NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Premeeting briefing

Tocilizumab for the treatment of systemic juvenile idiopathic arthritis

This briefing presents the key issues arising from the manufacturer's submission, Evidence Review Group (ERG) report and statements made by consultees and their nominated clinical specialists and patient experts. Please note that this briefing is a summary of the information available and should be read with the full supporting documents.

The manufacturer was asked to provide:

- Data clarifying the proportion of participants whose arthritis did not respond adequately to methotrexate in the TENDER trial
- Data clarifying the proportion of participants who were methotrexate naive in the TENDER trial
- Clarification on the health states as they relate to the CHAQ score and ACR responses in the economic model
- Clarification on the transition between the various ACR categories and how they had been represented in the Markov model structure
- Further details on cost effectiveness analysis and sensitivity analysis
- References for excluded citations

Proposed licensed indication

In May 2011 the Committee for Medicinal Products for Human Use recommended that tocilizumab (RoActemra, Roche) is indicated for the treatment of active systemic juvenile idiopathic arthritis (JIA) in patients aged 2 years and older, who have responded inadequately to previous therapy with non-steroidal anti-inflammatory drugs (NSAIDs) and systemic corticosteroids. Tocilizumab can be given as monotherapy (in case of intolerance to

methotrexate or where treatment with methotrexate is inappropriate) or in combination with methotrexate.

Tocilizumab is administered as an intravenous infusion over 1 hour and treatment is repeated at 2-week intervals. The recommended dose is 8 mg/kg in patients weighing 30 kg or more, and 12 mg/kg in patients weighing less than 30 kg.

Key issues for consideration

Clinical effectiveness

- Does the Committee consider that the population in the TENDER trial reflects the UK population with systemic JIA?
- Does the Committee agree that the 70% of patients in the TENDER trial who had been treated with methotrexate for at least 12 weeks before the trial should be considered to have systemic JIA with an inadequate response to methotrexate?
- Does the Committee agree that the 25% of patients in the TENDER trial who had been treated with methotrexate in the past should be considered to have systemic JIA with an inadequate response to methotrexate?
- Does the Committee consider that the TENDER trial adequately represents the populations described in the scope?
  - Population 1 is defined in the scope as children and young people 2 years and older with systemic JIA which has not responded adequately to prior NSAIDs and systemic corticosteroids. The comparator for this population in the scope is people treated with methotrexate. Does the Committee consider that the data for the 5% of patients who were methotrexate naive is the most appropriate evidence for population 1?
  - Population 2 is defined in the scope as children and young people 2 years and older with systemic JIA which has not responded adequately to prior NSAIDs, systemic corticosteroids and methotrexate. Does the Committee consider that the most appropriate evidence to use for

population 2 is the data for 95% of patients who had tried methotrexate, instead of data from all patients in the trial?

- Is the indirect comparison of tocilizumab with infliximab appropriate, given that the population in the infliximab trial included patients with pauciarticular and polyarticular JIA in addition to those with systemic JIA?
- Is the Committee satisfied with the indirect comparison of tocilizumab with anakinra? Does the Committee accept the ERG's results, using the updated data (for 95% of patients who had tried methotrexate) supplied by the manufacturer in response to clarification, for this comparison?

Cost effectiveness

- Does the Committee accept the manufacturer's economic model structure, given that the health states do not appear to be mutually exclusive?
- In the economic model the manufacturer assumes that transitions between the various American College of Rheumatology (ACR) states is not possible. Does the Committee accept this assumption?
- Are the assumptions used by the manufacturer to assign utility values for each health state appropriate? For example, does a difference in ACR response represent a corresponding difference in childhood health assessment questionnaire (CHAQ) score?
- Are the assumptions used by the manufacturer to map CHAQ scores to utilities appropriate? It is assumed that a CHAQ score of a child is equivalent to the health assessment questionnaire (HAQ) score of an adult and that adult EQ-5D is equivalent to the health related quality of life of a child?
- Does the Committee consider that the CHAQ scores for ACR response found in the TENDER trial are also valid for the anakinra and TNF-alpha inhibitors comparison?
- Does the Committee consider the costs for the heath states appropriate?
- Does the Committee consider the starting age of 2 years in the manufacturer's economic model to be appropriate?

- Does the Committee accept the ICER for tocilizumab compared with methotrexate, given that no data has been presented for population 1 in the TENDER trial?
- Does the Committee accept the ICERs for tocilizumab compared with the comparators given the concerns raised about the model structure?
- In the base case analysis for tocilizumab compared with anakinra the manufacturer used ACR30 response at week 12 as an input into the model. However, the primary outcome for tocilizumab at week 12 is ACR30 and absence of fever. The indirect comparison of tocilizumab with anakinra showed ACR30 and absence of fever not achieving statistical significance. 1.91 (0.84-4.37). Is the Committee satisfied with this base-case analysis?
- Does the Committee consider that when using a TNF alpha inhibitor as second or third line the effectiveness of the subsequent TNF alpha inhibitor in a sequence will be potentially different from if it were used first line?

1 Decision problem

1.1 Decision problem approach in the manufacturer's submission

Population	1. Children and young people 2 years and older with systemic JIA which has not responded adequately to prior NSAIDs and systemic corticosteroids 2. Children and young people 2 years and older with systemic JIA which has not responded adequately to prior NSAIDs, systemic corticosteroids and methotrexate.
Intervention	Tocilizumab with or without methotrexate
Comparators	For population 1 • methotrexate For population 2 • TNF-alpha inhibitors (for example etanercept and infliximab) • Anakinra
Outcomes	The outcome measures addressed include: • disease activity • physical function • joint damage (damage assessed by radiographic progression is not available from the 12 week data from the TENDER trial) • pain • steroid sparing • mortality • adverse effects of treatment • health-related quality of life The manufacturer also included 'fever' as an outcome
Economic evaluation	• Cost effectiveness of treatments is expressed in terms of incremental cost per quality-adjusted life year (QALY) • The time horizon considered is the lifetime of the patient • Costs are considered from an NHS and personal social services perspective

1.2 Evidence Review Group comments

1.2.1 Population

The ERG noted that the population addressed in the decision problem by the manufacturer matched that in the scope. However, the TENDER trial was not designed to address the question in the two distinct populations. The inclusion

criteria in the TENDER trial suggest that the population matched those whose disease responded inadequately to NSAIDs and systemic corticosteroids (that is, population 1). The ERG noted that the manufacturer was of the view that the population in the TENDER trial matched population 2, because 70% of patients (all with disease that responded inadequately to NSAIDs and systemic corticosteroids at baseline) were still receiving methotrexate and therefore could be considered to have disease with an inadequate response to methotrexate.

1.2.2 Intervention

The ERG noted that the intervention in the decision problem correctly matched the scope.

1.2.3 Comparators

The ERG noted that the TENDER trial compares tocilizumab plus standard care with placebo plus standard care. The ERG observed that the comparator in this study did not match the scope and decision problem. For population 1 the comparator in the scope is methotrexate. The manufacturer had used a post-hoc analysis to compare patients receiving tocilizumab with those patients in the placebo group also receiving methotrexate. The ERG noted that this was not methodologically acceptable because the trial participants were not originally randomised into those populations.

The ERG also noted that for population 2 the comparators should be TNF-alpha inhibitors (such as etanercept, adalimumab, and infliximab) and anakinra. The manufacturer did not identify any head-to-head evidence comparing tocilizumab with any of the comparators, so performed the analyses using indirect comparisons. For the tocilizumab arm, the manufacturer used the TENDER trial. The ERG noted that the indirect comparison used data for all patients, instead of data from patients who would represent those in the decision problem, (that is, the 95% of patients who had received methotrexate). For the comparators, the ERG noted that the manufacturer had decided to broaden the inclusion criteria to include all subtypes of juvenile arthritis, not just systemic JIA. The manufacturer had

taken this approach because of the dearth of clinical evidence in systemic JIA. The ERG was concerned that this approach had been taken despite the manufacturer's clinical experts stressing the differences between systemic JIA and other subtypes and advising against comparing the evidence from different JIA populations.

1.2.4 Outcomes

The ERG noted that the scope outcomes of disease activity, physical function, pain, adverse events and steroid sparing had been matched in the manufacturer's submissions. However, the ERG was not satisfied that the manufacturer had adequately addressed joint damage and health-related quality of life. The ERG also noted that, because of the clinical characteristics of systemic JIA, it would be important to consider the outcomes of lymph node enlargement, hepatomegaly, splenomegaly and serositis. The ERG also thought it would be appropriate to present macrophage inactivation syndrome more clearly within the adverse events.

1.2.5 Economic evaluation

The ERG considers the manufacturer's incremental cost-effectiveness ratio (ICER) for tocilizumab compared with infliximab is biased as a result of lack of data. This is largely related to the problems identified with the indirect comparison; basing the indirect comparison with TNF inhibitors on one study in a general JIA population leads to biases, even though the manufacturer attempted to correct for this using an adjustment factor.

1.3 Statements from professional/patient groups and nominated experts

Patient and professional groups state that systemic JIA is the least common form of juvenile arthritis, affecting fewer than 1 in 10 children with arthritis. Systemic JIA usually begins before the age of 5 years. It can manifest with systemic features, including temperature, rash and general malaise. Arthritis can be present but is often not an initial feature. Approximately 50% of people with systemic JIA still have active disease 15 years after onset. Patient and

professional groups state that people with systemic JIA will have considerable disability.

Patient groups highlight that etanercept is licensed for polyarticular JIA but not for systemic JIA. There are currently no treatments licensed for systemic JIA, but an expert stated that there was a broad agreement for the initial stages of treatment: namely high dose NSAIDs and steroids. Except in very mild disease, the steroid used initially is usually intravenous methylprednisolone, then oral prednisolone with boosts of intravenous methylprednisolone as needed. If disease activity persists, or if it is severe initially, then methotrexate is used.

If the person is intolerant to this regime or is not adequately treated by it, there is less clear agreement of the next steps. Subsequent options often have toxic side effects or require regular treatment. They include: anti-TNF-alpha therapy; steroid joint injections; high-dose intravenous immunoglobulin; oral ciclosporin; anakinra; tocilizumab; oral thalidomide; autologous stem cell rescue after marrow ablation; and cyclophosphamide. The experts considered that methotrexate was useful in some children with systemic JIA.

The experts noted that use of tocilizumab would allow a decrease in steroid use. The patient and professional groups highlighted that the advantages of better control of systemic JIA on lower dose steroids are wide ranging and produce lifelong benefits including: fewer steroid side effects (especially growth restriction, vulnerability to infection and osteoporosis); reduced risk of long-term joint damage and need for joint replacement; less pain and better energy, with benefits to school attendance, education, relationships with peers and social development; reduced risk of developing amyloidosis; reduced risk of cardiovascular disease associated with persistent inflammation; and reduced disruption to family, including sibling development and parent or carer employment.

The experts agreed that there was variation in the use of tocilizumab in the UK. The experts noted that tocilizumab would be a viable option in patients for whom current recommended treatment fails. There was some concern that

the long-term effects of tocilizumab in children whose treatment extends into adolescence or adulthood were not known. One of the experts noted that treatment of systemic JIA typically followed one of three patterns:

- in approximately 11% of patients, a one-off course with a good response and withdrawal of NSAIDs and systemic corticosteroids over a few months with no return of activity
- in about 34% of patients, a repeated course with intermittent relapse and remission
- in approximately 55% of patients, an unremitting course of treatment with difficulty in achieving remission.

The expert's view was that tocilizumab would be useful in the management of the second and third groups and that tocilizumab use should be reserved until remission develops, unless there is severe activity or there is only a partial response to steroids.

Tocilizumab requires hospital day-case attendance for infusion and experts agreed that it needs to be given by an appropriately qualified, specialist, expert multidisciplinary team, on a paediatric ward with appropriate guidelines.

Patients and professionals highlighted that tocilizumab being administered fortnightly, rather than weekly or biweekly as is the case for alternative treatments, is an attractive option for patients and clinicians.

Experts noted that everyone receiving tocilizumab should be registered with the British Society for Paediatric and Adolescent Rheumatology (BSPAR) Biologics Registry.

2 Clinical effectiveness evidence

2.1 Clinical effectiveness in the manufacturer's submission

2.1.1 Tocilizumab versus methotrexate

A systematic review was carried out, and identified one study as the most relevant to the decision problem. The TENDER trial is an ongoing three part, 5 year, phase III randomised controlled trial.

Part one consisted of a 12-week international multicentre randomised double-blind placebo-controlled parallel two-group study to evaluate the efficacy and safety of tocilizumab in children with active systemic JIA. The study enrolled 112 participants (from 17 countries, including the UK) unequally randomised 2:1 to tocilizumab (n = 75) or placebo (n = 37). Tocilizumab was administered every 2 weeks with a dose of 8 mg/kg for participants who weighed at least 30 kg (n = 37) and 12 mg/kg for those who weighed less than 30 kg (n = 38).

Part II is a 92-week single-group open-label extension and part III is a 3-year single-group open-label continuation of the study.

Ages of patients in the trial ranged from 2 to 17 years, with an average age of 10. Patients had to have documented persistent disease activity (at least five active joints, or at least two active joints with fever of above 38°C for any 5 out of 14 days screening) for at least 6 months, with inadequate response to NSAIDs and corticosteroids because of toxicity or lack of efficacy. Inadequate response to previous treatment was determined by the treating physician's clinical assessment. Before study entry 78/112 patients (70%) had been treated with methotrexate (36 entered the study on methotrexate that had been previously stopped then restarted; 42 were on their first course of methotrexate, which was ongoing). Twenty-nine patients (approximately 26%) had no background methotrexate at baseline but had received and stopped methotrexate previously. Five (approximately 4%) patients had never received methotrexate, and could be considered methotrexate naive. Patients taking

NSAIDs, corticosteroids and methotrexate were permitted but had to enter the study on a stable dose of the medicines. There was an 'early escape' option to allow children with more severe disease at baseline an opportunity to escape and receive active open-label tocilizumab.

The primary outcome measures were the proportion of patients achieving a JIA ACR30 response at 12 weeks and absence of fever (defined as no recorded temperature of 37.5°C or above in the preceding 7 days). JIA ACR30 response is defined as any three of six core outcome variables improved by at least 30% from the baseline assessments, with no more than one of the remaining variables worsened by more than 30%. Patients who withdrew, received escape medication, or for whom the endpoint could not be determined were classified as non-responders. The JIA core outcome variables are: physician global assessment of disease activity (100 mm visual analogue scale [VAS]); parent/patient global assessment of overall well-being (100 mm VAS); number of joints with active arthritis; number of joints with limitation of movement; erythrocyte sedimentation rate; and functional ability (using the Childhood Health Assessment Questionnaire, which measures eight everyday functional activities).

The secondary outcomes are: individual results for each JIA ACR component at 12 weeks: JIA ACR 50/70/90 responses at 12 weeks (that is, an improvement by at least 50%, 70% or 90% from the baseline assessments in any three of the six core outcome variables, and no more than one of the remaining variables worsening by more than 50%, 70% or 90%); corticosteroid reduction; fever; rash; pain; and laboratory outcomes (C-reactive protein [CRP]) levels, anaemia and haemoglobin levels, thrombocytosis and leucocytosis). For further details of the TENDER design, see pages 47–74 of the manufacturer's submission.

Results of the TENDER trial

Efficacy endpoints were analysed using the intention-to-treat population. All patients were classified as either responders or non-responders, and those patients who withdrew or escaped were classed as non-responders. For details of patients who escaped and withdrew from therapy see page 72 of the manufacturer's submission. The results of the TENDER trial are presented on pages 80–194 of the manufacturer's submission.

The TENDER study met its primary endpoint of a JIA ACR30 response and absence of fever at week 12 with 85.3% of the tocilizumab patients responding in contrast to 24.3% of the placebo patients, a statistically significant difference (p<0.0001).

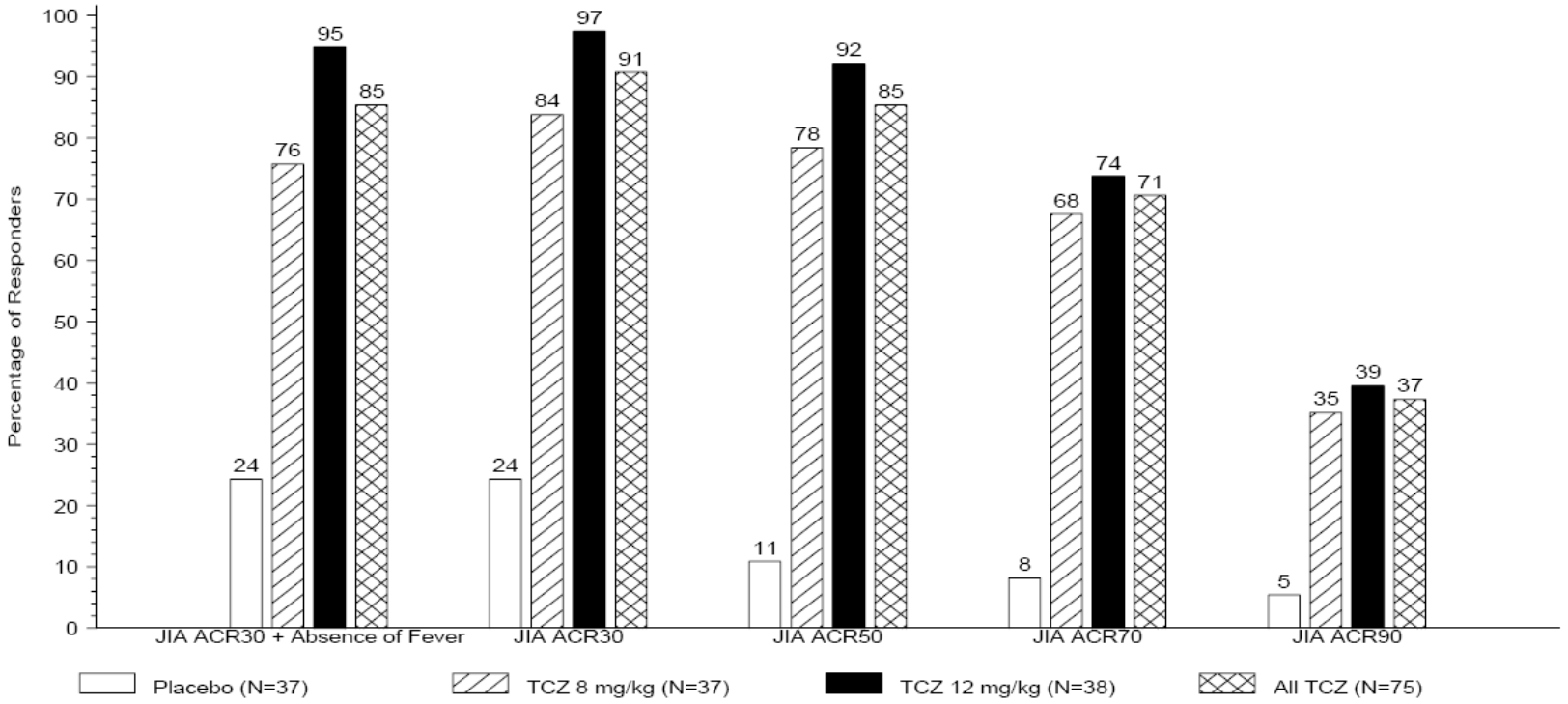
Table 1 Summary of key TENDER efficacy data (taken from page 81 of manufacturer's submission):

Summary and analysis of the percentage of patients with a JIA ACR30 response and absence of fever at week 12 – all tocilizumab compared with placebo (ITT population)

	Placebo (n = 37)	**Tocilizumab, all patients (n = 75)**
Number of responders (%) (95% confidence intervals)	9 (24.3%) (10.5 to 38.1)	64 (85.3%) (77.3 to 93.3)
Weighted difference versus placebo (95% confidence intervals) P value		61.5 (44.9; 78.1) < 0.0001

Tocilizumab patients had a greater chance of achieving JIA ACR30/50/70/90 responses at week 12 in comparison with the placebo patients. The differences in proportions of each JIA ACR response level were statistically significantly different ($p<0.0001$). The proportion of responders was higher in the tocilizumab 12 mg/kg patients compared with the tocilizumab 8 mg/kg patients. Figure 1 below illustrates the above. For further details of this analysis see page 87 of the manufacturer's submission.

Figure 1: Bar Chart of the Proportion of JIA ACR30 Responders with Absence of Fever and JIA ACR30/50/70/90 Responders at Week 12 (ITT Population)



Responders are patients who had a JIA ACR30 response and absence of fever or JIA ACR30/50/70/90 responses at Week 12.
Absence of fever (temperatures <37.5C) in the 7 days preceding the Week 12 assessment day.
Patients who withdrew, received escape medication, or for whom the endpoint could not be determined are classified as non-responders.
LOCF rule applied to missing JIA ACR core set components at Week 12.

The efficacy of tocilizumab with respect to a number of ACR core set components was analysed as part of the secondary efficacy analyses. A summary table of the results is presented below. The table is taken from page 89 of the manufacturer's submission.

Table 2: Analysis of variance of percentage change from baseline in the JIA ACR core set components at week 12 - all tocilizumab compared with placebo (ITT Population).

[redacted]

Adverse events

[redacted]

[redacted]

2.1.2 Tocilizumab versus TNF-alpha inhibitors or anakinra

No head-to-head trials were available analysing the efficacy of tocilizumab compared with etanercept, infliximab and anakinra for population 2. Therefore, the manufacturer searched for trials of the comparator interventions. It identified five studies, all of which compared one drug with placebo. Only one study evaluated a population of solely systemic JIA patients. Ruperto et al. 2008 studied abatacept, Lovell et al. 2008 studied adalimumab, Lovell et al. 2000 studied etanercept, Ruperto et al. 2007 studied infliximab and Quartier et al. 2010 studied anakinra.

Only data from Ruperto et al. 2007 and Quartier et al. 2010 were included in the indirect comparison analysis because the manufacturer stated that the design of these trials most closely matched the TENDER trial. In the Ruperto et al participants were blindly randomised to receive either placebo or infliximab before the open label phase. Quartier et al also had a randomised controlled phase before the open label phase.

Ruperto et al. 2007 compared infliximab with placebo in patients with juvenile rheumatoid arthritis (systemic 16%, pauciarticular 23%, polyarticular 61%) that was described as having suboptimal response to methotrexate. Participants were from North and South America and Europe, aged between 4 and 18 years, and 62 were randomised to infliximab and 60 to placebo. Patients received concomitant methotrexate alongside placebo or active treatment. The study was a randomised double blind placebo

controlled trial, and the primary outcome was the proportion of patients meeting a paediatric ACR 30 (ACR Pedi 30 score) response based on JIA core set parameters at week 14. Quartier et al. (2010) focused on people with systemic JIA and compared anakinra with placebo. This was a multicentre study with 24 participants (12 in each arm) aged 2 to 20 years, from North America and Europe. The study included patients whose systemic JIA had not responded to methotrexate and any of the disease modifying anti rheumatic drugs (DMARDs) and did not permit the administration of any DMARDs for the duration of the trial. The outcomes of the randomised controlled phase were reported after 1 month. The primary outcome was the ACR Pedi score, absence of fever and normalisation of CRP and erythrocyte sedimentation rate after 1 month.

Table 3 Evidence used in the indirect comparison analysis from page 141 of manufacturer's submission

Study	Treatment	Response	Total n	ACR only		ACR and absence of fever (< 38°C)	
				%	n	%	n
TENDER	TCZ +/-methotrexate	ACR 30	75	0.907	68	0.853	64
	TCZ +/-methotrexate	ACR 50	75	0.853	64	–	–
	TCZ +/-methotrexate	ACR 70	75	0.707	53	–	–
	PBO +/-methotrexate	ACR 30	37	0.243	9	0.243	9
	PBO +/-methotrexate	ACR 50	37	0.108	4	–	–
	PBO +/-methotrexate	ACR 70	37	0.081	3	–	–
ANAJIS (Quartier et al. 2010)	ANK	ACR 30	12	0.92	11	0.92	11
	PBO	ACR 30	12	0.58	7	0.5	6
NCT00036374 (Ruperto et al. 2007)	INFL +methotrexate	ACR 30	58	0.638	37	–	–
	PBO +methotrexate	ACR 50	58	0.5	29	–	–
	INFL +methotrexate	ACR 70	58	0.224	13	–	–
	PBO +methotrexate	ACR 30	59	0.492	29	–	–
	PBO +methotrexate	ACR 50	59	0.339	20	–	–
	PBO +methotrexate	ACR 70	59	0.119	7	–	–
Abbreviations: ACR=American College of Rheumatology; ANK=Anakinra; INFL=Infliximab PBO=Placebo							

Table 4 Results of the indirect comparison analysis from page 143 of manufacturer's submission

Comparison	Outcome	RR	95% CI
TCZ vs ANK	ACR30	2.37	1.10 to 5.10
	ACR30 and absence of fever	1.91	0.84 to 4.37
TCZ vs INF	ACR30	2.87	1.49 to 5.55
	ACR50	5.35	1.91 to 14.97
	ACR70	4.61	1.16 to 18.38
ACR:=American College of Rheumatology; ANK=Anakinra; INF=Infliximab			

The primary endpoint of the TENDER trial was ACR30 and absence of fever.

2.2 Evidence Review Group comments

The ERG noted that clearly documented search methods were presented in the submission and clarification response. Absence of comprehensive synonyms and poorly applied study design limits were identified in several of the search strategies, which may have affected the results of the search. For the most part, the ERG was unable to determine whether any relevant studies were not identified.

The manufacturer identified two studies: TENDER and Yokota 2008. Yokota was subsequently excluded from further analysis by the manufacturer. However the ERG thought that the Yokota trial should have remained in the analysis. The results for Yokota are presented in table 4.5, page 38 of the ERG report.

The ERG stated that the evidence presented in the manufacturer's submission is not in accordance with the NICE scope. The ERG considered that there is no evidence for any comparison in the NICE scope.

Table 5 12-week outcomes for the TENDER trial summarised by the ERG for each outcome and population in the scope (effect sizes calculated by the ERG)

Outcome/population	Population 1	Population 2		
		Tocilizumab[a] (Responders/patients analysed)	**Placebo (Responders/patients analysed)**	**Effect size Relative risk (95% confidence interval)**
Primary endpoint:				
JIA ACR 30 [b]	–	68/75	9/37	3.73 (2.1 to 6.61) *
JIA ACR 50		64/75	4/37	7.89 (3.11 to 20.11) *
JIA ACR 70		53/75	3/37	8.72 (2.92 to 26.0) *
JIA ACR 90		28/75	2/37	6.91 (1.74 to 27.4) *
Steroid sparing [c]	–	17/70	1/31	17.57 (2.49 to 123.89) *
Mortality	–	0/75	0/37	
Adverse events Patients with:				
• ≥1 adverse event	–	66/75	23/37	1.42 (1.09 to 1.85)*
• ≥1 serious adverse event	–	3/75	0/37	
• ≥1 infection	–	41/75	11/37	1.84 (1.08 to 3.14)*
• ≥1 serious infection	–	2/75	0/37	

		Tocilizumab [a] **Adjusted mean change from baseline (patients analysed)**	**Placebo Adjusted mean change from baseline (patients analysed)**	**Effect size Adjusted mean difference (95% confidence interval)**
Disease activity: Physician Global Assessment of Disease Activity (100 mm VAS)	–	−69.6 (73)	−41.1 (17)	−64.4 (−87.5 to −41.3)*
Number of active joints	–	−70.6 (73)	−37.2 (17)	−-33.4 (−53.2 to −13.6)*
Physical function No. of joints with limitation of movement	–	−51.6 (72)	−22.5 (17)	−29.1 (−53.4 to −4.9)*
CHAQ-DI score	–	−45.6 (72)	−10.3 (17)	−35.3 (−63.5 to −7.1)*
Joint damage	–	–		
Pain, visual analogue scale (0–100 mm)	–	−41.0 (73)	−1.1 (17)	−39.8 (−55.1 to −24.6)*
Health-related quality of life	–	–		
Definitions: Population 1, children with systemic JIA that has not responded adequately to treatment with NSAID and systemic corticosteroids; Population 2, children with systemic JIA that has not responded adequately to treatment with NSAID, systemic corticosteroids and methotrexate; Adjusted mean difference, analysis of variance adjusted for the randomisation stratification factors applied at baseline.				

* results are statistically significant

[a] Tocilizumab is a combination of 8 mg/kg and 12 mg/kg doses

[b] analysis was reported as intention to treat, but patients who withdrew, received escape medication, or for whom the endpoint could not be determined were classified as non-responders

[c] Patients receiving oral corticosteroids with JIA ACR70 response at week 6/8 who reduced oral corticosteroids dose by ≥ 20% without subsequent JIA ACR30 flare or occurrence of systemic symptoms

According to the manufacturer 95% of TENDER trial participants match population 2, because 'patients are included in the study if they have symptoms of active disease' and 'It follows that if patients have tried in the past or are currently administered methotrexate and continue to have persistent disease then they are inadequate responders' ('Response to clarification letter', question A2). Therefore the manufacturer states that there is some evidence for the second population but none for the first population, and no data were provided in the manufacturer's submission for population 1. The only comparison given is tocilizumab versus anti-TNFs or anakinra. The manufacturer performed a systematic review to identify trials for the comparators. One trial was identified in children with systemic JIA, comparing anakinra with placebo. The manufacturer decided to broaden the inclusion criteria to include all trials in juvenile arthritis regardless of subtype, despite advice from their clinical experts to the contrary (see manufacturer's submission, page 116). The ERG agrees with the advice from the clinical experts; therefore its report does not comment on trials in children with other types of juvenile arthritis.

For population 2 (children with systemic JIA with an inadequate response to NSAIDs, corticosteroids and methotrexate) the manufacturer's submission provided data for an indirect comparison of tocilizumab with anakinra, using data from the TENDER trial, and a trial of anakinra versus placebo. The ERG believes the 5% of participants in the TENDER trial who were methotrexate naive should be excluded from these analyses. The manufacturer's submission provided data only for all participants in the TENDER trial. However, in response to the clarification letter some data were provided in which methotrexate naive patients were excluded. These data were not reported for the TENDER trial, but only for the indirect comparison with anakinra. Where possible, the ERG used data for this population.

Table 6 Results of the indirect comparison analysis using data that excludes the 5% of methotrexate naive patients (supplied by manufacturer in response to clarification) page 45 of ERG report

Comparison	Outcome	Manufacturer's base-case analysis (TENDER[a])		ERG's sensitivity analysis (excl. methotrexate naive)	
		RR	95% CI	RR	95% CI
TCZ vs. ANK	ACR30	2.37	1.10 to 5.10	2.27	1.06 to 4.85
TCZ vs. INF	ACR30	2.87	1.49 to 5.55	2.75	1.44 to 5.26
	ACR50	5.35	1.91 to 14.97	5.04	1.81 to 14.04
	ACR70	4.61	1.16 to 18.38	4.33	1.09 to 17.20

Abbreviations: ACR=American College of Rheumatology; ANK=Anakinra; INF=Infliximab; TCZ=Tocilizumab

[a] Analysis was reported as intention-to-treat, but patients who withdrew, received escape medication, or for whom the endpoint could not be determined were classified as non-responders.

The ERG investigated heterogeneity within and across the TENDER and ANAJIS trials. The ERG reported that the inclusion criteria for both trials were similar (see page 46 of ERG report). Page 47 of the ERG report notes the differences between the two trials, the most important being the length of follow-up: 12 weeks in the TENDER trial and 1 month in the ANAJIS trial.

3 Cost effectiveness

3.1 Cost effectiveness in the manufacturer's submission

The manufacturer identified six published economic evaluations. Three were cost-utility studies, one was a willingness-to-pay study and two were stated as not clear. None of these were used by the manufacturer (see page 207 of manufacturer's submission).

The manufacturer submitted a Markov model to evaluate the costs and effectiveness of tocilizumab as part of a sequence of treatments (page 215 of the manufacturer's submission) in two different patient populations, using methotrexate and anakinra as comparators.

3.1.1 Model structure

Figure 2 shows the basic structure of the Markov model.

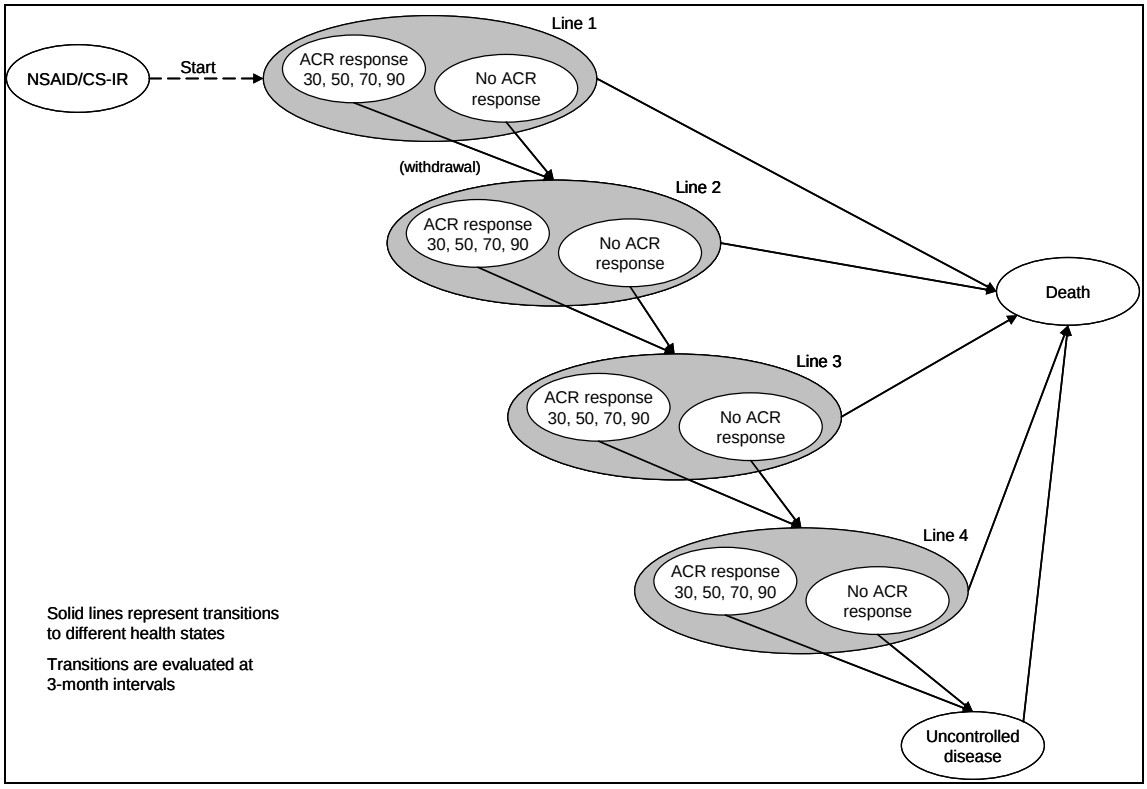


Figure 2 Structure of the Markov model (page 213 of the manufacturer's submission)

The Markov chain has 25 states. The model clusters the states into five groups: four are different lines of treatment and the fifth contains death and uncontrolled disease. Each line of treatment consists of five health states. Those five states are ACR response at the 30, 50, 70, or 90 level and 'no ACR response'. A patient can move from a particular ACR response from a particular line only to 'no ACR response' in the next line or to death. From 'no ACR response' the patient can move only to one ACR response level within this line of treatment or to 'no ACR response' in the next line. The patient cannot move within a given line to a better or worse health state (say, from ACR 50 to ACR 70). Only after being through all four lines does a patient move to the health state 'uncontrolled disease'.

The manufacturer states in the response to clarification letter that there is no evidence about transitions between ACR states within one treatment line, except that for tocilizumab the proportion of ACR70 and ACR90 responders increases following the first 12 weeks.

The probability of a response/non-response within a line of treatment depends on the treatment. The order in which the treatment is applied does not change these transitions. The probability of death is treatment- and health state-independent. The probability of withdrawal is health state independent, but higher for methotrexate than for other treatment options (all other treatment options have the same probability as each other). All transitions stay constant over time; that is, they are independent of age or disease duration. For further details of the transition and withdrawal probabilities for each comparison see pages 222–34 of the manufacturer's submission.

In each cycle, the proportion of patients in a given state is calculated. The distribution across states is used to calculate cycle-specific quality-adjusted life years (QALYs) and treatment costs, which are discounted and summed over the length of treatment.

Each health state leads to an absolute change in the initial CHAQ score. The initial CHAQ score is assumed to be the same for all patients, independent of treatment. For a given CHAQ score a utility is assigned to calculate QALYs. The health-state costs depend on the health state for and the treatment costs depend on the line of treatment.

The cohort characteristics at the start of the model are summarised in table 7.

Table 7 Cohort starting characteristics

Parameter	Value	Reference
Starting age	2 years	Assumption based on scope
Starting CHAQ score	1.702	Average CHAQ score at baseline from TENDER
Starting weight	13.25 kg	Assumption based on data extrapolated by a NICE rapid review on etanercept for JIA 2002

The main assumption of the model is that there were no transitions between ACR response categories. The analysis assumes that patients stay in the same health state unless they change treatment line. After 12 weeks of treatment the cohort is put on the next treatment in the sequence. In the

tocilizumab versus methotrexate model patients progressed to anakinra, etanercept and then adalimumab; in the tocilizumab versus anakinra model they progressed to etanercept, adalimumab and then abatacept.

The manufacturer's model has a time horizon of 16 years. That means that a patient in the model starting treatment at age 2 turns 18 and can be considered an adult at the end of the simulation. The model allows shorter and longer time durations for sensitivity analysis (up to 30 years). The discount rates applied were 3.5% for utilities and costs, and costs are considered from an NHS and personal social services perspective. A half-cycle correction was applied.

3.1.2 Utilities

The manufacturer stated that of the 34 studies identified in the review reporting on quality of life, none reported any utility values that were used in the economic evaluation presented here.

The TENDER trial included the Child Health Questionnaire (CHQ) as an instrument eliciting patient HRQL. The CHQ assesses a child's physical, emotional, and social well-being from the perspective of a parent or guardian (CHQ-PF50). The questionnaire was completed twice during the randomised period of the study: at baseline (visit 1) and at week 12 (visit 7).

During the design of the economic evaluation, a number of methods were attempted to translate CHQ scores to QALYs for the model. However, because of the lack of robust data and many other limitations (see pages 250–251 of the manufacturer's submission), an alternative method was selected to provide QALYs. A new mapping approach was developed, using CHAQ.

A CHAQ score is assigned to each of the four health states. The CHAQ score is mapped into utilities, using a mapping formula derived in adults with rheumatoid arthritis, mapping Health Assessment Questionnaire [HAQ] onto EQ-5D utilities. The manufacturer acknowledged that the assumptions that CHAQ is equal to HAQ and that adult EQ-5D is equal to health-related quality

of life of a child is not evidence based, and it was only because of the lack of other available data that this mapping method was preferred for the analysis to derive QALYs for the economic model. For details of the mapping formula see pages 251–254 of the manufacturer's submission.

The manufacturer states that health states reflect the condition of patients dependent on ACR response after a 12 week period on treatment. The utility of the health state is characterised by the resulting CHAQ triggered by the ACR response. The utilities for each health state presented below are based on the average CHAQ score per ACR response state, combined with the mapping formula.

Table 8 Summary of quality-of-life values for cost-effectiveness analysis

Health state name	CHAQ	Assumed QoL	Assumed SE	Adult RA values (for reference)
No response or uncontrolled disease	1.7442	0.4152	30% of the mean	0.4651
ACR 30	1.2699	0.5674		0.5660[a]
ACR 50	1.1351	0.6050		0.6084
ACR 70	0.8601	0.6736		0.6289
ACR 90	0.6692	0.7150		N/A
Abbreviations: ACR=American College of Rheumatology [a] Refers to ACR 20 and not ACR 30 in adult RA				

3.1.3 Costs

Treatment costs are a composite of the cost of the medication and the cost for administering the medication. In some cases the required dosage depends on the body weight of the patient. The manufacturer based its unit costs on UK reference costs, literature and expert opinion see section 6.5.5 and 6.5.6 of the manufacturer's submission. See table 5.18, page 69 of the ERG report for data on acquisition costs, cost of administration and dosage.

Table 9 Cost per year for a treatment, assuming wastage from varying package size (mean values) from ERG report page 69

Age	Abatacept	Adalimumab	Anakinra	Etanercept	Infliximab	Tocilizumab
2	£5,124	£9,383	£10,494	£9,566	£3,701	£10,570
5	£5,124	£9,383	£10,494	£9,566	£3,701	£13,233
8	£8,273	£9,383	£10,494	£9,566	£3,701	£17,226
10	£8,273	£9,349	£10,021	£9,431	£3,701	£13,233
12	£8,273	£9,349	£10,021	£9,431	£6,428	£17,226
14	£1,421	£9,349	£19,569	£9,431	£6,428	£19,889
16	£11,421	£9,349	£19,569	£9,431	£6,428	£19,889
18	£11,421	£9,349	£19,569	£9,431	£6,428	£23,882

A resource use schedule for a JIA patient was identified and modified for the current economic analysis. To establish the differences between the resource use for each health state, items from the combined cost schedule were presented to clinical experts in structured interviews to determine the proportion of patients that make use of an item, and frequency of use. For further details see section 6.3.5 manufacturer's submission. Several items were excluded from identified resource use schedule, see pages 70–74 of the ERG report for details.

The health state costs depend only on the ACR response level and are independent from any other health outcomes.

Table 10 Health state costs

Health state	Costs per year
Uncontrolled disease	£3,360
Response ACR 30	£504
Response ACR 50	£449
Response ACR 70	£396
Response ACR 90	£345

The manufacturer states that in all comparisons the identified adverse events are of minor severity and short duration, and have a minuscule cost impact for their management. Therefore, it can be assumed that they do not have a considerable bearing on the incremental cost of the two model arms.

3.1.4 Results

The base case results are given in tables 11 and 12 (reproduced from tables 108 and 109 of the manufacturer's report, page 302).

Table 11 Base-case results: comparison of tocilizumab with methotrexate

Technologies	Total costs (£)	Total LYG in response	Total QALYs	Incremental costs (£)	Incremental LYG in response	Incremental QALYs	ICER (£) incremental (QALYs)
Strategy TCZ	£141,716.09	6.4341	5.4465	£15,197.38	2.6071	0.7304	£20,806.31
Strategy methotrexate	£126,518.71	3.8270	4.7161				
Abbreviations: TCZ= Tocilizumab							

Table 12 Base-case results: comparison of tocilizumab with anakinra

Technologies	Total costs (£)	Total LYG in response	Total QALYs	Incremental costs (£)	Incremental LYG in response	Incremental QALYs	ICER (£) incremental (QALYs)
Strategy TCZ	£138,927.21	6.1284	5.3223	£11,697.43	1.7797	0.5038	£23,219.02
Strategy ANK	£127,229.78	4.3486	4.8185				
Abbreviations: TCZ = Tocilizumab , ANK= Anakinra							

Base-case probabilistic sensitivity analysis results

The manufacturer presented results based on the means costs and QALYs. In comparison with methotrexate, the probability of tocilizumab being cost effective was 0.39 at £20,000 per QALY and 0.72 at £30,000 per QALY. When compared with anakinra, the probability of tocilizumab being cost effective was 0.38 at £20,000 and 0.63 at £30,000 per QALY.

Sensitivity analysis

The manufacturer undertook deterministic sensitivity analysis and scenario analysis to examine how varying various parameters and assumptions affected the robustness of the model. Tables 110 and 111 on pages 303 to 306 of the manufacturer's submission show a summary of the sensitivity analysis. The manufacturer reported that doubling infusion administration costs had a substantial effect on the ICER – it increased to £57,350 in the tocilizumab versus anakinra model. Another parameter that seemed to be a key driver in the cost effectiveness analysis was the cost of inpatient stay. Reducing the cost of inpatient stay by 50% increased the ICER to £37,491 in the methotrexate strategy and £39,765 in the anakinra strategy.

3.2 Evidence Review Group comments

The ERG noted that the current economic model does not adhere to conventions in Markov modelling. In a Markov cohort model the health states defined should comprise the full range of conditions that are relevant to a patient population, and the states should be mutually exclusive. In the manufacturer's submission the health states were defined to reflect a change in a patients' condition (change in CHAQ based on ACR response) instead of the absolute condition of the patient. Change in a patients' condition should be included in a Markov model as a health state transition rather than a health state. The consequence of using a change in a patients' condition as a health state is that the Markov states are heterogeneous rather than mutually exclusive, depending on the disease variation of the cohort at the start of the model.

The ERG also noted the assumption in the model that patients move to a certain ACR response and stay in that state until they either withdraw (moves to the next treatment line) or die. The ERG thought that, given the nature of the disease, this assumption was unlikely.

The ERG noted the lack of health-related quality of life data both in the TENDER trial and in the literature, and recognised that very strong assumptions such as assuming the CHAQ of the child is equal to the HAQ score of an adult and adult EQ-ED is equivalent to the HRQoL of a child were needed to assign a utility to each health state in the model. The ERG considered the approach by the manufacturer reasonable and acceptable.

The ERG explored whether the assumed standard error of 30% on the mean is reasonable for use in the PSA. The ERG found that combining the uncertainty in baseline CHAQ with the uncertainty around the parameter estimates of the mapping formula led to a standard error of less than 10% of the mean. The ERG considered the 30% uncertainty used in the model reasonable, because it also takes into account the additional uncertainty of the assumptions used for the mapping procedure.

The ERG noted the possibility that the manufacturer assumed that the CHAQ scores for ACR response found in the TENDER trial are also valid for the anakinra and tocilizumab comparison. The ERG found it difficult to judge whether this is a reasonable assumption. The baseline CHAQ score in the anakinra study is 1.55 (standard deviation 0.74) and in the infliximab study (used for the indirect comparison of tocilizumab) the baseline CHAQ score was 1.5 (standard deviation 0.7).

The CHAQ baseline score in the anakinra group is slightly lower than that observed in the TENDER trial; the baseline score in the infliximab study is much lower. The latter reflects the fact that the infliximab study was done in the whole JIA population instead of the systemic JIA population.

The ERG considered the sensitivity analyses performed on the utilities were limited. The CHAQ scores, which are directly mapped into utilities, are only

varied slightly. Only the starting CHAQ value (base case: 1.7 ± 0.8) is varied slightly (to 1.63, 1.73 and 2, respectively) to reflect the mean starting CHAQ when accounting for different subgroups. In the response to clarification letter the manufacturer states that, in effect, absolute CHAQ scores are modelled but because of the way the model is set-up this:

- leads to the assumption that all patients have initially the same CHAQ score and
- means all relative improvement leads to the same absolute (improved) CHAQ score.

A clarifying example can be given by the initial CHAQ distribution, which has a mean of 1.7 ± 0.8. Assuming normality, this translates into 16% of all patients having an initial CHAQ of less than 0.9 and 16% of all patients having a CHAQ score higher than 2.4. As a relative increase is modelled, the heterogeneity in the treatment health states is of a similar magnitude. The ERG is of the opinion that the manufacturer should have addressed all these heterogeneities.

The ERG also noted that the incremental change in CHAQ score/utilities between health states when having a treatment response, which is also affected by uncertainty, is neither part of the deterministic sensitivity analysis nor the probabilistic sensitivity analysis.

The ERG highlighted that the cost estimates for health states have been defined by expert opinion and do not seem reasonable, because they present a cost for non-responders (£3300) that is six times higher that the costs for an ACR30 response (£500), whereas a ACR90 response is associated with only a 30% decrease (£350) compared with ACR30. Additionally, because of the wide variation in health status of the patients at baseline, patients may be assigned different costs even though at 12 weeks they have the same absolute health status.

The ERG stated that because no reliable data has been presented to inform population 1, the ICER for tocilizumab compared with methotrexate is biased. The ERG does not go on to explore this comparison further.

The ERG noted that some evidence suggests that the peak age of onset of systemic JIA is between 18 months and 2 years. In a UK cohort the peak age was 2 years with a mean of 6 years. In another UK prospective study, the Childhood Arthritis Prospective Study, the median age of onset is reported as 6.4 years (interquartile range 4.2 to 9.8). The ERG also noted that the average age in the TENDER trial was close to 10 years. The manufacturer's scenario analyses shows that patient age has a substantial effect on the cost-effectiveness of the intervention; for example, a starting age of 5 leads to an ICER of approximately £36,000 and a starting age of 10 leads to an ICER of approximately £47,000 when comparing tocilizumab with anakinra. The increase in the ICER with age is a result of larger doses of tocilizumab being used as the child gets older, so the cost of tocilizumab increases, whereas the dose for comparator treatments does not increase to the same extent. Therefore when comparing tocilizumab with the other options at a later starting age, the years with a relatively low cost of tocilizumab are no longer included in the analyses, hence the increase in the ICER.

The ERG noted that there was statistical uncertainty (variability) around the ACR responses, which is based on the assumption of a sample size of 100. On request, the manufacturer provided PSA output based on corrected sample sizes, but unfortunately did not provide an updated version of the electronic model. It appears from the explanation given by the manufacturer that only the sample sizes were adjusted to reflect those observed in the clinical studies with tocilizumab, anakinra and infliximab. However, it also appears that no account has been taken of the fact that the transition probabilities for the TNF alpha inhibitors are a multiplication of the transition probability for tocilizumab, a relative risk and an adjustment factor, and that they are all associated with uncertainty.

3.2.1 ERG's exploratory analyses

Based on the original submission of the manufacturer, the findings of the ERG, and the response of the manufacturer to the clarification letter, the ERG ran the manufacturer's cost-effectiveness model using the following assumptions and modifications:

- The starting age is 7 years (based on the observed average of 6 years, combined with 1 year for diagnosis and [failed] treatments with NSAIDs, corticosteroids and methotrexate), with a time horizon of 11 years.
- The cycle length is adjusted to 12 weeks instead of the current 3 months.
- The withdrawal rate is based on the exponential distribution.
- The ACR response probabilities for tocilizumab is adjusted to reflect the methotrexate non-responder population (95% of the whole populations).
- The relative risk of anakinra is adjusted to reflect the non-methotrexate-naive population in the indirect comparison.
- Parameters for the distribution of treatment response for anakinra and other TNF alpha inhibitors for the PSA, to include additional uncertainty around the relative risks and around the adjustment factor.

The results of this ERG base case analysis are presented in tables 13 and 14.

Table 13 ERG Base-case results: tocilizumab compared with anakinra

Technologies	Total costs (£)	Total life years gained in response	Total QALYs	Incremental costs (£)	Incremental life years gained in response	Incremental QALYs	ICER (£/QALYs)
Strategy tocilizumab	£121,952	4.9668	4.3065	£16,318	1.3630	0.3835	£42,552
Strategy anakinra	£105,634	3.6038	3.9230				

Table 14 Summary of costs by strategy: ERG comparison of tocilizumab with anakinra

	Strategy tocilizumab	Strategy anakinra	Incremental
Treatment cost	£76,193	£42,183	£34,010
Health state cost	£45,760	£63,451	−£17,692
Total cost	£121,952	£105,634	£16,318

The substantially higher ICER is explained almost entirely by the higher starting age of 7 years, compared with 2 years used in the manufacturer's submission.

The ERG did some exploratory analysis based on their base case. The ERG varied the withdrawal probabilities so that high-responders would have a lower probability of withdrawing than low responders. This was implemented by assuming withdrawal of 5% for ACR30, 3.5% for ACR50, 2.7% for ACR70 and 1.5% for ACR90. There is no evidence base for the specific values used; the main aim was to use realistic values so that the base case withdrawal risk of 3.13% would be between the ACR50 and ACR70 response. The resulting ICER was £40,916 per QALY gained, slightly lower than the base case ICER.

In the second scenario, the ERG explored the effect of the assumption that after the initial response, patients stay in their current health state, withdraw and move to next line or die. The ERG assumed that patients would move between all health states with a probability of 10% per transition, that is, patients in the ACR30 state had (per cycle) a 10% chance of moving to ACR50, a 10% chance of moving to ACR70 and a 10% chance of moving to ACR90. The resulting ICER was £53,051 per QALY gained, 24% higher than the base case ICER. This indicates that the assumption that patients who do not move to the next treatment line stay in the same health state indefinitely is an optimistic scenario.

As an alternative to the ERG's starting age of 7, which was derived from literature, the ERG also explored the starting age of 9.7 that is observed in the TENDER trial data across all patients; this increases the ICER to £46,611 per QALY gained.

Finally, the ERG explored the costs and effects from various alternative sequences for treatment. The decision problem states that tocilizumab should be compared with anakinra and TNF inhibitors. The main focus of the manufacturer's submission is on anakinra as comparator, and etanercept as comparator is explored in a scenario analysis. However, only a pair-wise comparison is done, instead of the full incremental analysis of the three

treatment options. Additionally, the ERG considers that anakinra as a second-line treatment after tocilizumab also a viable option. In the table below are the ERG's results of this full incremental analysis. It shows that starting with etanercept followed by anakinra is dominated by (that is, more costly and less effective than) anakinra followed by etanercept. The strategy of tocilizumab followed by etanercept is extendedly dominated by tocilizumab followed by anakinra. Thus, the ERG stated that the ICER of interest becomes that of tocilizumab followed by anakinra compared with anakinra followed by etanercept. This is £39,026, slightly lower than the ERG's base case ICER.

Table 15 Cost-effectiveness results for various treatment sequences, excluding infliximab

Strategy	QALY	Costs	Incremental QALY	Incr. costs	ICER
Etanercept - anakinra - adalimumab - abatacept	3.9113	£105,819			
Anakinra - etanercept - adalimumab - abatacept	3.9230	£105,634	0.0118	-£185	dominates
Tocilizumab - etanercept - adalimumab - abatacept	4.3065	£121,952	0.3835	£16,318	£42,552
Tocilizumab - anakinra - etanercept - adalimumab	4.4082	£124,569	0.1017	£2,617	£25,730

The ERG explored sequences that also considered infliximab. Various sequences were explored and the table below shows the most relevant options (others are all dominated or extendedly dominated).

Table 16 Cost-effectiveness results for various treatment sequences, including infliximab

Strategy	QALY	Costs	Incremental QALY	Incremental costs	ICER
Infliximab - etanercept - adalimumab - abatacept	3.7545	£98,250			
Anakinra - infliximab - adalimumab - abatacept	3.9230	£101,332	0.1685	£3,082	£18,287
Tocilizumab - anakinra - infliximab - adalimumab	4.4082	£122,490	0.4852	£21,158	£43,607

4 Equalities issues

No equality and diversity issues relating to population groups protected by equality legislation were highlighted when the scope for this appraisal was developed or in any of the submissions.

5 Authors

Alfred Sackeyfio (Technical Lead) and Joanna Richardson (Technical Adviser), with input from the Lead Team (Peter Selby, Wasim Hanif and Judith Wardle).

Appendix A: Sources of evidence considered in the preparation of the premeeting briefing

A The Evidence Review Group report for this appraisal was prepared by Kleijnen Systematic Review Limited in collaboration with Erasmus University Rotterdam and Maastricht University

- Reimsma R, Al M, Lhachimi S et al. Tocilizumab for the treatment of systemic juvenile idiopathic arthritis: a single technology appraisal, June 2011

B Submissions or statement were received from the following organisations

I Manufacturer/sponsor:

- Roche Products

II Professional/specialist, patient/carer and other groups

- Arthritis Care
- British Health Professionals in Rheumatology
- British Society of Paediatric and Adolescent Rheumatology (BSPAR)
- National Rheumatoid Arthritis Society
- Royal College of Pathologists