour, and radiographic evidence of at least 1 joint with a definite erosion attributable to RA, as determined by a central reading site (a centralized organization independent of the sponsors that provided blinded radiographic assessments). -Eligible patients had to be taking MTX (10–25 mg/week) for at least 12	

	weeks prior to screening, with the last 4 weeks at a stable dosage. Patients with a history of a rheumatic autoimmune disease other than RA (except secondary Sjo¨gren's syndrome), significant systemic involvement secondary to RA (vasculitis, pulmonary fibrosis, or Felty's syndrome), or ACR functional class IV disease were excluded. - each investigator reviewed their patients' current immunization status and, if required, provide appropriate vaccinations/boosters at least 4 weeks prior to enrollment in the trial. - Enrolled patients had experienced an inadequate response to previous or current treatment with the anti-TNF agents infliximab (_3 mg/kg; at least 4 infusions), adalimumab (40 mg every other week for _3 months), or etanercept (25 mg twice weekly for _3 months), or were intolerant to at least 1 administration of these agents. Patients discontinued etanercept for _4 weeks and infliximab or adalimumab for _8 weeks prior to randomization.	

REFLEX Keystone et al 2008a	As above (REFLEX)	As above (REFLEX)
OPTION Smolen et al 2008	- Adult patients with moderate to severe active rheumatoid arthritis (diagnosed according to American College of Rheumatology [ACR] criteria16) of more than 6 months' duration who had an inadequate response to methotrexate were recruited. -Active disease was defi ned by a swollen joint count of 6 or more plus a tender joint count of 8 or more and C-reactive protein (CRP) over 10 mg/L or ESR of 28 mm/h or more. - patients had to have received methotrexate for 12 weeks or longer before the start of the study (stable dose of 10–25 mg/week for 8 weeks or longer). - All other DMARDs were discontinued before the start of the study: lefl unomide for 12 weeks or more (or ≥4 weeks after 11 days of standard colestyramine washout), anakinra for 1 week or more, etanercept for 2 weeks or longer, and infl iximab or adalimumab for 8 weeks or longer. Oral glucocorticoids (≤10 mg/day prednisone or equivalent) and non-steroidal anti-infl ammatory drugs (NSAIDs) were permitted if doses were stable for 6 weeks or more before inclusion.	Main exclusion criteria were other autoimmune diseases or signifi cant systemic involvement secondary to rheumatoid arthritis (eg, vasculitis, pulmonary fibrosis, or Felty's syndrome), functional class IV rheumatoid arthritis, previous or current infl ammatory joint disease other than rheumatoid arthritis, currently active or previous recurrent bacterial, viral, fungal, or other infections including, but not limited to, tuberculosis and atypical mycobacterial disease, clinically signifi cant abnormalities on chest radiograph, hepatitis B and C, and recurrent herpes zoster. Investigators were encouraged to exclude potentially eligible patients if, in their judgment, they had a history of unacceptably frequent recurrent infections. Patients were also excluded if they had active liver disease, indicated by screening and baseline concentrations of alanine or aspartate aminotransferase of 1·5 times the upper limit of normal or more, or previous unsuccessful treatment with an anti-TNF agent (ie, lack of effi cacy or signifi cant safety issues; terminations due to cost or injection discomfort were not excluded).
RADIATE Emery et al 2008	- Patients 18 years of age and older with moderate to severe active RA and failure to respond or intolerance to one or more TNF antagonists within the past year were included. -Patients had active RA for 6 months or more, swollen joint count (SJC) of 6 or more, tender joint count (TJC) of 8 or more, and C-reactive protein (CRP) greater than 1.0 mg/dl or erythrocyte sedimentation rate (ESR) greater than 28 mm/h at baseline. -Patients discontinued etanercept (>2 weeks), infliximab or adalimumab (>8 weeks), leflunomide (>12 weeks) and all DMARD other than methotrexate before receiving study medication. -Patients had to be treated with methotrexate for 12 weeks or more before baseline (stable dose >8 weeks).	- Exclusion criteria included treatment with celldepleting agents, uncontrolled medical conditions, history of other inflammatory diseases or functional class 4 RA, history of malignancies or recurrent infections, primary or secondary immunodeficiency, haemoglobin less than 8.5 g/dl, leucopenia, neutropenia, thrombocytopenia, abnormal liver function, triglycerides greater than 10 mmol/l, or recognized active tuberculosis, hepatitis B, or hepatitis C.
SATORI Nishimoto et al 2009	- Eligible patients were between 20 and 75 years old, fulfilled the American college of Rheumatology (ACR; formerly, the American Rheumatism Association) 1987 revised criteria for the classification of RA , with disease duration of more than 6 months. -All candidates were treated with MTX 8 mg/week for at least 8 weeks until enrolment. They all had C6 tender joints (of 49 evaluated), C6 swollen joints (of 46 evaluated), ESR of C30 mm/h or CRP of C10 mg/l at enrolment.	-Patients were excluded if they had functional class IV using Steinbrocker's criteria, aspartate transaminase (AST), alanine transaminase (ALT) and serum creatinine C1.5-fold the upper limit of normal, were HBs antigen and/or HCV antibody positive, had pulmonary fibrosis or active pulmonary disease, a history of serious adverse drug reaction to MTX, concomitant pleural effusion, ascites, varicella infection, or were excessive users of alcohol on a regular basis.

	-An inadequate response to MTX was defined as the presence of active disease, as described above. - Patients were not allowed to have received prior anti-TNF agents or leflunomide (within 12 weeks prior to the first dose), plasma exchange therapy or surgical treatments (within 4 weeks prior to the first dose), DMARDs other than MTX or immunosuppressants (within 2 weeks prior to the first dose). - Oral corticosteroids (prednisolone, B10 mg/day) were allowed if the dosage had not been changed within 2 weeks. - Eligible patients had white blood cell counts C3.5 9 109/l, lymphocyte counts C0.5 9 109/l and platelet count of at least the lower limit of normal as defined by the respective local laboratory used.	-Patients were also excluded if they had significant cardiac, blood, respiratory system, neurologic, endocrine, renal, hepatic, or gastrointestinal disease, or had an active infection requiring medication within 4 weeks before the first dose or medical history of a serious allergic reaction. -Sexually active premenopausal women were required to have a negative urine pregnancy test at the entry to the study and to use effective contraception during the study period.
TOWARD Genovese et al 2008	-Patients aged at least 18 years with moderate-to severe RA of at lease 6 months' duration, diagnosed according to the American College of Rheumatology (ACR; formerly, the American Rheumatism Association) 1987 revised criteria for the classification of RA (21), with a swollen joint count (SJC) of _6, a tender joint count (TJC) of _8, and a C-reactive protein (CRP) level _1 mg/dl or an erythrocyte sedimentation rate (ESR) _28 mm/hour were enrolled. Eligible patients had received stable doses of permitted DMARDs (methotrexate, chloroquine, hydroxychloroquine, parenteral gold, sulfasalazine, azathioprine, and leflunomide) for _8 weeks prior to study entry. -Oral glucocorticoids (_10 mg/day prednisone or equivalent) and nonsteroidal antiinflammatory drugs (NSAIDs)/cyclooxygenase 2 inhibitors were permitted if the doses were stable for _6 weeks. - Patients who were unsuccessfully treated with an anti-TNF agent or were previously treated with any cell-depleting therapy were excluded. Tuberculosis screening was managed according to local practice.	
Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee		

Table 178 Primary and secondary outcomes of the RCTs

Trial no. (acronym)	Primary outcome(s) and measures	Reliability/ validity/ current use in clinical practice	Secondary outcome(s) and measures	Reliability/ validity/ current use in clinical practice.
Kremer et al., 2003	ACR20 response at 6 months	Well established endpoint and widely used in clinical practice.	- Secondary outcome measures were 50 percent improvement and 70 percent improvement according to ACR criteria (an ACR 50 response and an ACR 70 response, respectively). - Health-related quality of life was assessed at base line, 90 days, and 180 days with use of the Medical Outcomes Study 36-Item Short-Form General Health Survey (SF-36).	Well established endpoint and widely used in clinical practice.
Kremer et al., 2005	ACR20 response at 6 months	Well established endpoint and widely used in clinical practice.	Secondary end points were ACR50 and ACR70 responses and improvements in individual components of the ACR core data set. Pain and global assessment of disease activity (patient's and physician's) were evaluated using a 0–100-mm visual analog scale (VAS). The proportions of patients having low disease activity and experiencing remission were also determined by a post hoc analysis using the Disease Activity Score in 28 joints (DAS28), which assessed the number of swollen and tender joints, CRP levels, and the patient's global assessment of disease activity (as measured on a VAS).	Well established endpoint and widely used in clinical practice.
AIM Kremer et al, 2006	Our 3 primary objectives were to evaluate the -proportion of patients in each group with a 20% improvement in American College of Rheumatology (ACR)	Well established endpoint and widely used in clinical practice.	Secondary objectives included assessing: -ACR 50 and ACR 70 responses at 6 months -all ACR responses at 1 year	Well established endpoint and widely used in clinical practice.

	response crite- ria (ACR 20) at 6 months, - the proportion of patients in each group with clinically significant improvement (_0.3 unit) in the Health Assessment Questionnaire Disability Index (HAQ-DI) score at 1 year, - the radiographic progression of joint erosions (assessed by comparing changes from baseline in the Genant-modified Sharp score) at 1 year.		-the proportions of patients achieving a major clinical response and a protocol defined extended major clinical response at 1 year. - changes in disease activity by using the Disease Activity Score 28 (DAS28) (20, 21). -improvements in physical function over 1 year by using the HAQ-DI, which measures physical function during daily activities - changes in health-related quality of life by using the Medical Outcomes Study Short Form-36 Health Survey (SF-36)	

ASSURE (Weinblatt et al 2006)	- The primary end point of the ASSURE trial was to evaluate the safety of abatacept compared with placebo when added to a background of approved synthetic DMARDs and/or biologic DMARDs over the course of 1 year, in a blinded, randomized study.	Well established endpoint and widely used in clinical practice.	- Three patient-reported components of the ACR core data set were exploratory secondary efficacy objectives in this study. -Physical function was assessed using the Disability Index of the Health Assessment Questionnaire (HAQ). - Patient's global assessment of disease activity, patient's global assessment of pain, and physician's global assessment of disease activity were all assessed using a 100-mm VAS.	Well established endpoint and widely used in clinical practice.
ATTAIN Genovese et al 2005	There were two primary end points: -the proportion of patients with an ACR 20 response -proportion of patients with an improvement of at least 0.3 from baseline in the Health Assessment Questionnaire (HAQ) disability index (exceeding the minimal clinically important change of 0.22) at six months.	Well established endpoint and widely used in clinical practice.	Secondary objectives included - 50 percent and 70 percent improvement in the ACR response (ACR 50 and ACR 70, respectively) at six months. - Changes in disease activity were assessed with the use of the Disease Activity Score 28 (DAS28). - Clinical remission was defined by a DAS28 of less than 2.6, and a low level of disease activity was defined by a DAS28 of 3.2 or less. -The mean improvement in physical function at six months was based on the change from baseline in the HAQ disability index. -Changes from baseline in the health-related quality of life were assessed by the Medical Outcomes Study 36-Item Short-Form General Health Survey (SF-36) at six months.	Well established endpoint and widely used in clinical practice.

Kim et al 2007	The primary efficacy endpoint was the percentage of patients achieving an ACR20 response in the adalimumab group compared with the placebo group at week 24 in the intention-to-treat (ITT) population.	Well established endpoint and widely used in clinical practice.	Secondary efficacy parameters included the following: - the percentage of patients achieving an ACR50 response; - the percentage of patients achieving an ACR70 response; - the percentage of patients achieving improvement in individual ACR core components, including tender joint count, swollen joint count, the Physician's Global Assessment of Disease Activity, the Patient's Global Assessment of Disease Activity, the Patient's Global Assessment of Pain, Disability Index of the Korea Health Assessment Questionnaire (KHAQ), and C-reactive protein concentrations; and the percentage of patients reporting morning stiffness.	Well established endpoint and widely used in clinical practice.
van de Putta et al 2004	The primary efficacy end point was the rate of ACR20 response (>20% improvement in the ACR core criteria).	Well established endpoint and widely used in clinical practice.	Secondary efficacy end points included - the ACR50 and ACR70 response rates and improvements in ACR core components (patient global assessment of disease activity, physician global assessment of disease activity, patient assessment of pain, the Disability Index of the Health Assessment Questionnaire (HAQ DI), and serum levels of CRP).	Well established endpoint and widely used in clinical practice.
Weibnblatt et al 2003 ARMADA	The primary efficacy end point was the ACR20 response.	Well established endpoint and widely used in clinical practice.	Secondary efficacy end points included -the ACR50 and the ACR70 response rates -improvements in ACR core set of disease activity measures for RA clinical trials as follows: tender joint count, swollen joint count, patient's assessment of pain, patient's global assessment of disease activity, physician's global assessment of disease activity, the Disability Index of the Health Assessment Questionnaire (HAQ), and serum	Well established endpoint and widely used in clinical practice.

			levels of C-reactiveprotein. Other secondary efficacy end points were the score on the Short Form 36 (SF-36), which is a 36-item health survey, and the fatigue scale of the Functional Assessment of Chronic Illness Therapy (FACIT). Serum concentrations of the cartilage destruction markers pro–matrix metalloproteinase 1 (proMMP-1) and proMMP-3 were obtained during the study.	

Miayaska 2008 CHANGE	The primary efficacy endpoint was ACR20 response rate at Week 24 for the adalimumab 40 and 80 mg groups compared with placebo.	Well established endpoint and widely used in clinical practice.	The comparison between ACR20 response rates at Week 24 for the adalimumab 20 mg group and the placebo group was a secondary endpoint. The ACR components were evaluated at Weeks 0 (predose), 2, 4, 8, 12, 16, 20, and 24. - ACR20 response rate at Week 12; ACR50 and ACR70 response rates at Weeks 12 and 24; individual components of the ACR response (including TJC and SJC) at Weeks 0 (baseline), 12, and 24; and the Health Assessment Questionnaire Disability Index (HAQ DI) at Weeks 0 (baseline), 12, and 24. Morning Stiffness was evaluated at Weeks 0 (predose), 2, 4, 8, 12, 16, 20, and 24; and rheumatoid factor (RF) was evaluated at Weeks 0 (predose), 12, and 24. In addition, ACR20 area under the curve (AUC) over the 24-week study period was determined. ACR20 AUC was defined as the sum of the duration that patients achieved an ACR20 response.	Well established endpoint and widely used in clinical practice.
Keystone et al 2004 DE019	The ACR20 response at week 24 was the primary end point, and patients who did not achieve an ACR20 response, who withdrew from the study, or who received additional traditional DMARD therapy on or after week 16 were classified as nonresponders. .	Well established endpoint and widely used in clinical practice.	ACR20 responses as well as responses according to the 50% and 70% improvement levels (ACR50 and ACR70, respectively) were assessed at weeks 2 and 4, every 4 weeks from week 4 to week 24, every 8 weeks from week 24 to week 48, and a final time at week 52. Physical function was assessed at baseline and at each visit using the disability index of the Health Assessment Questionnaire (HAQ). Health-related quality of life was assessed at baseline and at weeks 12, 24, and 52 using the Medical Outcomes Study 36-item Short Form health survey(SF-36).	Well established endpoint and widely used in clinical practice.

			Safety was assessed through recording of adverse events, physical examinations, and standard laboratory tests. At baseline and at weeks 24 and 52, serum titers of antinuclear antibodies (ANAs) (positive titer _1:80) and anti–double-stranded DNA (anti-dsDNA) antibodies (positive titer _3.5 IU/ml, determined only if ANAs were elevated from baseline) were established by immunofluorescence on Hep-2 cells and by Farr radioimmunoassay, respectively. At baseline and at weeks 24 and 52, serum titers of anti-adalimumab antibodies (positive titer _20 ng/ml and not suppressed by _50% after the addition of human serum) were determined by a double-antigen immunoassay.	

STAR Furst et al 2003	Safety was the primary endpoint of this study and was assessed by types and frequencies of adverse events, physical examination findings, and standard laboratory test results. Adverse event data were stratified for adalimumab and placebo treatment, as well as for the number of concomitant traditional DMARD (i.e., 0, 1, or 2) administered with adalimumab or placebo. Serum titers of antinuclear antibodies (ANA) (positive ≥1:80) and anti-dsDNA antibodies (positive > 3.5 IU/ml) (performed if ANA titers increased from baseline) were determined by immunofluorescence and the Farr radioimmunoassay, respectively, at baseline and at Week 24.	Well established endpoint and widely used in clinical practice.	Efficacy was the secondary endpoint of this study and was assessed as ACR20, ACR50, and ACR70 responses10. Fulfillment of ACR20, ACR50, and ACR70 criteria was based on changes from baseline observed at Week 24 (using a nonresponder imputation technique so that all patients who withdrew from the study prior to Week 24 were counted as nonresponders). ACR20, ACR50, and ACR70 response rates were stratified for adalimumab and placebo treatment, as well as for the number of concomitant traditional DMARD (i.e., 0, 1, or 2) given with adalimumab or placebo.	Well established endpoint and widely used in clinical practice.
Chen et al 2009	Patients were assessed for the primary efficacy endpoint [the ACR20 response at week 12 in the modified full analysis set, and a subject was defined as a responder if the following three criteria were met: ≥ 20% improvement in tender joint count; ≥20% improvement in swollen joint count; and ≥ 20% improvement in at least three of the following assessments: pain visual analog scale (VAS; 0, no pain and 100, severe pain); patient's global assessment of disease activity (0,	Well established endpoint and widely used in clinical practice.	Secondary efficacy variables included the following: ACR50 and ACR70 responses at week 12, change from baseline in the individual components of the ACR response at week 12, and change from baseline in the presence/absence and duration of morning stiffness at week 12. ACR20 is defined as a reduction in tender and swollen joint counts of 20%, ACR50 of 50% and ACR70 of 70%, from baseline. Monitoring of vital signs, physical examinations, laboratory parameters (hematology, blood chemistry, CRP, routine urinalysis), and adverse events (AEs) was performed every month for	Well established endpoint and widely used in clinical practice.

	no disease activity and 10, extreme disease activity); physician's global assessment of disease activity (0, no disease activity and 10, extreme disease activity); the disability index of Health Assessment Questionnaire (HAQ; 0, without difficulty; 1, with some difficulty; 2, with much difficulty; and 3, unable to do so),20 and CRP values].		safety evaluation. The occurrence of treatment-emergent adverse events (TEAEs) was the primary safety variable.	

RAPID 1 Keystone et al 2008b	Co-primary end points were the ACR20 response rate at week 24 and the mean change from baseline in the modified total Sharp score at week 52.	Well established endpoint and widely used in clinical practice.	Major secondary end points included the change from baseline in modified total Sharp score at week 24, the change from baseline in the disability Index (DI) of the Health Assessment Questionnaire (HAQ) at weeks 24 and 52, the ACR20 responder rate at week 52, and the ACR50 and ACR70 responder rates at weeks 24 and 52. Additional secondary end points included mean changes from baseline in the following features: erosion and joint space narrowing scores, swollen (n _ 66 joints) and tender (n _ 68 joints) joint counts, physician's and patient's global assessments of disease activity, patient's assessment of arthritis pain, physical function (according to the HAQ DI), the Disease Activity Score 28-joint assessment (DAS28) (15), the ESR, and the CRP level.	Well established endpoint and widely used in clinical practice.
RAPID 1 Strand et al 2009	As above RAPID 1	Well established endpoint and widely used in clinical practice.	As above RAPID 1	Well established endpoint and widely used in clinical practice.
RAPID 2 Smolen et al 2009b	Assessments were made at baseline, weeks 1, 2, 4, 6, 8, 12, 14, 20 and 24, or withdrawal. The primary end point was ACR20 response at week 24, defined as a decrease of >20% from baseline in the number of tender (n=68) and swollen (n=66) joints, plus a 20% improvement in three or more of the following: patient's and physician's global assessment of disease activity, patient's assessment of arthritic pain, Health Assessment	Well established endpoint and widely used in clinical practice.	Secondary efficacy end points at week 24 included ACR50, ACR70, mean change from baseline in van der Heijdemodified Total Sharp Scores (mTSS), short Form-36 (SF-36) Health Survey, and individual ACR core set variables.	Well established endpoint and widely used in clinical practice.

	Questionnaire-Disability Index (HAQ-DI) and serum C-reactive protein or erythrocyte sedimentation rate (ESR).			

TEMPO Klareskog et al 2004	The primary efficacy endpoint was the numeric index of the ACR response (ACR N) area under the curve (AUC) over the first 24 weeks, as defined previously. (below)	Well established endpoint and widely used in clinical practice.	Other endpoints included ACR20, ACR50, and ACR70 responses and disease activity score. We defined remission as disease activity score less than 1·6.18	Well established endpoint and widely used in clinical practice.
TEMPO van der Heijde et al 2006b	As above	Well established endpoint and widely used in clinical practice.	As above	Well established endpoint and widely used in clinical practice.
Combe et al 2006	The primary efficacy end point was the percentage of patients achieving >20% improvement as assessed by the ACR 20 response at week 24 The patients' response to treatment was assessed at baseline and at weeks 2, 4, 8, 12, 16, 20 and 24.	Well established endpoint and widely used in clinical practice.	Secondary end points included the -ACR response rates (ACR 20, ACR 50 and ACR 70), -Disease Activity Scores (DAS; assessment of 44 swollen joints and ESR), -number of painful joints, number of swollen joints, morning stiffness (min), physician and patient global assessments (0–10 scales), Health Assessment Questionnaire (HAQ), pain Visual Analogue Scale (VAS), general health VAS, EuroQOL VAS,21 ESR and CRP at the aforementioned time points. Safety assessments were based on reports of adverse events and results of routine physical examinations and laboratory determinations. Treatment-emergent adverse events were defined as adverse events that were not present at baseline orevents that were present at baseline but worsened during the study.	Well established endpoint and widely used in clinical practice.
Moreland et al 1999	The primary efficacy end points were 20% and 50% improvement in disease	Well established endpoint and	Other efficacy end points included 70% ACR response at 3 and 6 months and percentage	Well established endpoint and widely

	activity at 3 and 6 months.	widely used in clinical practice.	change from baseline at 3 and 6 months in the following: tender joint count, swollen joint count, duration of morning stiffness, patient's global assessment, physician's global assessment, patient's assessment of pain, quality of life, erythrocyte sedimentation rate, and C-reactive protein level. Response was also evaluated according to the Paulus index, defined as a 20% or 50% improvement in at least four of the following variables: tender joint scores, swollen joint scores, duration of morning stiffness, erythrocyte sedimentation rate, patient's global assessment, and physician's global assessment.	used in clinical practice.

Weinblatt et al 1999	The primary end point with respect to efficacy was the proportion of patients meeting the ACR preliminary criteria for improvement in rheumatoid arthritis (ACR 20) at 24 weeks.	Well established endpoint and widely used in clinical practice.	Safety was evaluated according to the frequency of adverse events, laboratory abnormalities, and antibody formation. The other efficacy end points were the proportion of patients who reached the ACR 20 at 12 weeks and the proportions who met the ACR 50 and ACR 70 (defined in the same manner as ACR 20, but with improvements of 50 percent and 70 percent, respectively, in the various scores) at 12 and 24 weeks. Individual measures of disease activity, such as numbers of swollen and tender joints and physician's assessment, were evaluated at 12 and 24 weeks.	Well established endpoint and widely used in clinical practice.
GO AFTER Smolen et al 2009a	The primary endpoint was assessed at week 14 by achievement of a 20% or higher improvement in ACR criteria for assessment of rheumatoid arthritis (ACR20).	Well established endpoint and widely used in clinical practice.	Secondary endpoints were ACR20 at week 24; ACR50 and ACR70 at weeks 14 and 24; numeric index of the ACR response18 at weeks 14 and 24; DAS28 at weeks 14 and 24; HAQ-DI scores at weeks 14 and 24; fatigue score at weeks 14 and 24; DAS28 response according to EULAR (DAS28≤5·1 and improvement from base line>0·6, or improvement from baseline>1·2); and DAS28 remission (DAS28<2·6). Fatigue was assessed with the Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F) questionnaire, in which patients score (0–52) tiredness,weakness, and difficulty with usual activities because of fatigue; increased score indicates reduced fatigue.20,21 Serum samples taken at baseline and week 24 were assayed, for the presence of antibodies to golimumab. Safety was assessed by a general question to every patient about the number, type, and severity of	Well established endpoint and widely used in clinical practice.

			adverse events, which were coded according to MedDRA.	

GO FORWARD Keystone et al 2009b	There were two co-primary endpoints: the proportion of patients achieving an ACR20 response at week 14 and the improvement from baseline in HAQ-DI score at week 24.	Well established endpoint and widely used in clinical practice.	Response to treatment was assessed using the ACR response criteria (ACR20/50/70).19 ACR-N9 20 was also calculated. The disease activity score in 28 joints (DAS28)21 was calculated separately using both CRP and ESR. The health assessment questionnaire disability index (HAQDI) was used to evaluate physical function. The proportions of patients with a reduction from baseline in HAQDI of 0.25 or greater as a more conservative estimation of the minimum clinically important change. Serum samples taken at baseline and week 24 were assessed for the presence of antibodies to golimumab using a previously described assay. The presence of the study agent in the serum interferes with the detection of antibodies to the study agent in these types of assays.	Well established endpoint and widely used in clinical practice.
Kay et al 2008	The primary end point of the study was the proportion of patients meeting the ACR 20% improvement criteria (achieving an ACR20 response) at week 16. An ACR20 response was defined as _20% improvement in the tender joint count, swollen joint count, and 3 of the following 5 assessments: patient's assessment of pain, patient's global assessment of disease activity, evaluator's global assessment of disease activity, patient's assessment of physical function using the Health Assessment Questionnaire	Well established endpoint and widely used in clinical practice.	ACR50 and ACR70 responses were defined similarly using _50% and _70% improvements, respectively, and were included among the secondary end points. Other secondary end points included improvement from baseline at week 16 in the Disease Activity Score in 28 joints (DAS28), numeric index of the ACR response (ACR-N) at week 16, and ACR20, ACR50, and ACR70 responses over time through week 52.	Well established endpoint and widely used in clinical practice.

	(HAQ), and CRP level.			

ATTEST Schiff et al 2008	The primary endpoint was to evaluate a reduction in disease activity, measured by Disease Activity Score 28 (based on erythrocyte sedimentation rate levels; DAS28 (ESR)) with abatacept vs placebo at 6 months.	Well established endpoint and widely used in clinical practice.	Secondary endpoints included mean reduction in DAS28 (ESR) with infliximab vs placebo at 6 months. Additional secondary endpoints at 6 months and 1 year included: mean reduction in DAS28 (ESR) with abatacept vs infliximab; DAS28 (ESR) European League Against Rheumatism (EULAR) responses;6 low disease activity score (LDAS; DAS28 (ESR) (3.2); DAS28 (ESR)-defined remission (DAS28 (ESR) ,2.6); ACR 20, 50 and 70 responses; Health Assessment Questionnaire Disability Index (HAQ-DI) response rates (>0.3 improvement from baseline); and mean changes in the physical and mental component summary (PCS and MCS, respectively) scores, and eight subscales of the Short Form-36 (SF-36).	Well established endpoint and widely used in clinical practice.
ATTRACT Lipsky et al 2000	ACR 20	Well established endpoint and widely used in clinical practice.	ACR 50 and 70 Arthritis-related functional disability was measured with the Health Assessment Questionnaire. General health status was assessed by the Medical Outcomes Study Short-Form Health Survey (SF-36). Eight aspects of health status were assessed: general and mental health, physical function, social function, physical and emotional health, pain, and vitality; the score on each subscale ranges from 0 (worst) to 100 (best). The individual aspects of the survey were grouped into physical-component and mentalcomponent summary scores, each of which was assigned a mean (±SD) score of 50±10 on the basis of an	Well established endpoint and widely used in clinical practice.

			assessment of the general U.S. population of persons without chronic conditions. Individual scores were compared with the normalized scores for the general population.	

ATTRACT Maini et al 1999	The primary endpoint was prospectively defined as 20% improvement, according to the American College of Rheumatology at the week 30 visit without requiring a surgical joint procedure (ie, arthrodesis and joint replacement); initiation of new drugs for rheumatoid arthritis, or increases in dose of medication for rheumatoid arthritis. Patients received their baseline dose of methotrexate or corticosteroids during the trial.	Well established endpoint and widely used in clinical practice.	Secondary measurements of response to therapy included documentation of 50% and 70% improvement, reduction in individual measurements of disease activity, and a general health assessment.	Well established endpoint and widely used in clinical practice.
START Westhovens et al 2006b	The primary end point of the study was the proportion of patients who reported experiencing a serious infection within the first 22 weeks after initiating therapy.	Well established endpoint and widely used in clinical practice.	Safety evaluations included assessments of adverse events, measurements of vital signs, and the results of routine laboratory tests. A serious infection was defined as any infection identified by the investigator on the basis of clinical judgment or the results of culture, microscopy, serology, biopsy, or imaging that also met the definition of a serious adverse event Clinical response was evaluated at baseline and at weeks 0, 2, 6, 14, and 22, using the ACR criteria for an improvement response. Disease activity at week 22 was assessed using the Disease Activity Score in 28 joints (DAS28). Disease remission was defined as a DAS28 of less than 2.6.	Well established endpoint and widely used in clinical practice.
Abe et al 2006	The primary endpoint of the DBT was a response rate of a 20% improvement according to the ACR criteria (ACR20) at Week 14.	Well established endpoint and widely used in clinical practice.	Evaluations were made in terms of improvement of 20%, 50%, and 70% according to the ACR response (ACR20, ACR50 and ACR70) and individual measurements of the ACR core set at Weeks 0, 2, 6, 10, and 14 in the DBT and every 4 weeks from Weeks 0 to 36 in the OLT.	Well established endpoint and widely used in clinical practice.

			In the DBT, patients were monitored for safety until just before the first infusion of the OLT. Patients who did not enter the OLT were assessed until 20 weeks after the last infusion. In the OLT, safety assessments were performed until 36 weeks. An infusion reaction was defined as any adverse event occurring during or within 2 hours after the completion of each infusion. Vital signs including body temperature, blood pressure, and pulse rate were recorded every 30 min during and for 2 hours after the completion of each infusion.	

Maini et al 1998	The primary efficacy measurement was the total time (in weeks) that the patient exhibited a response to treatment. A composite index, as defined by the criteria of Paulus et a1, referred to as the Paulus 20% index, was used.	Well established endpoint and widely used in clinical practice.	Other secondary assessments of the magnitude of response and remission of disease werc evdluatcd as follows: 1) the proportion of patients with a *250%* improvement in disease activity according to the Paulus criteria 2) disease remission as defined by Pinals et a1, requiring the fulfillment of at least *S* of the following criteria for at least 8 consecutive weeks: no swollen or tender joints, morning stiffness <I5 minutes, no fatigue or pain (<0.5 cm on a 10-cm visual analog scale [VAS]), or an ESK <30 mm/hour for women and <20 mm/hour for men. Partial remission was prespecified as fulfilling *2 3* criteria for at least 8 consecutive weeks. The following measures were also evaluated for changes from baseline at weeks 1, 2, 4, 8, 12, 16, and 26: swollen joint count (SJC), the tender joint count (TJC), patient and physician global assessments, pain score (on 10-cm VAS), disability as assessed by a modified Stanford Health Assessment Questionnairc (HAQ) (22), and serum CRP levels. Safety was monitored until the end of the twenty-sixth week of the trial, whether or not the patient was continuing with trial medication. Adverse experiences observed by personnel at a study center, spontaneously reported by the patient at or between visits, or elicited from the patient by questioning at each visit wcre recorded.	Well established endpoint and widely used in clinical practice.
Edwards et al 2004	The primary end point of the study was	Well established	Secondary outcomes included ACR 20 and ACR	Well established

	the proportion of patients with an ACR 50 response at week 24. An ACR 50 response was defined as an improvement of at least 50 percent from baseline in counts of both tender and swollen joints, as well as in three of the five remaining disease-activity measures of the ACR core set: physician's assessment of disease activity, patient's assessment of disease activity, patient's assessment of pain, patient's assessment of physical function (by means of the health-assessment questionnaire), and the value for one acutephase reactant (either serum C-reactive protein level or erythrocyte sedimentation rate).	endpoint and widely used in clinical practice.	70 responses (20 percent and 70 percent improvement, respectively, according to the ACR criteria), a change in the disease-activity score (which includes the physician's assessment of 28 joints and the patient's self-assessment of disease activity), and the response according to the criteria of the European League against Rheumatism (EULAR response).	endpoint and widely used in clinical practice.

Strand et al 2006	The primary end-point of the study was the proportion of patients at week 24 with an ACR50 response, defined as _50% improvements from baseline in TJC and SJC, and three of the five components of patient pain [by visual analogue scale (VAS) score], patient assessment of global disease activity (VAS), physical function (HAQ-DI), physician global assessment of disease activity (VAS) and ESR or CRP.	Well established endpoint and widely used in clinical practice.	Secondary outcome measures included EULAR responses based on 'good' and 'moderate' improvements in Disease Activity Scores (DAS) derived from TJC, SJC, patient assessment of global disease activity and ESR or CRP.	Well established endpoint and widely used in clinical practice.
REFLEX Cohen et al 2006	The primary end point was the proportion of patients with an ACR20 response at week 24, defined as at least a 20% improvement from baseline values in the swollen joint count and the tender joint count as well as in 3 of the 5 remaining disease activity measures: physician's global assessment of disease activity; patient's global assessment of disease activity, patient's assessment of pain, patient's assessment of physical function, and either the CRP level or the ESR.	Well established endpoint and widely used in clinical practice.	Secondary end points included ACR50 and ACR70 responses (50% and 70% improvement from baseline according to the ACR criteria, respectively), changes from baseline to week 24 in scores on the Disease Activity Score 28-joint assessment for swelling and tenderness (DAS28), the EULAR response criteria (30), and the individual parameters of the ACR improvement criteria: swollen joint count, tender joint count, patient's and physician's global assessments of disease activity, patient's assessment of pain, patient's assessment of disability (using the Disability Index [DI] of the Health Assessment Questionnaire [HAQ]), the CRP level, and the ESR. Additional end points included changes from baseline to week 24 in the Functional Assessment of Chronic Illness Therapy–Fatigue (FACIT-F) score, the Short Form 36 (SF-36) health survey score (assessing health-related quality of life), and the Genant-modified Sharp radiographic score.	Well established endpoint and widely used in clinical practice.

			Regardless of their status in the study, all patients were to return for scheduled radiographs.	

REFLEX Keystone et al 2008a	As above (REFLEX)		As above (REFLEX)	
OPTION Smolen et al 2008	The primary efficacy endpoint was the proportion of patients with a 20% improvement in rheumatoid arthritis signs and symptoms according to ACR criteria (ACR20 response) at 24 weeks.	Well established endpoint and widely used in clinical practice.	Secondary efficacy endpoints included the proportion of patients with an ACR50 and ACR70 response at 24 weeks, change from baseline in disease activity score using 28 joint counts (DAS28)18,19 at 24 weeks, the proportion of patients in DAS28 remission (DAS28 <2·6) at 24 weeks, and categorical DAS28 (European League Against Rheumatoid Arthritis [EULAR]) response by 24 weeks.20 Haemoglobin concentrations were also assessed.	Well established endpoint and widely used in clinical practice.
RADIATE Emery et al 2008	The primary endpoint was the ACR20 response at week 24.	Well established endpoint and widely used in clinical practice.	Secondary endpoints included further efficacy measures. Safety outcomes included adverse events, infections and infusion reactions.	Well established endpoint and widely used in clinical practice.
SATORI Nishimoto et al 2009	The primary end point was the ACR20 response at week 24 with the last observation carried forward (LOCF) method, using an intent-to-treat (ITT) analysis.	Well established endpoint and widely used in clinical practice.	Secondary endpoints were ACR50 and ACR70 at week 24.	Well established endpoint and widely used in clinical practice.
TOWARD Genovese et al 2008	The primary end point was the proportion of patients who had achieved a response according to the ACR criteria for 20% improvement (ACR20) at week 24.	Well established endpoint and widely used in clinical practice.	Secondary end points included the proportion of patients with 50% or 70% improvement (ACR50/70) at week 24, as well as the time to onset of ACR20/50/70 responses. The Disease Activity Score in 28 joints (DAS28) based on the erythrocyte sedimentation rate (ESR), the European League Against Rheumatism (EULAR) response, changes in hemoglobin levels, and disability index of the Health Assessment	Well established endpoint and widely used in clinical practice.

			Questionnaire (HAQ) (25), Functional Assessment of Chronic Illness Therapy–Fatigue (FACIT-F) , and Short Form 36 (SF-36) scores were also assessed at week 24. Safety was assessed by examining adverse events (AEs), serious AEs, infections, withdrawals due to AEs, deaths, and clinically significant changes in vital signs and laboratory test results.	

Table 179 Summary of statistical analyses in RCTs

Trial no. (acronym)	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
Kremer et al., 2003	Hypothesis testing: if there was a significant difference in the rates of ACR 20 responses between the group given 10 mg of cytotoxic T-lymphocyte–associated antigen 4–IgG1 (CTLA4Ig) per kilogram and the placebo group with use of a chisquare test, then we compared the group given 2 mg of CTLA4Ig per kilogram with the placebo group. This testing strategy was also used to identify differences in the rates of ACR 50 and ACR 70 responses.	. A closed testing procedure based on an ordered analysis of variance was established for hypothesis testing - Descriptive statistics were used to compare the demographic and base-line characteristics of the patients in the three treatment groups. - efficacy analyses included all patients who received at least one dose of study medication. -When assessing the change from base line in the health-related quality of life and the individual components of the ACR response in patients who discontinued the study for any reason, we used the values obtained at the last assessment and carried them forward. - A secondary analysis was performed in which all patients who discontinued the study for any reason were classified as having had no response. -Fisher's exact tests were used to compare the incidence of adverse events in the CTLA4Ig groups and the placebo group. For other end points, analysis of covariance (adjusted for base-line values) with linear contrasts was used for continuous variables and chi-square tests were used for proportions. All statistical tests were two-sided and conducted at the 5 percent level.	A sample of 107 patients per treatment group was determined to yield 94 percent power at the 5 percent level (two-sided) to detect an absolute difference of 25 percent between the group given 10 mg of CTLA4Ig per kilogram and the group given placebo plus methotrexate, on the basis of an expected ACR 20 response rate at six months of 25 percent in the placebo group and a dropout rate of 15 percent in each treatment group	Study medication was administered to 339 patients: - 119 patients were randomly assigned to receive placebo plus methotrexate, - 105 to receive 2 mg of CTLA4Ig per kilogram plus methotrexate, - 115 patients to receive 10 mg of CTLA4Ig per kilogram plus methotrexate. A total of 259 patients completed six months of treatment. More patients in the placebo group discontinued the study than in either of the CTLA4Ig groups. The most common reason for discontinuation was lack of efficacy as indicated by worsening arthritis. -To account for missing data in the assessment of the ACR responses in the primary, prespecified analysis, we considered patients who discontinued the study because of worsening disease not to have had a response, and we carried forward the values obtained at the last assessment for patients who discontinued the study for any other reason. Thus, all patients were assessed for an ACR response.
Kremer et al., 2005	To determine the clinical efficacy, safety, and immunogenicity of abatacept (CTLA-4Ig), a selective costimulation modulator, in patients with	-All statistical analyses were carried out on the intent-to treat (ITT) population, defined as all patients who received at least 1 treatment infusion. -Baseline demographics and disease history were analyzed using descriptive statistics. -Differences in ACR20, ACR50, and ACR70	A sample size of 107 patients per treatment group was calculated to yield 94% power to detect a difference of 25% in ACR20 responses between the 2 abatacept groups and the placebo group at	When ACR response rates were assessed, all patients who discontinued from the study due to worsening RA disease (lack of efficacy) were considered nonresponders from that time point.

	rheumatoid arthritis (RA) that has remained active despite methotrexate (MTX) therapy.	response rates on day 360 were analyzed by comparing each abatacept treatment group with the placebo group using a Dunnett-adjusted chi-square test -ACR response rates at other time points were compared between each abatacept treatment group and the placebo group using a chi-square test unadjusted for multiple comparisons. -Differences in percentage change from baseline to the last observation carried forward (LOCF) for all ACR core components were analyzed using analysis of covariance with the baseline value as the covariate and without adjustment for multiple comparisons. -Fisher's exact tests were used to compare the incidence of AEs between the abatacept treatment groups and the placebo group. -For all other end points, discrete variables were analyzed using chi-square tests, and all continuous variables were analyzed by *t*-tests unadjusted for multiple comparisons. -All statistical tests were conducted at the 5% significance level (2-tailed).	the 5% significance level (2-tailed), adjusted for a discontinuation rate of 15%.	However, patients who discontinued for other reasons had their last observations carried forward.

AIM (Kremer et al 2006).	To evaluate the effects of abatacept in patients with persistent, active rheumatoid arthritis despite methotrexate treatment.	-all efficacy and safety analyses on a modified intention-to-treat population, defined as all randomly assigned patients who received at least 1 dose of study medication. - all statistical tests on a 2-sided 5% level of significance and used SAS software, version 8.2 (SAS Institute, Cary, North Carolina), for all analyses. - co-primary analyses of ACR 20 at 6 months and HAQ-DI responses at 1 year, a 2-sided, continuity-corrected chi-square test was used to compare the responses of the abatacept group with those of the placebo group.	The protocol estimated that 680 patients would need to be enrolled to randomly assign 540 patients. - Sample sizes were based on a 5% level of significance (2-tailed). -99% power was used to detect a difference of 20% in ACR 20 between the 2 groups. - On the basis of the hierarchical testing procedure for the co-primary measures, the sample size allowed detectection of an 18% difference in HAQ-DI response rate between the 2 groups, with 98% power, and a treatment effect of 60% reduction from placebo (assuming an increase of 1.27 units in placebo for the change from baseline), with 90% power, for change from baseline in the Genant-modified Sharp erosion score.	Missing data for patients who discontinued as nonresponders subsequent to the discontinuation was inputed; thus, we based these analyses on the full modified intention-to-treat denominator. -additional sensitivity analyses to assess the effect of the imputation of missing data was performed. - including a "modified worst-case" analysis, -missing data for placebo recipients who discontinued for reasons other than lack of efficacy by using their last observed response was inputted - a "worst-case" analysis, where we imputed missing data for placebo recipients who discontinued as responders. Both cases, -missing data for abatacept recipients as nonresponders was inputted. -additional longitudinal analyses by using the generalized estimating equations to assess the treatment effect over time.
ASSURE Weinblatt et al 2006	To assess the safety of abatacept, a selective costimulation modulator, in patients with active rheumatoid arthritis (RA) who had been receiving >1 traditional nonbiologic and/or biologic diseasemodifying	- Data were assessed for all patients treated with either abatacept or placebo in order to ascertain the safety and patient- and physician-reported benefit of abatacept in the overall population (all patients, regardless of background therapy) of this study. Furthermore, data from each treatment group were assessed according to background therapy (nonbiologic DMARDs versus biologic DMARDs)	- The study was powered to detect AEs occurring at a rate of 0.2%. - No further information is stated regarding the sample size calculation.	

	antirheumatic drugs (DMARDs) approved for the treatment of RA for at least 3 months prior to entry into the study.	to determine whether differences in the effects of abatacept between these patient populations would be evident. -Safety analyses were based on a data set containing all available assessments from all patients who received at least 1 infusion of study medication (treated patients).		

ATTAIN Genovese et al 2005	To evaluate the efficacy and safety of abatacept, a selective costimulation modulator, in patients with active rheumatoid arthritis and an inadequate response to at least three months of anti–TNF*a* therapy.	All efficacy analyses included all randomized patients who received at least one dose of study medication. For the primary analyses of the ACR 20 and HAQ responses, the proportion of patients who had a response at six months was summarized according to the treatment group. A two-sided Cochran–Mantel–Haenszel chi-square test (with stratification according to baseline anti–TNF*a* use [current or former]) was used to compare response rates in the abatacept group with those in the placebo group at the 0.05 level of significance. The primary and the multiple secondary end points were tested in a prespecified sequence after the use of a closed testing procedure, thus controlling the overall type I error rate at the 0.05 level. All reported P values are two-sided.	The statistical power of the study with respect to the two primary end points of the ACR 20 response and the HAQ response was 96 percent and 87 percent, respectively, at a two-sided alpha level of 5 percent, to detect absolute differences of 20 percent and 18 percent, respectively. No interim analyses were planned or conducted.	For patients who discontinued treatment, the last observation was carried forward in subsequent analyses. Safety was evaluated according to the frequency of adverse events, changes in laboratory values, and abnormal clinical findings. P values for safety comparisons were obtained with the use of a chi-square test or, where appropriate, Fisher's exact test. For the analyses of ACR 20 and HAQ responses, all patients who discontinued treatment were subsequently considered not to have had a response
Kim et al 2007	To investigate the efficacy and safety of 40 mg every-other-week (eow) subcutaneous injections of adalimumab with methotrexate (MTX) versus placebo with MTX in Korean patients with RA with insufficient responses to MTX.	-Patients receiving at least one injection of the study drug were included in the ITT analysis set. -Patients with missing data at week 24 and patients switching to open-label rescue treatment prior to week 24 were counted as non-responders. -All ACR20 response rates,including the primary endpoint, were compared using Pearson's chi-square test. -For secondary efficacy assessments of ACR50 and ACR70, Pearson's chi-square test was also used. Differences in the change from baseline to the last observation carried forward (LOCF) to week 24 in other secondary efficacy endpoints were compared between the adalimumab and placebo groups using the Wilcoxon rank sum test.	A sample size of 44 patients per group was estimated to provide 80% power for detecting a 29% difference in ACR20 response between placebo and adalimumab at week 24 at an ◎-level of 0.05, assuming a placebo response rate of 26% and an adalimumab response rate of 55%. Assuming a degree of uncertainty in determining sample size in a country in which adalimumab had not been used, 60 patients per study group were planned.	

		Treatment comparisons were performed at an ⍺-level of 0.05. All patients who received at least one dose of study drug were included in the safety analysis. Incidence rates of treatment-emergent AEs of the adalimumab group were compared with the placebo group using Pearson's chi-square test across treatment groups or Fisher's exact test, if 20% of expected cell counts were < 5. The treatment comparison was performed at an ⍺-level of 0.05.		

van de Putte et al 2004	To evaluate the efficacy and safety of monotherapy with adalimumab in patients with RA for whom previous DMARD treatment has failed	Demographic and baseline characteristics were summarized by descriptive statistics and compared between treatment groups to assess baseline homogeneity using one way analysis of variance or the Kruskal-Wallis test for continuous variables and the Cochran-Mantel-Haenszel test or Pearson's x2 test for discrete variables. An intention to treat analysis was performed that included patients who were randomised and who received at least one injection of the randomised study drug. The response rates for the primary efficacy end point (ACR20 at week 26) for each adalimumab group were compared with that of placebo using a two sided Pearson's x2 test. The Bonferroni-Holm procedure was applied to account for multiplicity of testing. Improvement in RA (that is, fulfilment of ACR20 response criteria) was defined as change from baseline. Patients not completing the trial (that is, who were withdrawn or required rescue) despite fulfilling ACR criteria were considered nonresponders upon withdrawing or entering rescue. Improvements in the seven ACR core components were compared between adalimumab and placebo treated patients using an analysis of covariance model, with baseline values as covariates. No a correction for multiple testing was applied to secondary efficacy variables. Comparisons of the active treatment groups with placebo were performed on the basis of the adjusted means resulting from these models.	A sample size of 90 patients in each treatment group was required to detect a difference of 30% in ACR20 response rates between an adalimumab dose group and placebo (with a predicted ACR20 response rate of 20% for placebo and with statistical power of at least 95% probability). The overall level of significance was set at p=0.05. The study was not powered to detect differences between individual adalimumab groups. To account for the few patients who might not be evaluable for any reason, the sample size was set at 100 patients in each treatment group	
ARMADA	To detect a difference of 35%	Demographic and baseline characteristics were	To detect a difference of 35% in	To adjust for the higher

Weinblatt et al 2003	in ACR20 response rates between any of the tested doses of active drug and placebo,	compared among all dosage groups, as determined using the Kruskal-Wallis test for continuous variables and the chi-square test for discrete variables. Efficacy end points were analyzed on an intent-to-treat basis and included all patients who received at least 1 dose of study drug (adalimumab or placebo). Differences in the percentage of patients achieving an ACR20 response at week 24 (the primary efficacy end point) were compared between each of the adalimumab dosage groups and the placebo group by use of Dunnett's test. Differences in the percentages of patients achieving ACR50 and ACR70 responses at week 24 (secondary efficacy end points) were compared between each of the adalimumab dosage groups and the placebo group by use of an unadjusted *t*-test, without correction for multiple comparisons. Differences in the change from baseline to the last observation carried forward (LOCF) to week 24 in other secondary efficacy end points were compared between each of the adalimumab dosage groups and the placebo group by use of analysis of covariance, with baseline as the covariate and without correction for multiple comparisons. The significance of the change from baseline in efficacy end points within each treatment group was assessed using the corresponding 95% confidence interval (if "0" was not contained in the	ACR20 response rates between any of the tested doses of active drug and placebo, assuming a placebo response rate of 20% and 90% power, a sample size of 67 patients was calculated to be required for each of the 4 treatment arms. Statistical significance was set at *P* _ 0.05 for all tests. The study was not powered to show a difference among adalimumab treatment groups.	rate of withdrawals and shorter amount of treatment time in the placebo group, adverse events were also analyzed by the total number of patients experiencing a particular adverse event per total years of treatment (number of patients/patientyear). Fulfillment of the ACR20, ACR50, and ACR70 criteria was based on changes observed from baseline to week 24. Patients who dropped out before week 24 and patients who did not achieve an ACR20 response were classified as nonresponders.

		interval, the change was considered significant). Adverse events were analyzed by frequency and percentage and were compared between the adalimumab and placebo groups by use of Pearson's chi-square test, without correction for multiple comparisons.		

Miyasaka 2008 CHANGE	To detect a difference of 25% in ACR20 response rates between the placebo group and the adalimumab 40 mg group	The primary efficacy endpoint, ACR20 response rate at Week 24, was compared for the placebo group against that of the 40 and 80 mg eow adalimumab groups, using the Pearson v2 test. The Hochberg procedure was applied to control for multiplicity of testing. If both P values were less than 0.05, then the individual null hypotheses (no treatment difference between adalimumab and the placebo) were rejected. If one P value did not show significance at 0.05, then the other hypothesis was tested against and adjusted at the 0.025 level. If the test was significant at the adjusted 0.025 level, then the null hypothesis was rejected.	To detect a difference of 25% in ACR20 response rates between the placebo group and the adalimumab 40 mg group, assuming an ACR response rate of 20% in the placebo arm and 45% in the 40 mg eow arm, a sample size of 74 patients per treatment group was estimated to be required to provide 80% power for a two-sided test (continuity corrected) with an alpha of 0.025. Therefore, taking the exclusion analysis into consideration, a total of 320 subjects (80 subjects per treatment group) needed to be equally allocated to one of the four treatments: 20 mg adalimumab, 40 mg of adalimumab, 80 mg adalimumab, or placebo.	Patients who discontinued the study prior to Week 24 or who moved to the rescue arm following at least eight weeks of treatment were classified as nonresponders.
Keystone et al 2004 DE019	To assess the ability of adalimumab, a human anti–TNF monoclonal antibody, to inhibit the progression of structural joint damage, reduce the signs and symptoms, and improve physical function in patients with active RA receiving concomitant treatment with methotrexate (MTX).	Demographic and baseline clinical characteristics were analyzed by Kruskal-Wallis test for continuous variables and Pearson's chi-square test for discrete variables. An intent-to treat population was formed for efficacy analyses and was defined as all patients who received at least 1 dose of study drug. The 3 primary efficacy end points were analyzed in hierarchic order, beginning with the ACR20 response rates, followed by the modified total Sharp scores, and ending with HAQ scores. A closed testing procedure was chosen to control the overall significance level at 0.05. An initial global	The power calculation was based on both the predicted ACR20 response rate and radiographic findings. A sample size of 200 patients per treatment group was estimated to provide _95% power for detecting a difference of _20% in ACR20 response rates at week 24 between the placebo group and the adalimumab groups at a significance level of _ _ 0.05, assuming a placebo response rate of 35%.	

		null hypothesis was tested for the first hierarchic primary efficacy end point, the ACR20 response. If this was significant (*P* _0.05), pairwise comparisons between each adalimumab group and the placebo group would be performed. If all individual hypothesis tests were significant, then a repeat of the aforementioned testing procedure for the second hierarchic primary efficacy end point, the modified total Sharp score, would be done. This was repeated a third time for the HAQ if the modified Sharp score showed a significant difference. Tests of normality for the change from baseline in the total Sharp score and HAQ score were conducted using the Shapiro-Wilk test. If the data were normal, analysis of covariance (ANCOVA) would be performed.	Assuming that 70% of patients would have evaluable radiograph films at 12 months, a sample size of 140 patients per treatment group was estimated to provide 90% power for detecting a difference in the mean increase in the modified total Sharp scores at a significance level of _ _ 0.05, assuming a mean change of 2.0 in the placebo group, 0.5 in each adalimumab group, and a pooled standard deviation of 4.0. The study, however, was not powered to distinguish differences between adalimumab groups.	

STAR Furst et al 2003	This study evaluated the safety and efficacy of adalimumab (HumiraTM), a fully human monoclonal tumor necrosis factor-alpha (TNF-α) antibody, when given with standard antirheumatic therapy in patients with active rheumatoid arthritis (RA) not adequately responding to such therapies. Standard antirheumatic therapy included traditional disease modifying antirheumatic drugs (DMARD), low dose corticosteroids, nonsteroidal antiinflammatory drugs (NSAID), and/or analgesics.	Demographic and baseline disease characteristics were analyzed using Wilcoxon rank sum test for continuous variables and Pearson's chi-square test for discrete variables. Statistical comparisons for the frequency of adverse events were made between the adalimumab and placebo groups using Pearson's chi-square test. The efficacy analysis was performed on an intent-to-treat basis, including all patients who received at least one injection of study drug and had at least one efficacy assessment. ACR20, ACR50, and ACR70 response rates observed at Week 24 were compared between the adalimumab and placebo groups using Pearson's chi-square test with a 2-sided level of significance of ◎◎= 0.05. No correction was made for multiple statistical comparisons.	A sample size of 300 patients per group was determined to demonstrate a specific adverse event rate of 1%, or less, with 95% confidence.	
Chen et al 2009	The objective of this study was to compare the efficacy and safety of adalimumab plus methotrexate (MTX) and MTX alone in Taiwanese patients with active RA.	The efficacy analysis was performed on an "intentto-treat (ITT)" population, which was defined as all patients with baseline data and at least one posttreatment evaluation. The change and percentage change from baseline in the treatment group were determined by nonparametric Wilcoxon signed rank test. The differences between treatment groups for the efficacy endpoints were compared by nonparametric Wilcoxon rank sum test. The differences in the ACR 20%, 50%, and 70% response rate were analyzed by Fisher's exact test. The safety analysis was performed on all patients who received randomized study medication.	Not stated	Not stated

		TEAEs included all AEs that either began on or after administration of study drugs, or preexisting conditions that worsened on or after study drug administration. The number and percentage of subjects reporting TEAEs were tabulated by MedDRA21 preferred terms and system organ class. Vital signs and laboratory data profiles were analyzed based on change from baseline using nonparametric Wilcoxon rank sum test for the analysis between treatment groups, and Wilcoxon signed rank test for the analysis within treatment groups. The number of patients with AEs was compared between treatment groups using Fisher's exact test.		

RAPID 1 Keystone et al 2008b	To evaluate the efficacy and safety of 2 dosage regimens of lyophilized certolizumab pegol (a novel PEGylated anti–tumor necrosis factor agent) as adjunctive therapy to methotrexate (MTX) in patients with active rheumatoid arthritis (RA)	Efficacy analyses were con based on the larger estimate to control for Type II error. Efficacy analyses were conducted on an intent-to-treat (ITT) population, which consisted of all patients who were randomized into the study. Primary analyses were performed using nonresponder imputation. Hypothesis testing for the co-primary end points was performed in a hierarchical manner. First, comparisons of the ACR20 responses at week 24 between the placebo group and each of the 2 certolizumab pegol dosage groups were performed using logistic regression, with treatment and geographic region as factors. The treatment effect was estimated using odds ratios and corresponding 97.5% confidence intervals obtained by fitting this model. Rejection of the null hypothesis for the ACR20 response enabled comparison of each active treatment with placebo in terms of the change from baseline in the modified total Sharp score at week 52. This latter analysis was performed using analysis of covariance (ANCOVA) on the ranks, with treatment and geographic region as factors and with the ranked baseline modified total Sharp score as the covariate. For patients who withdrew early (before week 52) and who had radiographs taken at their withdrawal visit, the modified total Sharp score at week 52 was estimated by linear extrapolation of the scores on	Sample size was determined on the basis of anticipated differences between the certolizumab pegol groups and placebo for both of the primary efficacy end points. For the ACR20 response, a sample size of 590 patients was required in order to have 90% power to detect a statistical difference of _20% between the certolizumab pegol groups and placebo with a 2-sided significance level of 2.5%. For the modified total Sharp score, a sample size of 950 patients was determined to be sufficient to detect differences of _2.2 Sharp units between an active drug group and a control group with _90% power (assuming an SD of 7 Sharp units). The sample size was based on the larger estimate to control for Type II error.	Patients who received rescue medication or who withdrew for any reason, including safety, were considered nonresponders from that time point onward.

		the radiographs taken at the early withdrawal visit or, if this was not performed, at week 24. Multiple sensitivity analyses were performed on the radiographic data under various assumptions on the imputation of missing values, including an analysis of the per-protocol pop- ulation, which consisted of a subset of the ITT population, excluding patients who had at least 1 major protocol deviation, as confirmed during a preanalysis review prior to unblinding of the data. Sensitivity analyses were also performed using the last observation carried forward (LOCF) method for imputation of missing scores. Comparison of active treatment versus placebo for the major secondary end points was tested at the 5% level of significance. Analysis of the ACR20, ACR50, and ACR70 responder rates was performed using logistic regression, with treatment and geographic region as factors. Analysis of secondary continuous efficacy end points was performed using ANCOVA, with geographic region and treatment as factors and baseline values as the covariate. Analysis of the responders according to the MCID for the HAQ DI values was post hoc and was analyzed using a repeated-measures logistic regression. Safety analyses were conducted on the safety population, which consisted of all patients who received at least 1 dose of medication. Adverse events are presented		

		as either the number of events or the incidence rate per 100 patient-years to adjust for differences between certolizumab pegol and placebo exposure.		

RAPID 1 Strand et al 2009	As above		As above	
RAPID 2 Smolen et al 2009	To evaluate the efficacy and safety of certolizumab pegol versus placebo, plus methotrexate (MTX), in patients with active rheumatoid arthritis (RA).	Efficacy analyses were conducted on the intention-to-treat (ITT) population (all randomised patients). Primary analysis used non-responder imputation; patients who received rescue drugs (any non-biological disease modifying antirheumatic drug other than MTX, any other biological agent, intravenous corticosteroids, or intra-articular hyaluronic acid) or withdrew for any reason were considered non-responders from that time point onward.	Sample size was determined on expected differences in the ACR20 responder rate between certolizumab pegol and placebo at week 24 Five hundred and ninety randomised patients would provide 90% power to detect a difference of >20% in ACR20 response at week 24 between each certolizumab pegol group and placebo at a two-sided significance level of a=0.025, assuming a placebo rate of 30%.	Not stated
TEMPO Klareskog et al 2004	The objective of the study was to assess combination treatment with etanercept and methotrexate versus the monotherapies in patients with rheumatoid arthritis.	For both the clinical and radiographic endpoints, we did two primary comparisons—combination versus methotrexate and etanercept versus methotrexate— with Hochberg's approach21 for multiple comparisons. Statistical tests were two-sided with significance defined as p<0·05.	The planned enrolment of 205 patients per group gave 90% power to detect a pairwise difference between groups of 4·5 units in ACR-N AUC, with a two-sided test at _=0·05 and assuming an SD of 14.	For patients who dropped out before 1 year, we did a radiographic examination at the time of discontinuation and estimated the 52-week total Sharp score by linear extrapolation.
TEMPO van der Heijde et al 2006b	To compare patient reported measures of function, health related quality of life (QoL), and satisfaction with medication among patients with rheumatoid arthritis (RA) treated with methotrexate (MTX), etanercept, or both for up to 1 year	Analyses were conducted for all the enrolled patients; to reduce bias and loss of statistical power, missing data due to study drop-out or for other reasons were imputed using the last observation carried forward (LOCF) method.14 All PRO measures (HAQ disability index, eight HAQ subscale scores, EQ-5D VAS, GHVAS, PGAD) were compared between treatment groups using the mean change from baseline and area under the curve (AUC).	As above	As above

		Comparisons between treatment groups were also performed using least squares means (means adjusted for potential imbalance in baseline values using a model fitted by the least squares method17) and 95% confidence intervals (CIs) for EQ-5D VAS, PGAD, and GHVAS. These comparisons used an analysis of covariance model that included baseline score as a covariate and factors for study centre, treatment, and prior MTX use.		

Combe et al 2006	To compare the efficacy and safety of etanercept and sulfasalazine, alone and in combination, in patients with active rheumatoid arthritis despite sulfasalazine treatment.	Binary efficacy end points of ACR 20, ACR 50 and ACR 70 response rates were analysed using the Mantel–Haenszel x2 test, stratified by study centre. For continuous and ordinal efficacy end points, including physician and patient global assessment of disease activity, pain VAS, number of swollen joints and painful joints, HAQ, ESR, CRP and morning stiffness, the changes from baseline were analysed with a two-way analysis of covariance with treatments and centre as factors and the baseline value as a covariate. For comparisons of baseline demographic characteristics among treatment groups, a one-way analysis of variance with treatment as a factor was used for continuous or ordinal variables and Mantel–Haenszel x2 test for binary variables. Adverse events (all and treatment-emergent) were summarised and compared among treatment groups, using Fisher's exact test. The baseline data and adverse event comparisons were based on an intent-to-treat population, including all randomly assigned patients who received any test article.	Not stated	Not stated
Moreland et al 1999	To confirm the benefit of etanercept therapy of longer duration and simplified dosing in patients with rheumatoid arthritis.	The ACR response rates and Paulus indices were compared by using the likelihood ratio chi-square test. The Fisher exact test was substituted when necessitated by low response rates (50% ACR response at 2 weeks and 70% ACR response). Individual measures of disease activity were compared by using analysis of variance in which treatment, study site, and their interaction were the factors. The last	Not stated.	Patients who withdrew for any reason were counted as nonresponders subsequent to withdrawal.

		available observation was used for dropouts. If the initial comparison of the three treatments was significant at the *P* 5 0.05 level, each pair of treatments was compared (also at the 0.05 level). This procedure controls the type I error at the 0.05 level. The Stuart–Maxwell chi-square test was used to test for normalization of laboratory values (within treatment groups). We conducted all analyses by using version 6.12 of SAS software (SAS Institute, Cary, North Carolina).		

Weinblatt et al 1999	This study was undertaken to determine whether the addition of etanercept, a soluble tumor necrosis factor receptor (p75):Fc fusion protein (TNFR:Fc), to methotrexate therapy would provide additional benefit to patients who had persistent rheumatoid arthritis despite receiving methotrexate.	Response rates measured by the ACR 20 and ACR 50 were compared with use of the chi-square test. Fisher's exact test (twotailed) was used for response rates according to the ACR 70 and for data on safety. With regard to the ACR response measures, patients who withdrew from the study were considered not to have had a response at all points after withdrawal, irrespective of the clinical response. For individual measures (tender and swollen joints and global assessments), the last observation was used in analysis if the patient withdrew. Patients who received intraarticular injections of corticosteroids during the study were counted as having or not having a response according to their overall evaluation, but the joint or joints injected were counted as tender and swollen for the remainder of the study. Patients who received increased doses of oral corticosteroids were considered not to have had a response at all time points after the increase	The power of the study with respect to the primary efficacy end point (based on the ACR 20) was estimated to be approximately 80 percent, on the assumption that the response rates would be 25 percent in the placebo-plus-methotrexate group and 55 percent in the etanercept-plus-methotrexate group. At the planned sample size of 75 patients, if the underlying rate of adverse events was 5 percent, the probability of observing at least one adverse event was 92 percent in the etanercept-plus-methotrexate group (50 subjects) and 72 percent in the placebo-plus-methotrexate group (25 subjects).	With regard to the ACR response measures,patients who withdrew from the study were considered not to have had a response at all points after withdrawal, irrespective of the clinical response. For individual measures (tender and swollen joints and global assessments), the last observation was used in analysis if the patient withdrew. Patients who received intraarticular injections of corticosteroids during the study were counted as having or not having a response according to their overall evaluation, but the joint or joints injected were counted as tender and swollen for the remainder of the study.
GO AFTER Smolen et al 209a	The efficacy and safety of the TNFα inhibitor golimumab in patients with active rheumatoid arthritis who had previously received one or more TNFα inhibitors.	The primary endpoint was tested with a hierarchical approach. If a Cochran-Mantel-Haenszel test, stratifi ed by baseline methotrexate use, showed a signifi cant diff erence between the proportion of patients with ACR20 in the combined golimumab groups (50 mg and 100 mg) and on placebo, then pairwise comparisons were made between 50 mg golimumab and placebo, and between 100 mg golimumab and placebo. Achievement of the primary endpoint required that the proportion of	A sample size of 140 patients per treatment group was calculated to provide more than 90% power at the 5% level of significance. This calculation assumed that 50% of patients used methotrexate at baseline, and ACR20 occurred in 30% of the placebo group (irrespective of methotrexate use), 45% of the 50 mg golimumab group that used methotrexate, 40% of the 50 mg golimumab group that did not use methotrexate, 55% of the 100 mg golimumab group that used	Patients were included in the statistical analysis if they discontinued the study drug for reasons unrelated to lack of eff ectiveness and returned for assessment, but they were regarded as non-responders if they met any of the failure criteria above.

		patients with ACR20 on combined golimumab and 50 mg or 100 mg golimumab, or both, was signifi cantly greater than was those on placebo. Secondary endpoints with discrete data were assessed with a Cochran-Mantel-Haenszel test, stratified by baseline methotrexate use. Secondary endpoints with continuous data were assessed by ANOVA from the van der Waerden normal scores. Subgroup analysis of DMARD use at baseline, and number of previous TNFα inhibitors and reason for their discontinuation were done to compare the combined golimumab group with the placebo group. Patients who were missing all components of the ACR or DAS response criteria were regarded as non-responders. Additionally, patients were deemed to have failed to achieve the primary endpoint if they had initiated treatment with a new DMARD, systemic immuno suppressive, or biologically derived drug for rheumatoid arthritis; increased methotrexate, sulfasalazine, or hydroxychloroquine dose above the baseline dose for treatment of rheumatoid arthritis; initiated treatment with or increased the dose of a corticosteroid; or discontinued the study drug because of an unsatisfactory treatment effect. All effi cacy data were analysed by intention to treat. All safety data were analysed according to the study	methotrexate, and 50% of the 100 mg golimumab group that did not use methotrexate.	

		drug that the patient received; patients who were randomised but never treated were not included. For patients who received rescue therapy, effi cacy data from week 16 were carried forward for analysis at week 24 to ensure that the results were not biased by the increased dose the patient received. No statistical tests were done on safety data. All statistical tests were two-sided (α=0·05) and done with SAS software (version 8.2).		

GO FORWARD Keystone et al 2009b	The phase III GO-FORWARD study examined the efficacy and safety of golimumab in patients with active rheumatoid arthritis (RA) despite methotrexate therapy.	There were two co-primary endpoints: the proportion of patients achieving an ACR20 response at week 14 and the improvement from baseline in HAQ-DI score at week 24. A two-sided x2 test was used to analyse the ACR20 data and a two-sided analysis of variance on the van der Waerden normal scores was used to analyse the HAQ-DI data; both were conducted at a significance level of a = 0.05. The coprimary endpoints were analysed sequentially (ACR20 response at week 14 first and HAQ-DI at week 24 second) to maintain an overall type I error rate of 0.05. Additional details regarding statistical testing of the primary endpoint and data handling guidelines are provided in supplemental material 2 (available online only).	Assuming 55% or more of patients in groups 3 and 4 and 35% of patients in group 1 would achieve an ACR20 response, a sample size of 120 patients in group 1 and 80 patients in groups 3 and 4 was required to achieve greater than 90% power (two-sided x2, a = 0.05). Assuming 55% of patients in group 2 and 35% of patients in group 1 would achieve an ACR20 response, a sample size of 120 patients in both groups 1 and 2 was needed to achieve greater than 85% power (two-sided x2 test, a = 0.05). This sample size would also provide greater than 90% power to detect a difference in the change from baseline in HAQ-DI score between treatment groups (two-sided t test on the van der Waerden normal scores, a = 0.05), assuming an improvement from baseline in HAQ-DI of 20.21 for group 1, 20.47 for group 3 and 20.39 for group 4.	Not stated.
Kay et al 2008	To assess the efficacy, safety, and pharmacology of subcutaneous administration of golimumab in patients with active rheumatoid arthritis (RA) despite treatment with methotrexate (MTX).	Simulations were performed to evaluate the power of the chi-square test to detect a significant treatment effect for the combined golimumab plus MTX groups versus the placebo plus MTX group and at least 1 individual golimumab dose group versus the placebo plus MTX group (_ _ 0.05, 2-sided test).	Assuming that 60% of golimumab-treated patients and 25% of placebo-treated patients achieved the primary end point, the study required 35 patients in each treatment group to achieve _90% power.	For the primary analysis, a last observation carried forward procedure was used for patients who did not return for an evaluation or for whom we had insufficient data to determine their ACR20 response. Patients who initiated treatment with oral

		The primary analysis was conducted using a 2-sided chi-square test comparing the combined golimumab plus MTX treatment groups with the placebo plus MTX group. If a statistically significant difference (_ _ 0.05) was evident in favor of the combined golimumab plus MTX groups, pairwise comparisons between each individual golimumab dose group and the placebo group were to be performed separately.		corticosteroids or disease-modifying antirheumatic drugs (other than MTX but including biologics), increased MTX or oral corticosteroid dosages above baseline levels, or discontinued the study agent because of lack of efficacy before week 16 were considered to have not achieved the primary end point at week

ATTEST Schiff et al 2008	This trial evaluated the efficacy and safety of abatacept or infliximab vs placebo. The primary objective of this study was to evaluate the mean change from baseline in Disease Activity Score (based on erythrocyte sedimentation rates; DAS28 (ESR)) for the abatacept vs placebo groups at day 197.	All patients who received at least one dose of study medication were assessed for efficacy and safety (intent-to-treat population). At day 197, the abatacept or infliximab groups were compared with the placebo group by analyses of covariance (ANCOVA) for mean changes from baseline in DAS28 (ESR) and in the SF-36 (PCS and MCS). The model included the change as the dependent variable, with treatment group as a main effect and the baseline score as an additional covariate. The proportion of patients with ACR 20, 50 and 70 responses, LDAS, DAS28-defined remission, a good EULAR response and a clinical meaningful HAQ-DI response was calculated. The x2 test was performed to evaluate the differences (and 95% CIs) between the abatacept or infliximab groups and placebo. At day 365, the reference group was changed to infliximab. tabulated by treatment group at days 197 and 365	The sample size and power were calculated to detect a treatment difference in the primary analysis of a mean change from baseline in DAS28 (ESR) for the abatacept vs placebo groups at day 197. Prospectively, this study was not powered with a superiority or non-inferiority design to compare the two active arms.	Patients who discontinued the study prematurely were considered as non-responders subsequent to the time of discontinuation for ACR 20, 50 and 70 responses, good ULAR responses and clinically meaningful HAQ-DI responses. For all continuous measurements (mean changes in DAS28, SF- 6 and the HAQ-DI score), LDAS and DAS28-defined remission he last observations prior to the discontinuation were carried orward (LOCF).
ATTRACT Lipsky et al 2000	See below	Pairwise comparisons of the infliximab and placebo groups were made when the overall effect of treatment had a significant (P< 0.025) effect on the primary end point — a clinical response. We used the chi-square test to evaluate categorical variables and analysis of variance to evaluate continuous variables. The proportion of patients who had a response was analyzed by chi-square test, and we used Fisher's exact tests for pairwise comparisons of adverse effects. For	As below	As below

		continuous variables, we made pairwise comparisons using linear contrasts. All statistical tests were two-sided.		

ATTRACT Maini et al 1999		All statistical testing began with a test for an overall treatment effect across the five treatment groups. Categorical variables were tested by the _2 test, and continuous variables were tested by analysis of variance on the van der Waerden normal scores. Pairwise analyses comparing infliximab treated groups with the placebo group were only done if the test for an overall treatment effect was significant (p<0·05). Pairwise testing of categorical efficacy endpoints used the _2 test, and pairwise testing of categorical safety endpoints used Fisher's exact test. For continuous variables, pairwise comparisons of infliximab groups versus placebo were made with linear contrasts. All statistical testing was two-sided.	The sample size of about 80 patients per treatment group provided more than 90% power to detect a difference in proportions between treatment groups by use of the two-sided _2 test at _=0·01, where the methotrexate-alone group was predicted to have a 20% clinical response rate and the infliximab group with the highest response was predicted to have a 65% clinical response rate (the remaining infliximab treatment groups were predicted to be midway between these two extremes).	Not stated
START Westhoven s et al 2006b	To assess the risk of serious infections following 22 weeks of infliximab therapy, and to further characterize the safety profile of infliximab in combination with background treatments during 1 year in patients with rheumatoid arthritis (RA) with various comorbidities.	The primary end point of the study was the proportion of patients who reported experiencing a serious infection within the first 22 weeks after initiating therapy. The Mantel-Haenszel chi-square test, stratified by baseline corticosteroid use (no corticosteroid use or any corticosteroid use), was used to analyze the data for serious infections. Other categorical data were analyzed using the chi-square test. Continuous variables were compared using an analysis of variance on the van der Waerden normal scores. All statistical tests were 2-sided, with _ _ 0.05.	A 1-sided equivalence test model supported a sample size of 334 patients in the placebo group and 666 patients evenly distributed in the 2 infliximab groups. Thus, with 1,000 patients, the study had _80% power, at a 5% significance level, to rule out a 2-fold increase in serious infections based on the assumption that the rates are the same in the placebo group and the combined infliximab group. The presumed rate of infection for the placebo group was 6%, which was the rate reported in patients receiving placebo plus	Not stated

			MTX for 30 weeks in a previous trial.	

Abe et al 2006	A placebo controlled, double-blind trial (DBT) was conducted for Japanese patients with active rheumatoid arthritis (RA) despite treatment with low dose methotrexate (MTX) to evaluate the efficacy and safety of infliximab. Extended treatment with infliximab was conducted in an openlabel trial (OLT).	The analysis set of demographics and efficacy was the full analysis set. The analysis set for safety consisted of patients who received at least one infusion of the study drug. Demographics across treatment groups were analyzed using the chisquare test for categorical data, the Kruskal-Wallis test for ordered categorical data, and ANOVA for quantitative data. Response rates between treatment groups, based on the ACR criteria, were analyzed using logistic regression.	not stated.	In the DBT, patients who discontinued treatment before Week 14 received assessments at discontinuation as the primary endpoint. For other efficacy values, assessments up to discontinuation were adopted, but assessments after discontinuation were removed. For the efficacy values of patients discontinuing the OLT, assessments up to discontinuation were adopted and assessments at discontinuation were carried as those after discontinuation.
Maini et al 1998	To evaluate the efficacy, pharmacokinetics, immunogenicity, and safety of multiple infusions of a chimeric monoclonal anti-tumor necrosis factor ***a*** antibody (cA2) (infliximab; Remicade, Centocor, Malvern, PA) given alone or in combination with lowdose methotrexate (MTX) in rheumatoid arthritis **(RA)** patients.	Demographic variables and total time of response were compared using an analysis of variance on the van der Waerdcn normal scores, which had bcen blocked by investigational site. The proportions of patients responding to treatment according to the Paulus criteria, as well as the categorical demographic and safety variables, were compared among treatment groups using the Cochran-Mantel-Haenszel chisquare test for general association stratified by investigationalsite. Analyses comparing each of the cA2 treatment groups with thc placebo infusion plus MTX group (control) were performed only when the overall treatment effect P value was <0.05. The Mann-Whitney test was used to compare serum cA2 concentrations for each dosage of cA2 with and without MTX. All computations were performed using SAS software (SAS Institute, Cary, NC). All *P* values are 2-sided.	Not stated.	For the purposes of this analysis, patients who were unable to complete the 26 weeks of the trial for any reason (e.g., discontinuation at their own or their physician's request or as a result of an adverse event, or if an increase in the dosage of MTX or corticosteroids or treatment with a new DMARD was required) were considered nonresponders from the day of withdrawal from the study.

Edwards et al 2004	An open-label study indicated that selective depletion of B cells with the use of rituximab led to sustained clinical improvements for patients with rheumatoid arthritis. To confirm these observations, we conducted a randomized, double-blind, controlled study.	The primary analyses were based on the intention-to-treat principle. For patients who withdrew before week 24, a last-observation-carried-forward method of imputation was applied. Statistical analyses (with the two-sided Fisher's exact test) were performed only for comparisons of each rituximab group with the control group. Exploratory secondary analyses were performed for ACR response rates at week 48 with use of a nonresponder imputed rule for all patients who withdrew before that time. Roche was the study sponsor and was responsible for data collection. Statistical analyses were conducted by suitably qualified statisticians who were employees of the sponsor.	Sample-size calculations were based on the assumption that the proportion of patients continuing to receive only methotrexate and achieving an ACR 50 response at week 24 would be 5 percent and that the proportion of patients in any of the rituximab treatment groups would be 30 percent. On the basis of these assumptions and with the use of Fisher's exact test with a two-sided significance level of 0.05, we calculated that a sample of 40 patients per treatment group would provide the study with 82 percent power to detect a difference between the two proportions.	The primary analyses were based on the intention-to-treat principle. For patients who withdrew before week 24, a last-observation-carried-forward method of imputation was applied. Statistical analyses (with the two-sided Fisher's exact test) were performed only for comparisons of each rituximab group with the control group. Exploratory secondary analyses were performed for ACR response rates at week 48 with use of a nonresponder imputed rule for all patients who withdrew before that time.
Strand et al 2006	To evaluate the long-term impact on physical function of a single course of rituximab in rheumatoid factor, seropositive patients with active rheumatoid arthritis (RA) despite ongoing methotrexate treatment.	Descriptive statistics were used to compare baseline demographic and disease parameters in the initial intent-to-treat (ITT) population with those completing either 1 or 2 yrs of protocol participation. ACR responses were calculated using the ITT population. The significance of the change from baseline was determined by analysis of variance using all available data.	Retrospective analysis	Patients with insufficient data to calculate an ACR response and patients who withdrew from protocol participation were classified as non-responders.
REFLEX Cohen et al 2006	To determine the efficacy and safety of treatment with rituximab plus methotrexate (MTX) in	Efficacy analyses were performed on the intent-to-treat (ITT) population, defined as all randomized patients who received any part of an infusion of study medication,.	Sample size calculations were based on the assumption of detecting a difference in the proportion of patients with an	Patients who withdrew prematurely from the study or who started rescue therapy were included in the ITT population as nonresponders.

	patients with active rheumatoid arthritis (RA) who had an inadequate response to anti–tumor necrosis factor (anti-TNF) therapies and to explore the pharmacokinetics and pharmacodynamics of rituximab in this population.	Patients who withdrew prematurely from the study or who started rescue therapy were included in the ITT population as nonresponders.	ACR20 response in the rituximab plus MTX (0.45) and the placebo plus MTX (0.30) groups. On the basis of these assumptions and using a conservative exact test (Fisher's) with a 2-sided 5% significance level, a sample size of 500 patients randomized to rituximab and placebo at a ratio of 3:2 provided the study with 91% power to detect a difference between the treatments.	

REFLEX Keystone et al 2008a	As above (REFLEX)	As above (REFLEX)	As above (REFLEX)	As above (REFLEX)
OPTION Smolen et al 2008	The aim was to assess the therapeutic eff ects of blocking interleukin 6 by inhibition of the interleukin-6 receptor with tocilizumab in patients with rheumatoid arthritis.	To control for false positive conclusions for the primary endpoint, the tocilizumab 8 mg/kg arm was fi rst compared with the placebo arm, and a p value was derived. Only if this comparison gave a p value of 0·05 or less was a comparison of the tocilizumab 4 mg/kg with the placebo arm made. For the secondary endpoints, a prespecifi ed fi xed sequence approach was applied, allowing us to test each of the null hypotheses at the same signifi cance level of α without any adjustment. ACR20 response at 24 weeks was compared between treatment groups with a Cochran-Mantel-Haenszel χ^2 test with adjustment for site.	A sample size of 210 patients per arm (630 patients) was calculated to provide 90% power to detect a diff erence between tocilizumab and placebo (two-sided test, corrected for multiple comparisons), assuming ACR20 responses of 60% with the study drug versus 40% with placebo, allowing for a 15% dropout rate.	Patients who withdrew before week 24, patients who received rescue therapy, and patients whose week 24 categorical endpoints could not be determined due to insufficient data were deemed to be non-responders in the analysis.
RADIATE Emery et al 2008	The phase III RADIATE study examined the efficacy and safety of tocilizumab, an anti-IL-6 receptor monoclonal antibody in patientswith rheumatoid arthritis (RA) refractory to tumour necrosis factor (TNF) antagonist therapy.	Primary endpoint analysis was performed on all participants receiving one or more administration of study treatment (the intent to treat (ITT) population). Safety data are presented using the safety population, comprising all ITT patients with one or more postrandomisation assessments of safety. Primary endpoint analysis for ACR20 response (with secondary analyses for ACR50/70, DAS28 and European League Against Rheumatism	A sample size of 450 patients was calculated to provide more than 80% power to detect a difference of 20 points between tocilizumab and control arms at week 24 for the ACR20 response and to enable the reporting of safety and efficacy for this unique patient population for registration	Patients on rescue therapy or with insufficient data to calculate the change from baseline ACR score at a specific time point were classified as non-responders at that time point.

		(EULAR) responses) compared the proportion of patients in each of the tocilizumab plus methotrexate groups versus controls with a response at week 24 using a Cochrane–Mantel–Haenszel x2 test with adjustment for site.		

SATORI Nishimoto et al 2009	The clinical efficacy and safety of tocilizumab (a humanized anti-IL-6 receptor antibody) monotherapy in active rheumatoid arthritis (RA) patients with an inadequate response to low dose methotrexate (MTX).	The primary end point was the ACR20 response at week 24 with the last observation carried forward (LOCF) method, using an intent-to-treat (ITT) analysis. The incidences of clinical improvements were analyzed by the chi-square test. All statistical analyses were two-sided and P values less than 0.05 were considered significant. All patients receiving at least one dose of tocilizumab or tocilizumab placebo, and at least 4 weeks of MTX or MTX placebo administration were included in the clinical efficacy analysis.	We determined that a sample size of 57 patients per treatment group was required to provide 90% power for detecting a significant (P\0.05) difference in ACR20 response between the control group and the tocilizumab group by use of the two-side chi-square test, where ACR20 response rates in the population were assumed to be 35 and 65% in the control group and the tocilizumab group, respectively. 60 patients were recruited per treatment group to allow for anticipated withdrawals.	Not stated.
TOWARD Genovese et al 2008	To examine the efficacy and safety of the humanized anti–interleukin-6 receptor antibody tocilizumab combined with conventional diseasemodifying antirheumatic drugs (DMARDs) in patients with active rheumatoid arthritis (RA).	The primary end point, the proportion of patients with an ACR20 response at week 24, was compared using a Cochran-Mantel-Haenszel chi-square test with adjustment for site; this methodology was also used for the ACR50/70 response, DAS28 remission, and the EULAR response. Changes from baseline in the individual ACR core set parameters the DAS28, the hemoglobin concentration, and FACIT-F and SF-36 scores were summarized by descriptive statistics. The difference between treatment groups for each component at week 24 was compared using an analysis of variance model with adjustment for	Evaluation of at least 1,200 patients was planned to provide a sufficient number of patients to examine the safety of conventional DMARDs administered in combination with tocilizumab. The sample size with 2:1 randomization provided _90% power to detect an efficacy difference between the tocilizumab and control arms at week 24.	Patients who did not have the required data for a specific time point, who withdrew from the study, or who received rescue therapy were classified as nonresponders.

		site.		

Figure 7 Participant Flow

Where graphs were available, they have been presented below. However a small number of papers did not include enough information in the results to present a completed CONSORT graph, and therefore text is used to describe the patient flow of the trials instead.

Kremer et el 2005

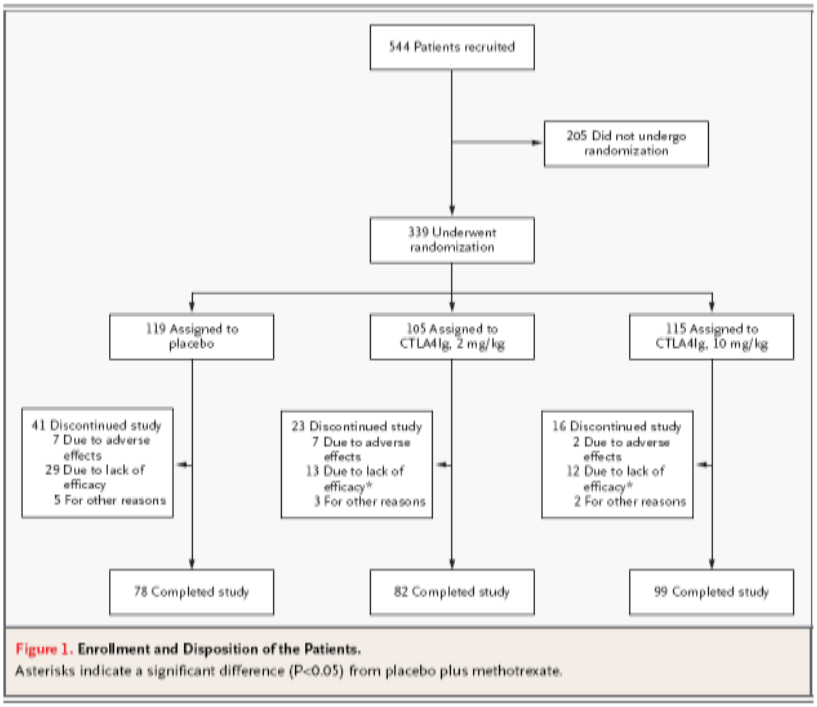


Figure 1. **Enrollment and Disposition of the Patients.**
Asterisks indicate a significant difference (P<0.05) from placebo plus methotrexate.

AIM Kremer 2006

Baseline characteristics. Three hundred thirty-nine patients were recruited and randomly assigned to t

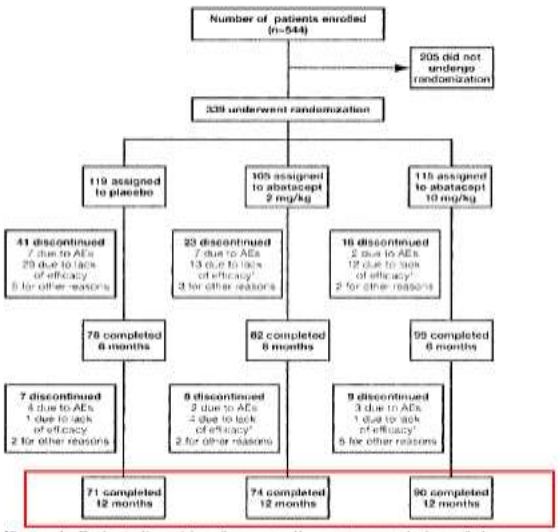
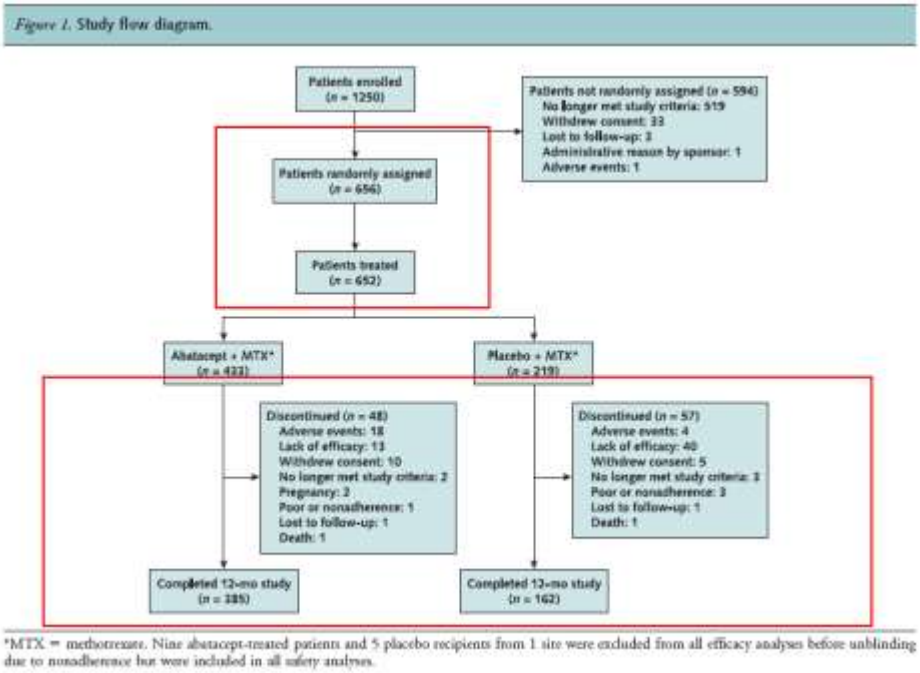


Figure 1. Patient disposition from enrollment to completion of the 12-month study. * = $P < 0.05$ versus placebo plus methotrexate. AEs = adverse events. Adapted, with permission, from ref. 11.

Kremer et al 2006



Weinblatt et al 2006 ASSURE

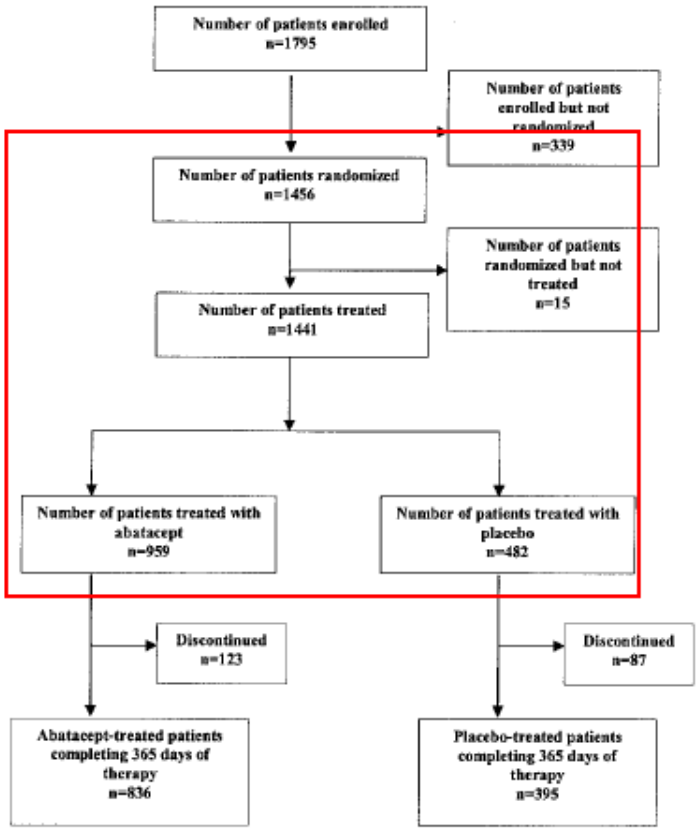


Figure 1. Patient disposition to 1 year.

Genovese et al 2005

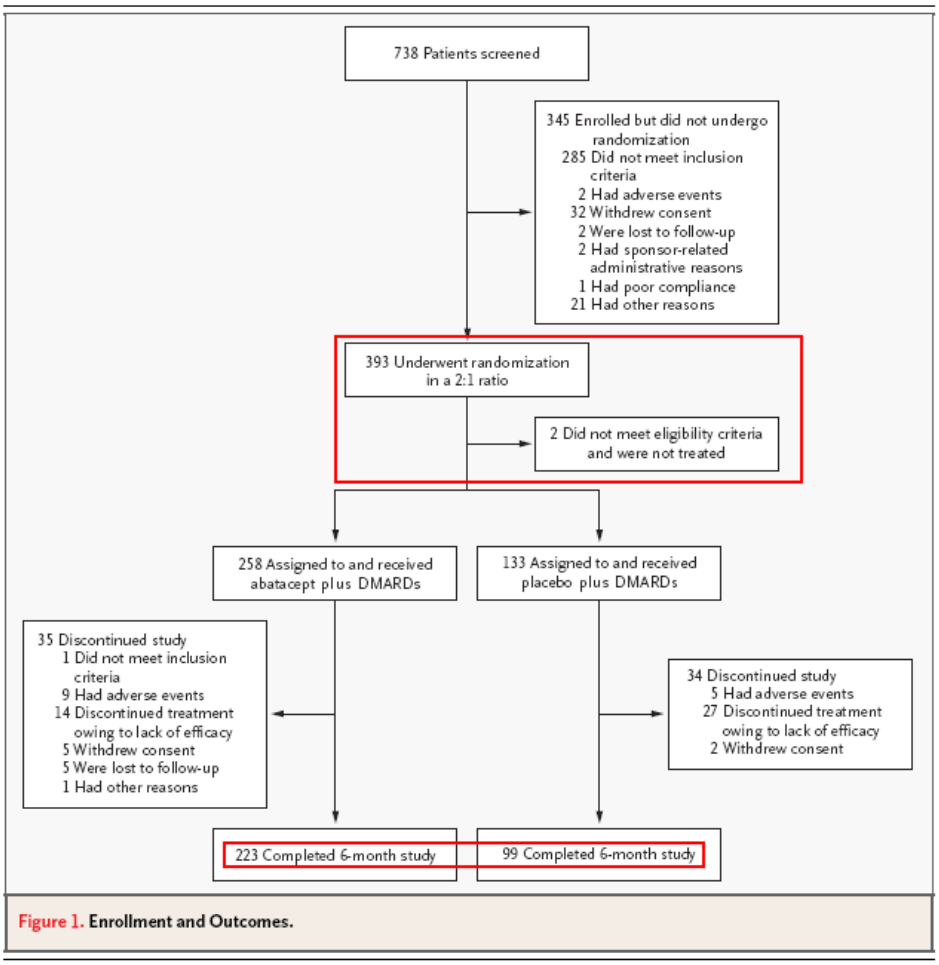


Figure 1. Enrollment and Outcomes.

Kim et al 2007

A total of 128 patients were enrolled at six sites. Sixtythree patients were randomized to the placebo treatment group, and 65 patients were randomized to the adalimumab treatment group. All 128 patients received at least one injection of the study drug and were thus included in the ITT analysis set (Table 1). Fifty-one of 65 patients (78.5%) randomized to adalimumab and 40 of 63 patients (63.5%) randomized to placebo completed the 24-week, double-blind period without openlabel rescue treatment. Eight of 65 patients (12.3%) randomized to adalimumab and 19 of 63 patients (30.2%) randomized to placebo had insufficient response, switched to rescue arm, and completed the 24-week study on treatment with open-label adalimumab.

Van de Putte et al 2004

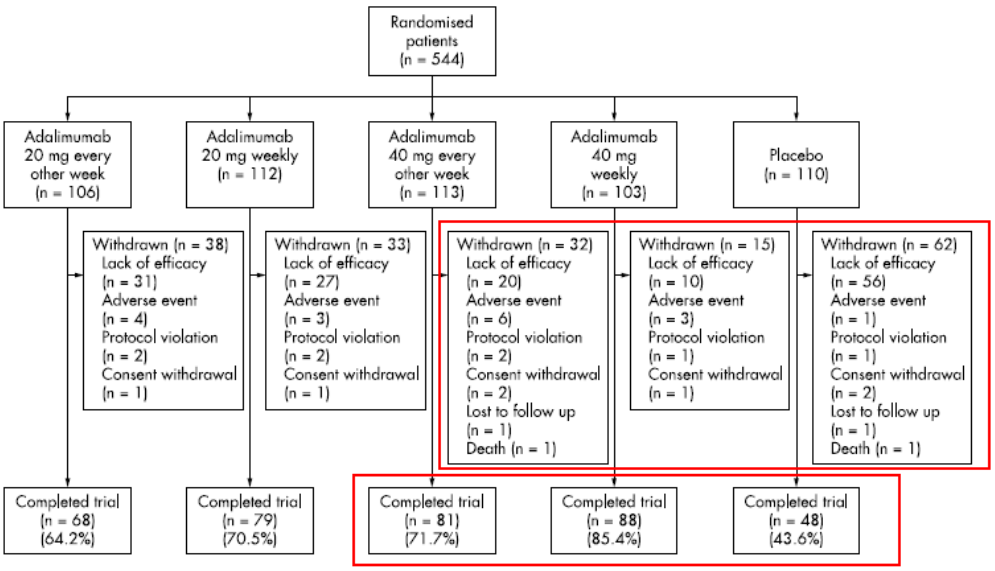
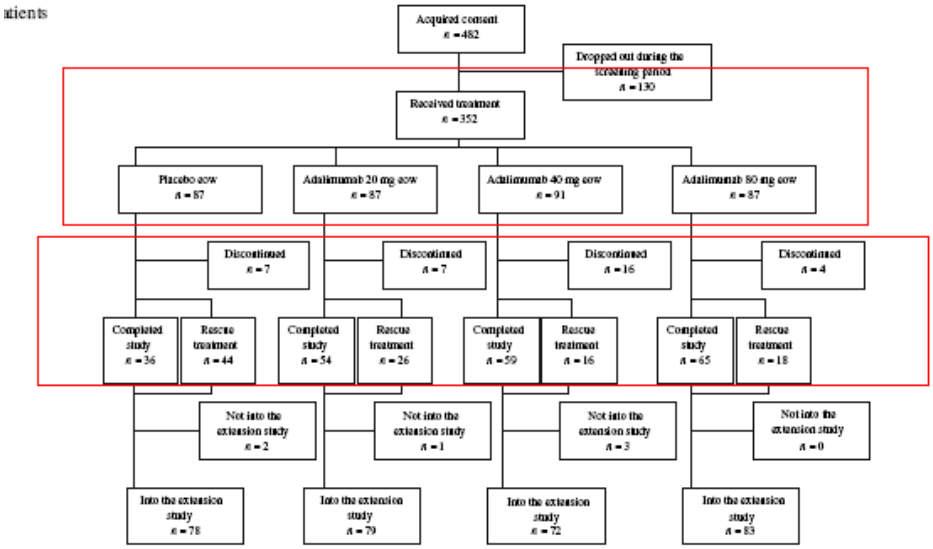


Figure 1 Patient disposition.

Weinblatt et al 2003 ARMADA

A total of 336 RA patients were screened, and 271 patients met the entry criteria and were randomized to 4 treatment groups: 62 (22.9%) in the placebo group, 69 (25.5%) in the 20-mg adalimumab group, 67 (24.7%) in the 40-mg adalimumab group, and 73 (26.9%) in the 80-mg adalimumab group. Among the 271 patients that entered the study, 161 completed the 24 weeks. Ninety-two patients who did not achieve an ACR20 response elected to enter the open-label continuation study between weeks 16 and 24. Of these 92 rollover patients, 23, 27, and 27 were in the adalimumab 20 mg, 40 mg, and 80 mg groups, respectively, and 35 were in the placebo group. In addition, 18 patients withdrew from the study prematurely because of adverse events (n _ 7), withdrawal of consent (n _ 5), lack of efficacy (n _ 3), protocol violation (n _ 1), or loss to followup (n _ 2). All of the 18 patients who withdrew from the study did so before week 16 and were not eligible to roll over into the open-label extension trial.

Miyasaka 2008 CHANGE



Keystone et at 2004

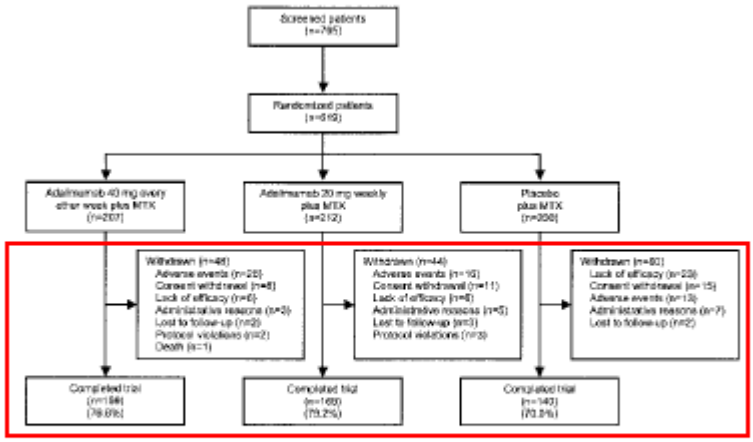


Figure 1. Profile of trial involving treatment with adalimumab (a human anti–tumor necrosis factor monoclonal antibody) as compared with placebo in patients with active rheumatoid arthritis receiving concomitant methotrexate (MTX) therapy.

STAR Furst et al 2003

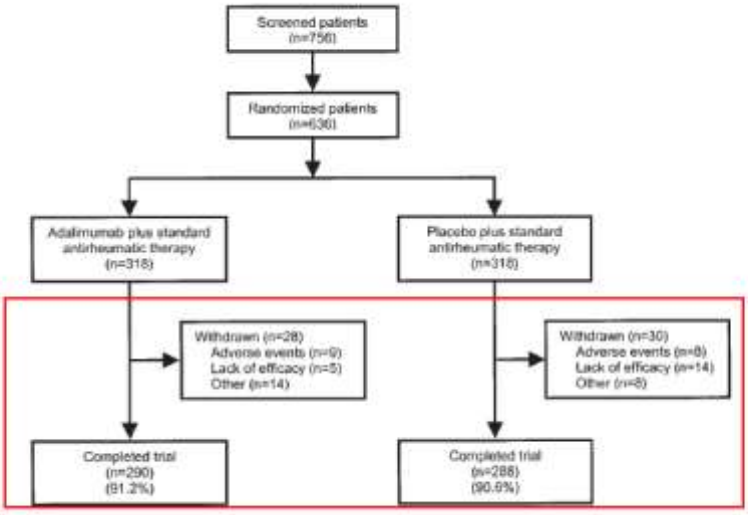


Figure 1. Patient disposition

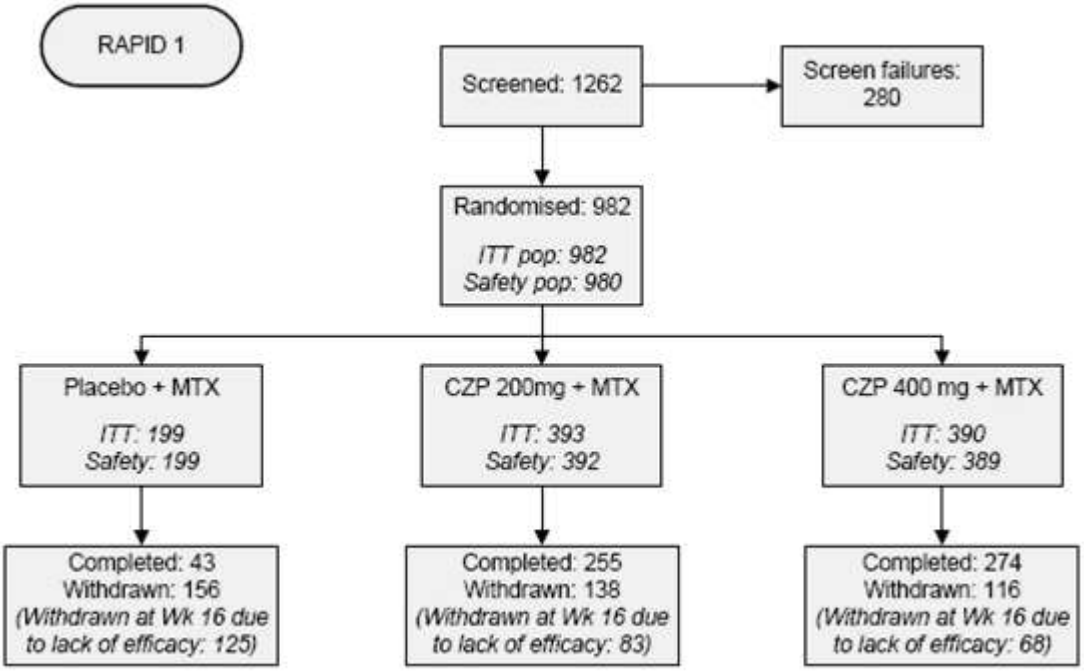
Chen et al 2009

Forty-seven RA patients who were enrolled in the ITT population were randomized into treatment groups (35 in the adalimumab plus MTX group and 12 in the MTX alone group).

Severe AEs were reported in 14.3% of adalimumab-treated subjects compared to 8.3% of placebo-treated subjects, and AEs leading to discontinuation occurred in 8.6% of adalimumab-treated subjects compared to none of the placebo-treated subjects. Of the six severe AEs reported in the adalimumab group, three were of infectious etiology (1 case of TB, 1 of pneumonia, 1 of sinusitis). No deaths, immunological reactions or malignancies were reported during the study.

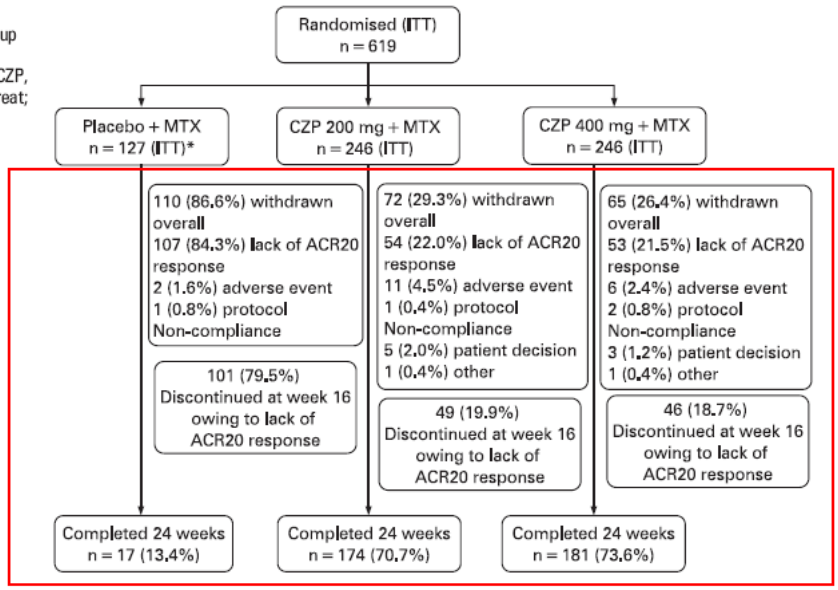
RAPID 1 Keystone et al 2008b, RAPID 1 Strand et al 2009

Figure 4: Patient numbers in RAPID 1

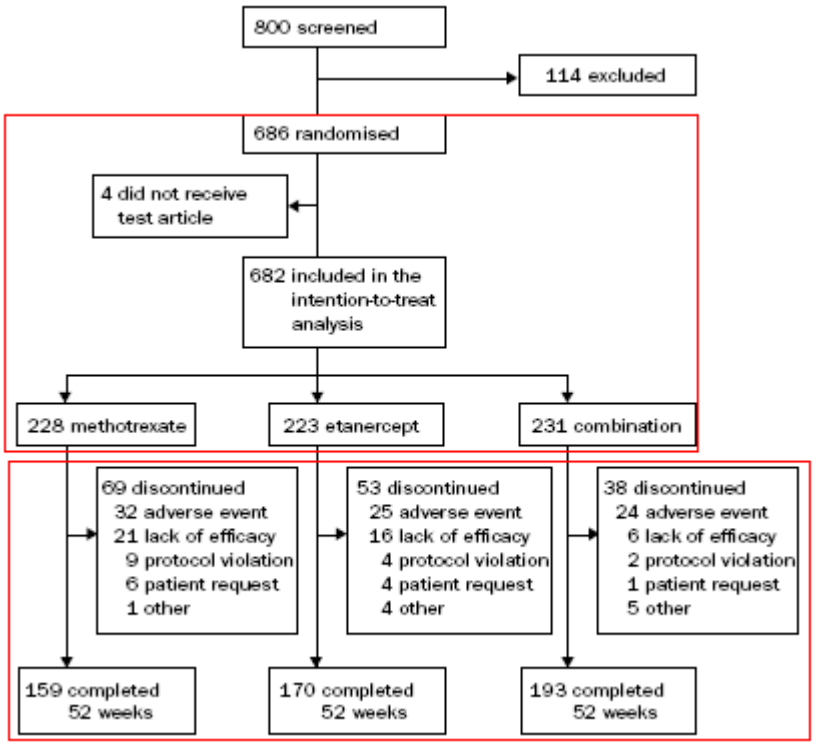


Smolen et al 2009b RAPID 2

Figure 1 Patient disposition. *Two patients in the placebo treatment group received CZP and are included in the 200 mg group for safety evaluations. CZP, certolizumab pegol; ITT, intention-to-treat; MTX, methotrexate.



TEMPO Klareskog et al 2004



Combe et al 2006

Of the total of 260 patients who were randomly assigned in the study, 254 received at least one dose of the study treatment (sulfasalazine, n=50; etanercept, n=103; and combination, n=101). A total of 221 (87%) patients completed the study. Unsatisfactory response to treatment, the most common primary reason for discontinuation, was reported by more patients receiving sulfasalazine alone (24%) than by those receiving etanercept alone (1%) or etanercept and sulfasalazine (4%; p,0.001, sulfasalazine v etanercept or combination therapy). We found no significant difference in the percentage of patients (6% sulfasalazine, 6% etanercept, 1% combination) who withdrew because of adverse events.

Moreland et al 999

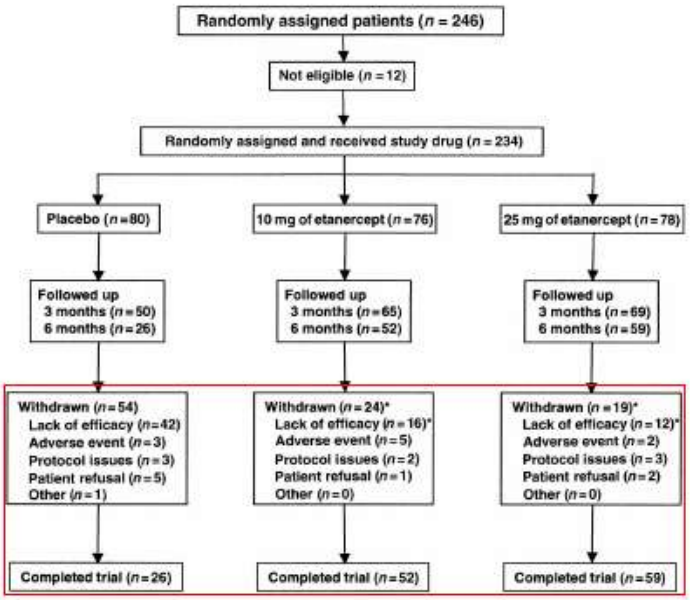
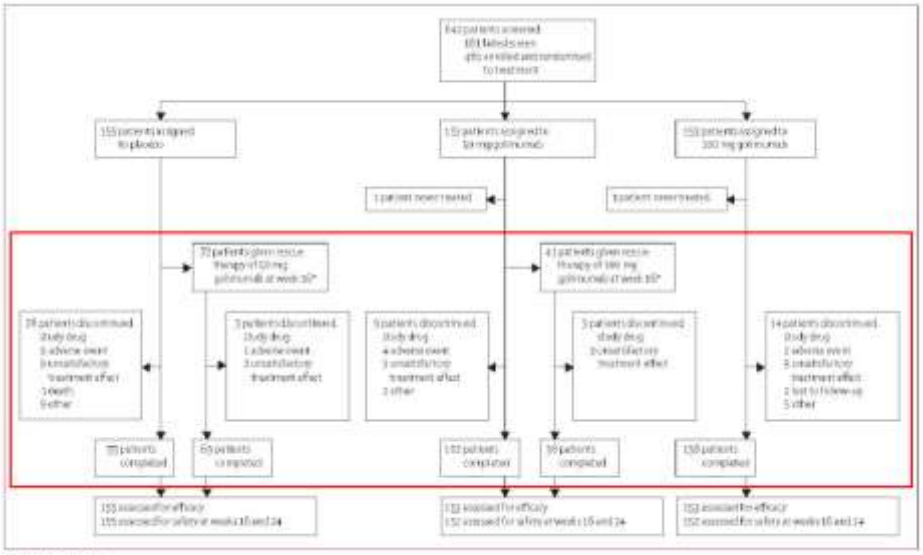


Figure 1. Study completion and withdrawal before 6 months. *$P < 0.001$ for each etanercept group compared with the placebo group (likelihood ratio chi-square test).

Weinblatt et al 1999

Fourteen men and 75 women were enrolled. Of the 59 patients randomly assigned to receive etanercept plus methotrexate, 57 (97 percent) completed the 24-week study and 2 withdrew because of adverse events unrelated to etanercept (abdominal pain due to an incisional hernia from prior surgery in 1 patient, and traumatic fractures of the shoulder and calcaneus in the other). Of the 30 patients randomly assigned to receive placebo plus methotrexate, 24 (80 percent) completed the study, 4 withdrew because of lack of efficacy, 1 had a myocardial infarction, and 1 was lost to follow-up. All the patients received at least one dose of study drug and could be included in the evaluation of the safety and efficacy of the treatment. The mean number of doses of study drug received was 47 in the etanerceptplus- methotrexate group and 43 in the placeboplus-methotrexate group.

GO AFTER Smolen et al 2009a



GO FORWARD Keysone et al 2009b

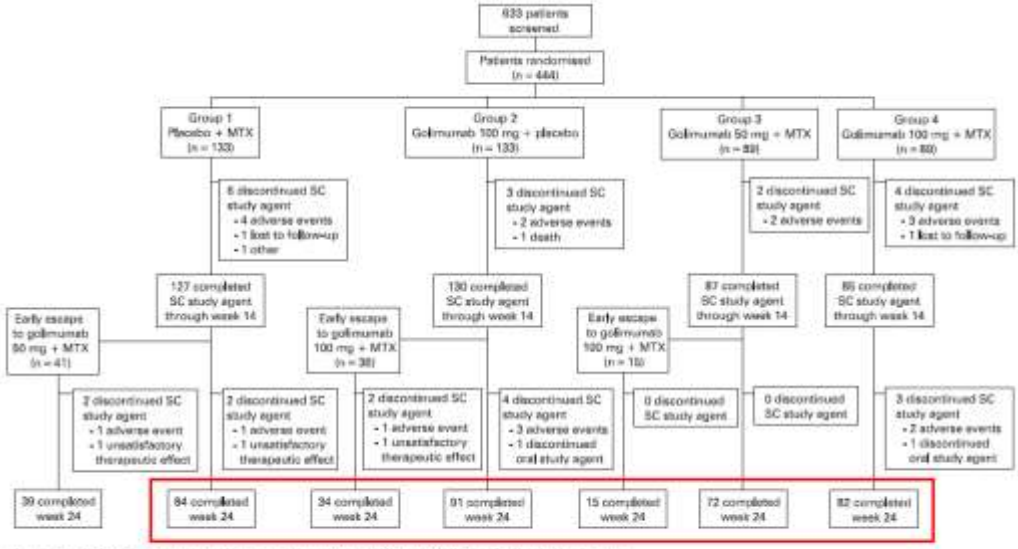


Figure 1 Disposition of patients during the study. MTX, methotrexate; SC, subcutaneous.

Kay et al 2008

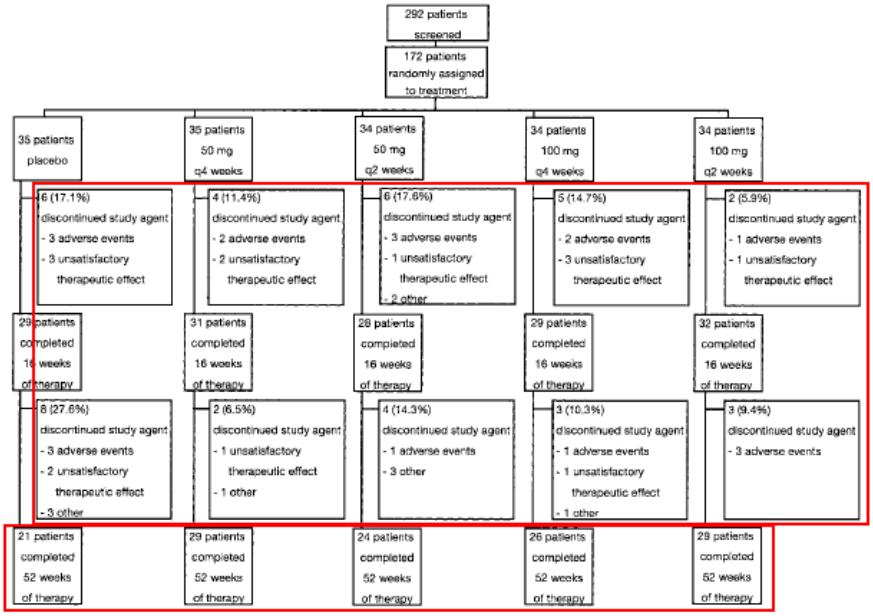
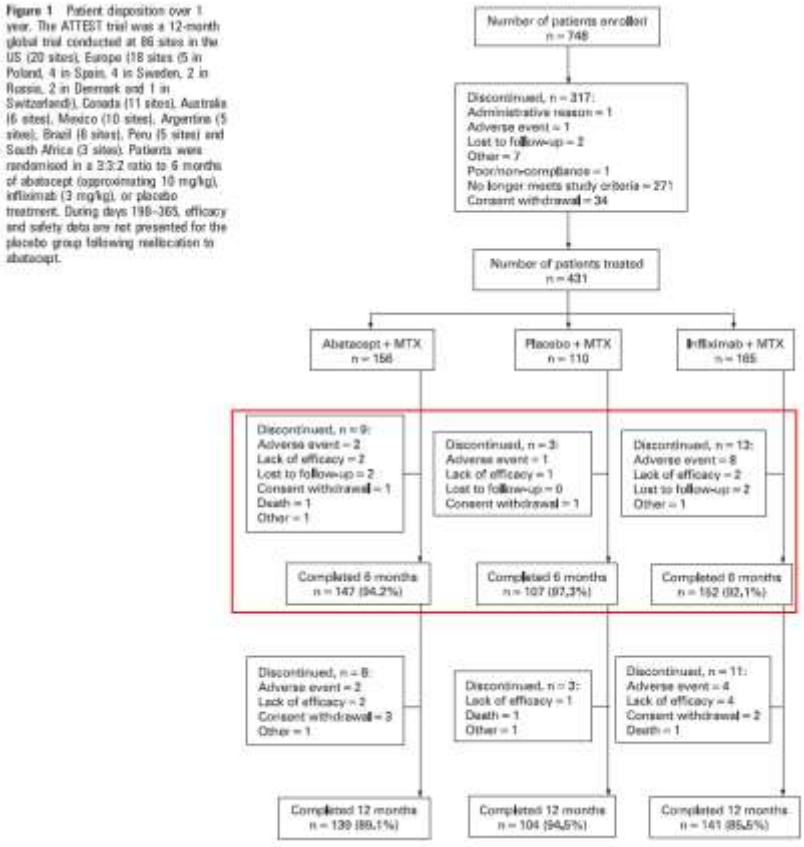


Figure 1. Disposition of patients during the study. q4 weeks = every 4 weeks.

ATTEST Schiff et al 2008



ATTRACT Lipsky et al 2000

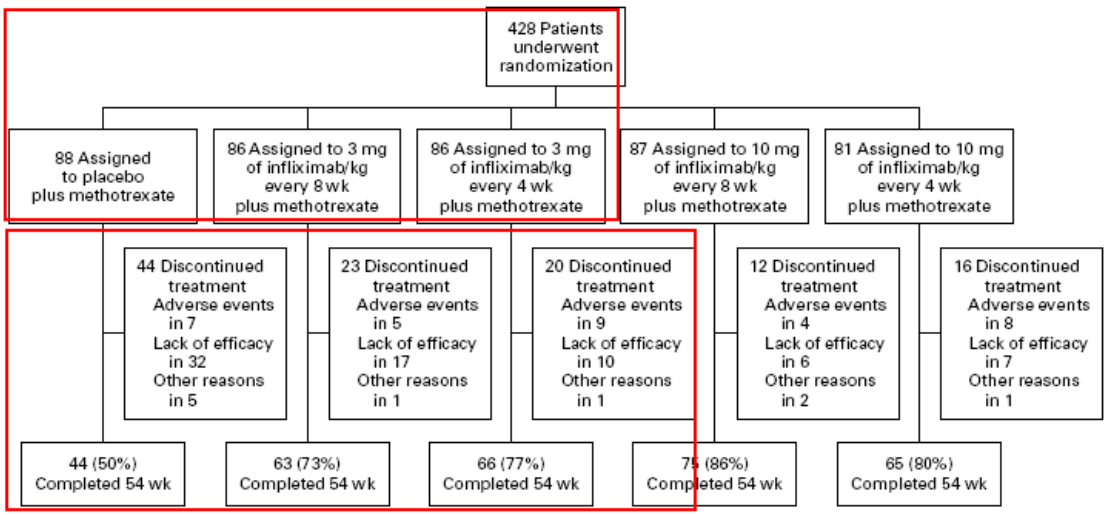


Figure 1. Randomization, Reasons for Discontinuing Treatment, and the Numbers of Patients Who Completed the Trial. Other reasons for discontinuing treatment included withdrawal of consent and withdrawal because of noncompliance.

ATTRACT Maini et al 1999

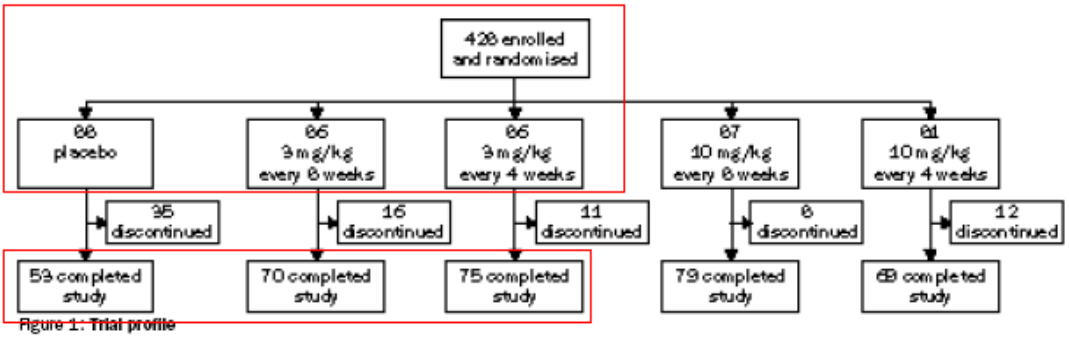


Figure 1: Trial profile

START Westhovens et al 2006b

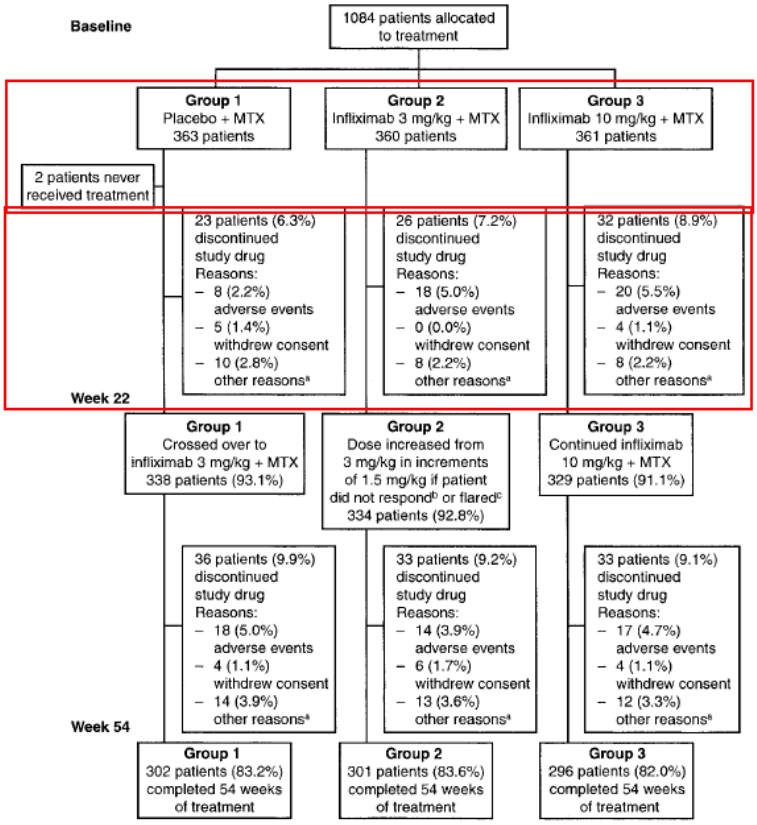


Figure 1. Randomization scheme and distribution of patients in the safety study of infliximab. All patients received concomitant methotrexate (MTX).

Abe et al 2006

Out of 151 patients enrolled in the DBT, 147 received at least one infusion of study drugs (47,49, and 51 patients in the placebo, 3 mg/kg, and 10 mg/kg groups, respectively). Five patients receiving the placebo discontinued treatment, including 3 due to lack of efficacy, one due to an adverse event, and one due to a protocol violation. Five patients receiving infliximab discontinued treatment due to adverse events. Baseline demographics were comparable among the 3 groups, with the exception of body weight. The difference had no influence on the result of the primary endpoint using covariance adjustment. The mean dose of MTX was 7.2 ± 2.0 mg/week. The doses of MTX among the treatment groups were well balanced. A large number of patients were treated with NSAID and corticosteroid concomitantly.

Maini et al 1998

One hundred one eligible patients were enrolled from 6 centres. Within 1 week of scrccning, they were allocated by randomization to 1 of 7 groups: 4 groups received either 1, 3, or 10 mg/kg of cA2 or a placebo infusion concomitantly with 7.5 mg/wcck of oral MTX, and 3 groups received 1,3, or 10 mgikg of cA2 with placebo tablets. Randomization was performed centrally, and the nature of the codcd study medications was not revealed to the paticnts or the assessors. Prior to the first infusion (day 0, week O), repeat baseline measurcments of clinical and laboratory parameters were made. Patients returned to the study center for followup assessments at wecks 1 and 2, then every 2 weeks until week 22, and then for a final visit at wcck 26. Infusions of cA2 or placebo were repeated (aft.er the assessments wcre made) at 2, 6, 10, and 14 weeks. The proportion of patients who discontinued treatment because of lack of efficacy was highest in the placebo infusion plus MTX and 1 m a g cA2 without MTX groups (57% and ***33%,*** respectively), and was notably lower in the groups receiving ***3*** and 10 mg/kg of cA2 alone (7% and 13%) or the 3 dosages of cA2 plus MTX (0%, 095, and 7%). These findings support the conclusion that the most consistent response rates were obtained in patients who were receiving 3 or 10 mg/kg of cA2 alone or a combination of cA2 at any dosage plus MTX.

Edwards et al 2004

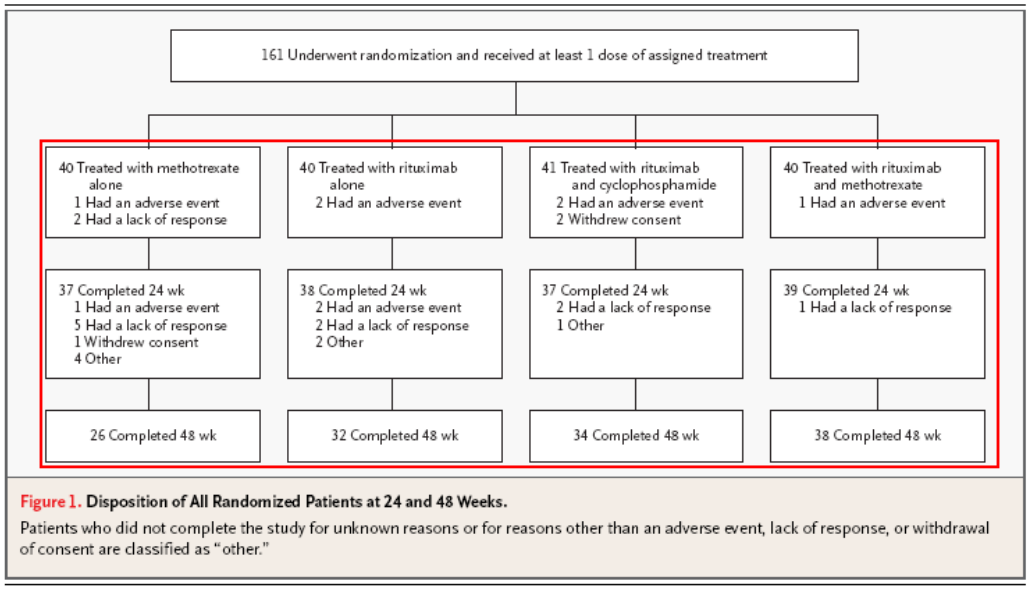


Figure 1. Disposition of All Randomized Patients at 24 and 48 Weeks.
Patients who did not complete the study for unknown reasons or for reasons other than an adverse event, lack of response, or withdrawal of consent are classified as "other."

Strand et al 2006

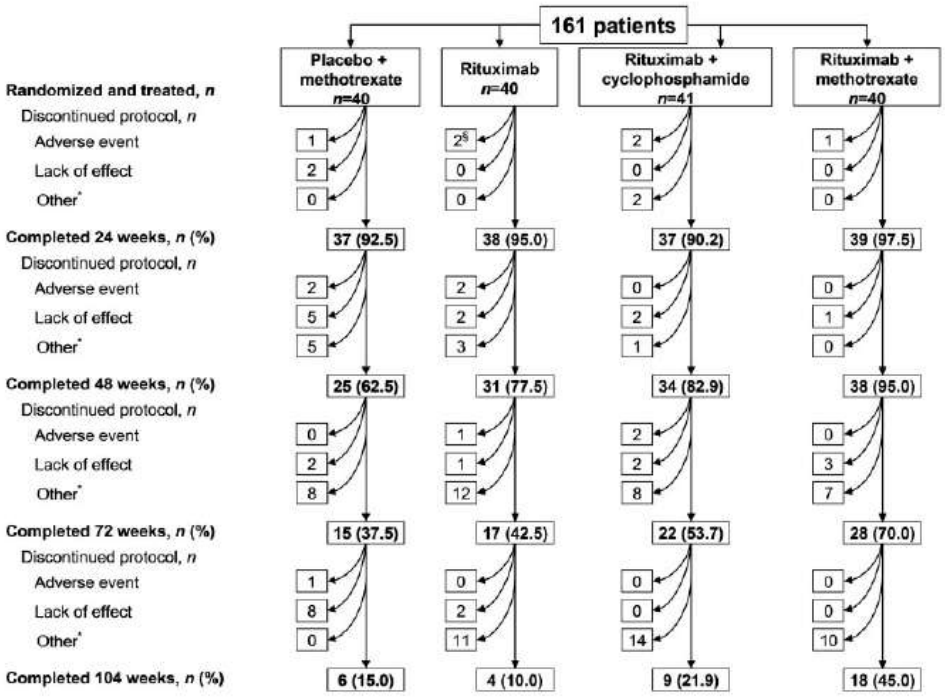


FIG. 1. Disposition of patients over 2 yrs. *The majority of protocol discontinuations classified as 'other' were owing to the requirement for retreatment with rituximab under a separate protocol; §includes 1 death due to pneumonia.

REFLEX Cohen et al 2006

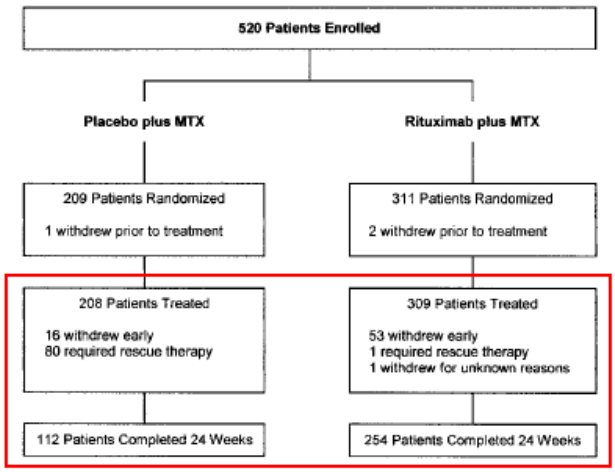


Figure 2. Disposition of patients from enrollment to week 24. With regard to the number of patients treated, 1 patient who was randomized to the rituximab group actually received placebo and was therefore included in the placebo subgroup for the safety analysis, resulting in 209 patients receiving placebo plus methotrexate (MTX) and 308 patients receiving rituximab plus MTX.

OPTION Smolen et al 2008

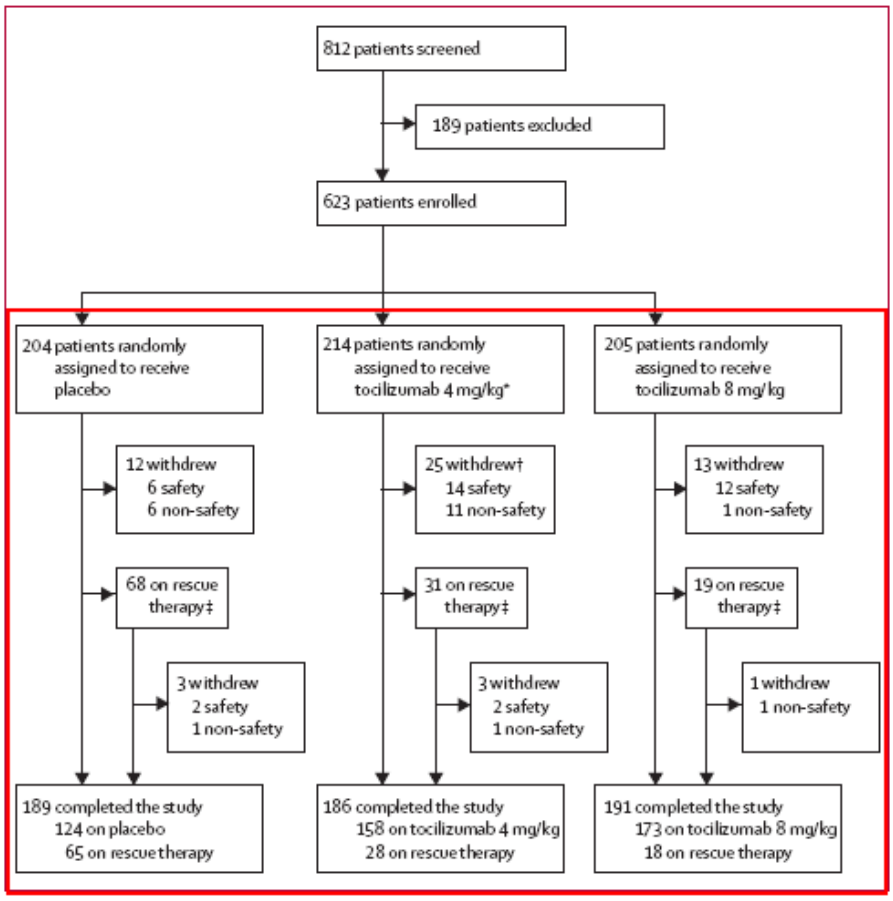
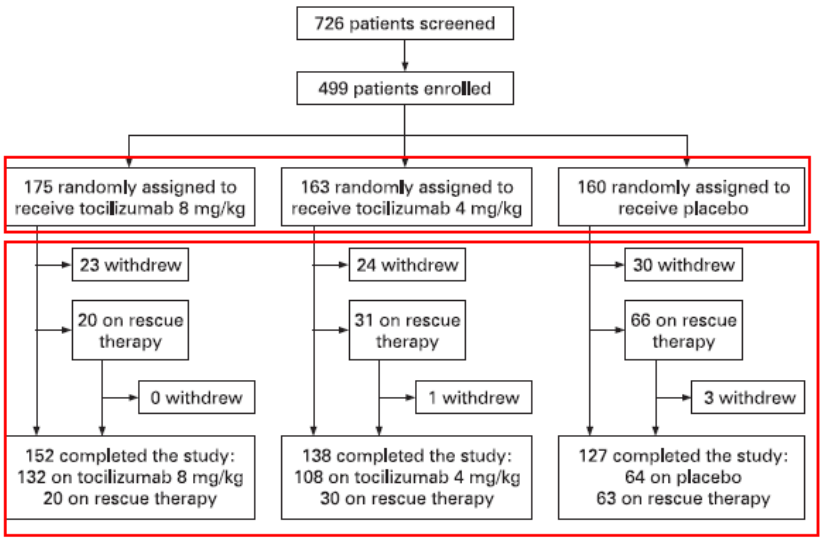


Figure 1: Trial profile
*Includes one patient who was randomised to tocilizumab 4 mg/kg but received tocilizumab 8 mg/kg throughout the study. †Includes one patient who did not receive study treatment and was subsequently withdrawn. ‡Patients who did not achieve >20% improvement from baseline in both tender and swollen joint count at week 16 were offered rescue therapy.

RADIATE Emery et al 2008

Figure 1 Numbers of patients undergoing enrolment, random selection and study completion. *One randomly assigned patient was withdrawn from the study before receiving any study medication because of a latex allergy.



SATORI Nishimoto et al 2009

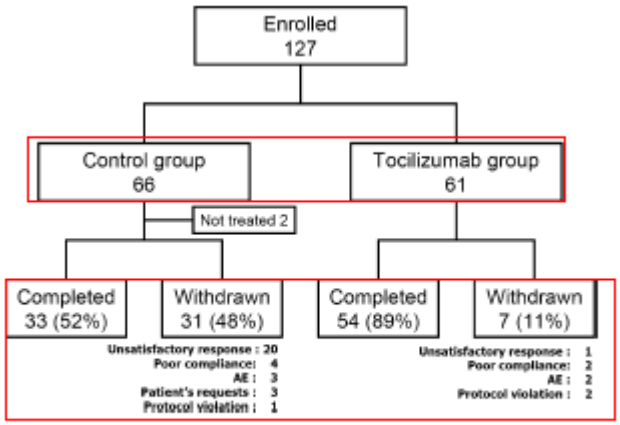


Fig. 1 Randomization, reasons for withdrawal, and numbers of patients who completed the trial. *Tocilizumab* humanized anti-interleukin-6 receptor antibody

TOWARD Genovese et al 2008

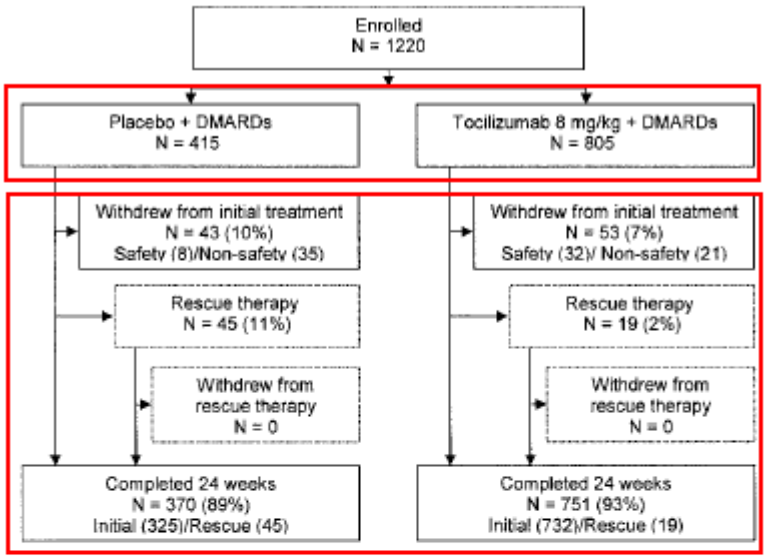


Figure 1. Patient disposition. The 415 patients in the disease-modifying antirheumatic drugs (DMARDs) plus placebo group included 2 patients who were randomized to this group but did not receive any study medication and 2 patients who were randomized to this group but received tocilizumab 8 mg/kg plus DMARDs throughout the study. The 805 patients in the tocilizumab plus DMARDs group included 2 patients who were randomized to this group but did not receive any study medication and 3 patients who were randomized to this group but received placebo plus DMARDs throughout the study. Patients who did not achieve 20% improvement from baseline in both the swollen joint count and the tender joint count at week 16 were offered rescue therapy (adjustment of the background DMARD dosage and/or treatment with a different conventional DMARD). Patients receiving rescue therapy were considered nonresponders in the primary efficacy analysis at 24 weeks.

[redacted]

[redacted]

8.15 *Appendix 16. Estimating the Weibull Distribution*

It was not possible to fit the Weibull distribution using patient level data for this model. Therefore it was necessary to estimate the rate of withdrawal from etanercept, infliximab and methotrexate using data reported in published articles. Two studies with a five year follow-up reported withdrawal data for infliximab, etanercept and methotrexate (25;26). Both studies published Kaplan-Meir survival estimate diagrams showing the rate of withdrawal from treatment for the three drugs. Point estimates for number of patients remaining on treatment were extracted using Grafula 3 software.

Using these estimates a regression was run in excel to generate 2 parameters loglambda and gamma.

The Weibull distribution was estimated in the model using the following equation:

$$pwithdrawa\ \ l = EXP\left(-\lambda\left(t^{\gamma}\right)\right)$$

The lambda and gamma parameters were varied according to a normal distribution in the PSA. The measures of precision for lambda and gamma were taken from the regression Variance/Covariance matrix.

The regression outputs for infliximab, etanercept and methotrexate can be found below.

SUMMARY OUTPUT- Infliximab						
Regression Statistics						
Multiple R	0.965843					
R Square	0.932852					
Adjusted R Square	0.930166					
Standard Error	0.244788					
Observations	27					
ANOVA						
	df	*SS*	*MS*	*F*	*Significance F*	
Regression	1	20.8114	20.8114	347.3129	3.55E-16	
Residual	25	1.498029	0.059921			
Total	26	22.30943				
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-3.21521	0.128009	-25.1169	2.95E-19	-3.47885	-2.95156
X Variable 1	0.872721	0.046829	18.63633	3.55E-16	0.776275	0.969167

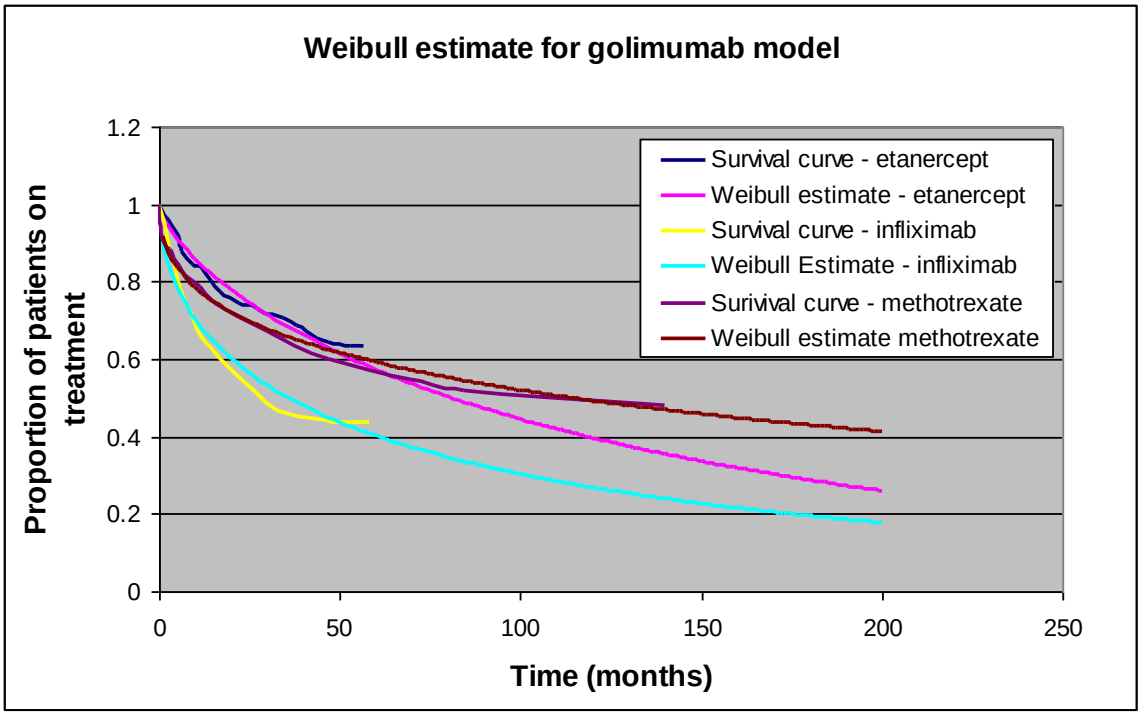
SUMMARY OUTPUT - etanercept						
Regression Statistics						
Multiple R	0.98515					
R Square	0.97052					
Adjusted R Square	0.968554					
Standard Error	0.116969					
Observations	17					
ANOVA						
	df	*SS*	*MS*	*F*	*Significance F*	
Regression	1	6.756275	6.756275	493.815	6.82E-13	
Residual	15	0.205227	0.013682			
Total	16	6.961502				
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-3.45111	0.097001	-35.5783	6.66E-16	-3.65786	-3.24436
X Variable 1	0.692736	0.031174	22.22195	6.82E-13	0.626291	0.759181

SUMMARY OUTPUT - methotrexate						
Regression Statistics						
Multiple R	0.991013					
R Square	0.982106					
Adjusted R Square	0.98048					
Standard Error	0.103417					
Observations	13					
ANOVA						
	df	*SS*	*MS*	*F*	*Significance F*	
Regression	1	6.45704	6.45704	603.7408	5.82E-11	
Residual	11	0.117646	0.010695			
Total	12	6.574686				

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-2.41641	0.051934	-46.528	5.53E-14	-2.53071	-2.3021
X Variable 1	0.433048	0.017624	24.57114	5.82E-11	0.394257	0.471839

A graph depicting the survival curves for each treatment and the estimated Weibull distribution can be found in Figure 11.

Figure 8: Weibull estimates for duration of treatment



8.16 Appendix 17. Source Code

[redacted]

[redacted]

[redacted]

[redacted]