Thursday, 1 September, 2011

[redacted]
National Institute for Health and Clinical Excellence
Level 1A, City Tower
Piccadilly Plaza
Manchester
M1 4BD

BY E-MAIL

Dear [redacted]

SINGLE TECHNOLOGY APPRAISAL
Tocilizumab for the treatment of systemic juvenile idiopathic arthritis

We welcome this opportunity to comment on the ACD for tocilizumab in systemic juvenile idiopathic arthritis (sJIA).

We have provided additional health economic analyses and clinical information requested respectively in sections 1.2 and 1.3 of the ACD. Following these analyses and information, we submit our comments on other sections of the ACD for the Institute's consideration.

Should you have any questions regarding this response or supporting materials, please do not hesitate to contact me.

Kind Regards,

[redacted]

Part A. Additional economic analysis requested in ACD section 1.2

The following economic modeling requests were made in the ACD:

1.2 The following further information on clinical and cost effectiveness should be provided by the manufacturer:

- A revised economic model that allows transitions between ACR (American College of Rheumatology)-response categories within a treatment line and
    - restructures the health states defined by Childhood Health Assessment Questionnaire (CHAQ) scores and uses treatment response to define transition probabilities or
    - uses a patient-level simulation to define the cost and utility values of a health state depending on both the starting CHAQ score and the change in CHAQ score in relation to the ACR (American College of Rheumatology) response.
- A revised manufacturer's base case with the assumption that treatment starts at age 5 years.
- For the comparison of tocilizumab and anakinra a revised base case using the primary outcome (ACR30 response and no fever).
- A fully incremental cost-effectiveness analysis of tocilizumab compared with anakinra and infliximab.
- Sensitivity analysis on the revised base cases that includes:
    - the uncertainty around the adjustment factor derived from the etanercept study used to take account of the other juvenile idiopathic arthritis subgroups in the infliximab study
    - a 'stopping rule' for tocilizumab after 2 years of treatment
    - a decreased frequency of administration of tocilizumab after 6 months to a 4-weekly regimen.
- Scenario analyses that include a fully incremental cost-effectiveness analysis of tocilizumab compared with anakinra and infliximab with:
    - infliximab as first-line treatment followed by tocilizumab
    - infliximab as first-line treatment followed by anakinra.

To respond to the Appraisal Committee's request, we describe our amendments to the economic model under each of the points from section 1.2 of the ACD.

A revised economic model that allows transitions between ACR (American College of Rheumatology)-response categories within a treatment line and:

- **restructures the health states defined by Childhood Health Assessment Questionnaire (CHAQ) scores and uses treatment response to define transition probabilities OR**
- **uses a patient-level simulation to define the cost and utility values of a health state depending on both the starting CHAQ score and the change in CHAQ score in relation to the ACR (American College of**

Rheumatology) response.

Background

The Roche economic model uses a published regression equation (originally developed to map HAQ score to utility values in adult rheumatoid arthritis [RA]) to transform CHAQ scores into utility scores. Therefore, the way in which CHAQ is modeled has the potential to significantly influence cost-effectiveness results.

Our economic modeling approach compares sequences of treatment rather than single treatments. In the original model, the following comparison made up the base case for the methotrexate inadequate-responding (MTX-IR) population[1]:

Tocilizumab → Anakinra → Etanercept → Adalimumab → Palliation

versus

Anakinra → Etanercept → Adalimumab → Palliation

In the original model, hypothetical patients entered the model at baseline, all with the same, assumed CHAQ score (and utility level derived from this CHAQ score). The assumed CHAQ score corresponded with the mean CHAQ score observed in the TENDER trial at baseline (pooled across all treatment arms).

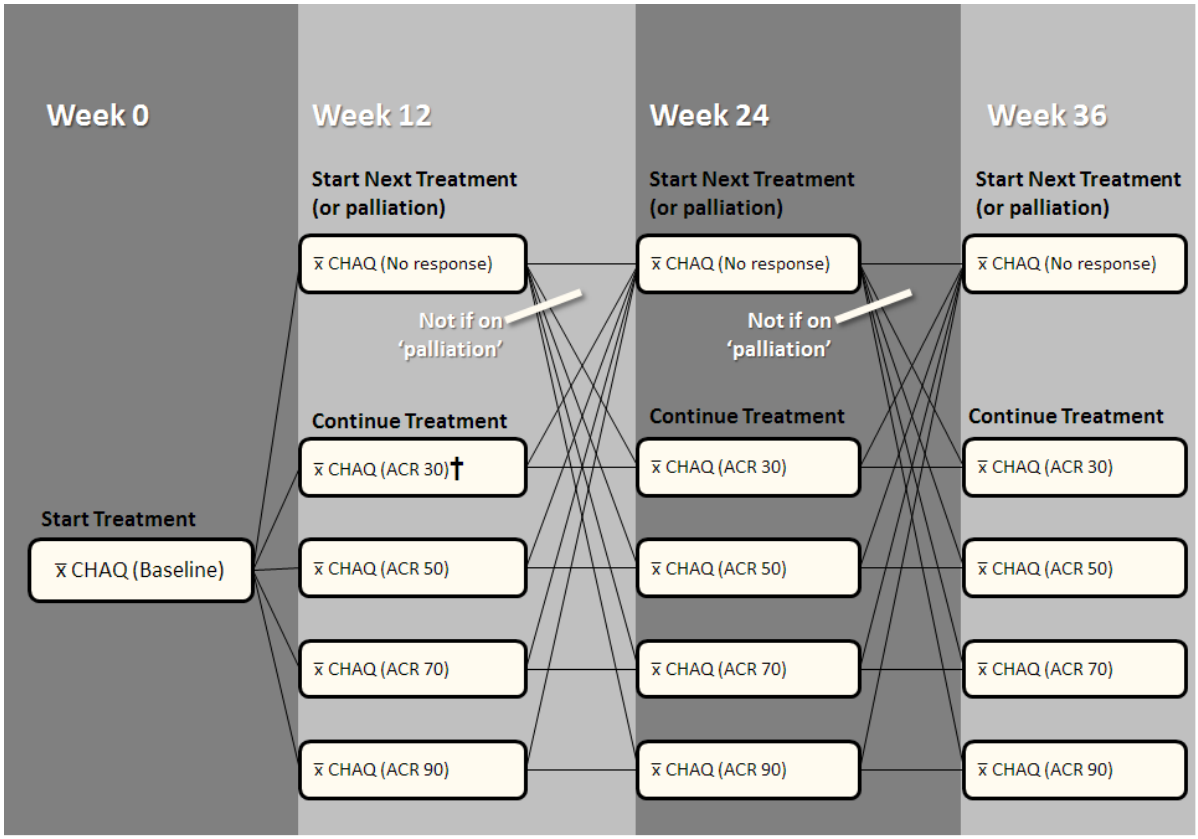
Patients remained on this baseline level of CHAQ (and associated utility) for 12 weeks. At week 12, patients were assigned to one of five health states in proportions which matched the week 12 ACR response (or non-response) observed for their allocated treatment. To each of the five 'ACR response category' health states we applied an average CHAQ score (and corresponding utility); this average was derived from the mean change in CHAQ observed within each ACR response category of the TENDER trial at week 12 (irrespective of treatment arm).

With each successive 12-week treatment cycle, patients remained in the 'ACR response category' health state they attained at week 12, unless they became 'ACR non-responders' due to discontinuation or loss of treatment efficacy. In such cases their treatment was switched to the next in their allocated sequence. Upon switching treatments, health states were re-assigned according to the expected ACR response distribution of the next treatment. Once active treatments were exhausted and 'palliation' was reached, patients were assumed to remain in the 'ACR non-responder' category.

Figure 1 provides an illustration of the old model flow.

Figure 1. Allocation of health states and CHAQ in old Roche model

[1] We have since learned at the Appraisal Committee hearing that the biologics subsequent to tocilizumab and/or anakinra should not be included in a model as this over-complicates the analysis and is likely to misrepresent clinical practice.



† x̄ CHAQ scores by ACR response type come from the TENDER trial dataset at week 12

We interpret the Appraisal Committee's primary concern to be that using average CHAQ scores and simply linking these to ACR response category -a relative measure of symptom improvement- may have misrepresented the natural history of sJIA. Whilst we will take the opportunity to comment on this assessment by the Committee later in this document, we also recognize the Committee's right to request an alternative analysis and have endeavoured to meet this request.

Redesign – Introduction

Our new model incorporates elements from both designs suggested by the Appraisal Committee in section 1.2 of the ACD. We feel that the new design appropriately captures the spirit of the Committee's information request. The economic model has been modified such that health states are defined according to categories of CHAQ, rather than being based on ACR response categories in which an average CHAQ/utility is applied. However, given the importance of ACR as an indicator of clinical response and its commonality as an outcome measure in sJIA (importantly allowing indirect comparisons to be made), we have not discarded ACR from the modeling process altogether. Rather, we have allowed ACR to remain as a potential predictor of CHAQ score.

Comparisons

In response to the Appraisal Committee's requests in 1.2, we have simplified our comparisons of sequences of treatment such that our base case now involves the

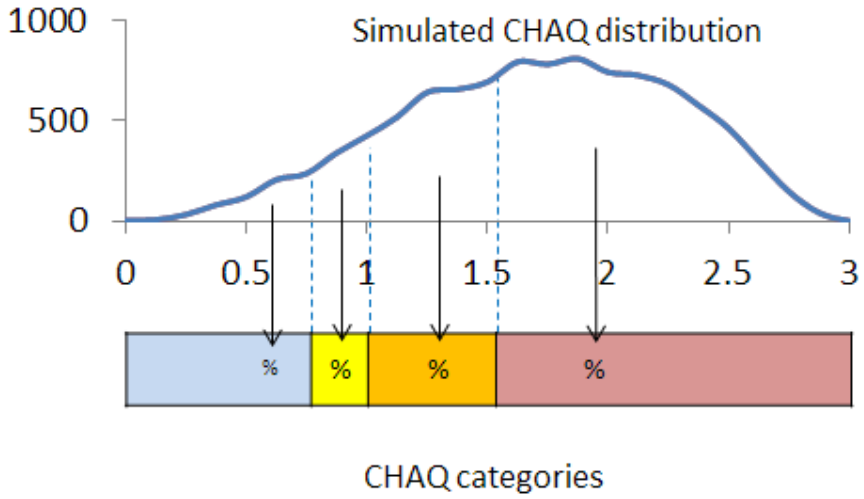
following comparison:

Tocilizumab → Anakinra → Palliation
versus
Anakinra → Palliation

Model structure

In our model redesign, we have adopted an approach whereby CHAQ categories define health states. However, a simulated patient distribution (n=10,000) of CHAQ score, based on TENDER, is used in the background to establish what proportion of patients will fall into each category. This principle is illustrated in Figure 2.

Figure 2. Simulated distribution of CHAQ scores is divided into 'CHAQ category' health states



In order to estimate the expected CHAQ distribution associated with each treatment, we developed a regression model based on the week 12 dataset from the TENDER trial. All patients were included in the regression dataset regardless of treatment type. The regression analysis tested the influence of the following variables on week 12 CHAQ score:

- ACR response category, i.e.:
  - No response
  - ACR30
  - ACR50
  - ACR70
  - ACR90
- Baseline CHAQ score
- Age (modeled as a continuous variable)

The results of this regression analysis can be found in Table 1. The model generally

provided a reasonable fit for the data with an R-square value of 48% and statistically significant model F-statistic ($F_{(6, 104)}$ = 15.73, p<0.001).

Table 1. Week 12 CHAQ regression analysis results

Variable	Coefficient	Standard error	p-value	95% Confidence interval	
Baseline CHAQ	0.459	0.076	<0.001	0.308	0.609
ACR30	-0.437	0.244	0.075	-0.920	0.046
ACR50	-0.428	0.225	0.060	-0.873	0.018
ACR70	-0.883	0.192	<0.001	-1.264	-0.502
ACR90	-1.098	0.186	<0.001	-1.466	-0.730
Age	0.016	0.013	0.229	-0.010	0.042

Number of observations = 111
F(6, 104) = 15.73, p<0.001
R-squared = 0.48

F-test for ACR terms combined:
F(4, 104) = 10.59, p<0.001

Not all variables representing ACR response were associated with statistically significant coefficients, however the combined effect of the ACR categories' inclusion to the model was statistically significant ($F_{(4, 104)}$ = 10.59, p<0.001). Age was tested in the regression, however because it was not an important predictor of week 12 CHAQ (β=0.016 units per year of age, p=0.229), it was subsequently excluded. Excluding age was also felt to be justified as our new base case (as per the Appraisal Committee's request) sets all subjects' starting ages to 5 thus reducing the need to build the effect of starting age variability into the model. The regression coefficients used in our economic model can be found in Table 2. Their application is explained in the next section.

Table 2. Final week 12 CHAQ regression analysis results (excluding age)

Variable	Coefficient	Standard error	p-value	95% Confidence interval	
Baseline CHAQ	0.466	0.076	<0.001	0.315	0.617
ACR30	-0.425	0.244	0.085	-0.908	0.059
ACR50	-0.425	0.225	0.062	-0.871	0.022
ACR70	-0.889	0.193	<0.001	-1.271	-0.507
ACR90	-1.123	0.185	<0.001	-1.489	-0.756

Number of observations = 111
F(5, 105) = 18.50, p<0.001
R-squared = 0.44

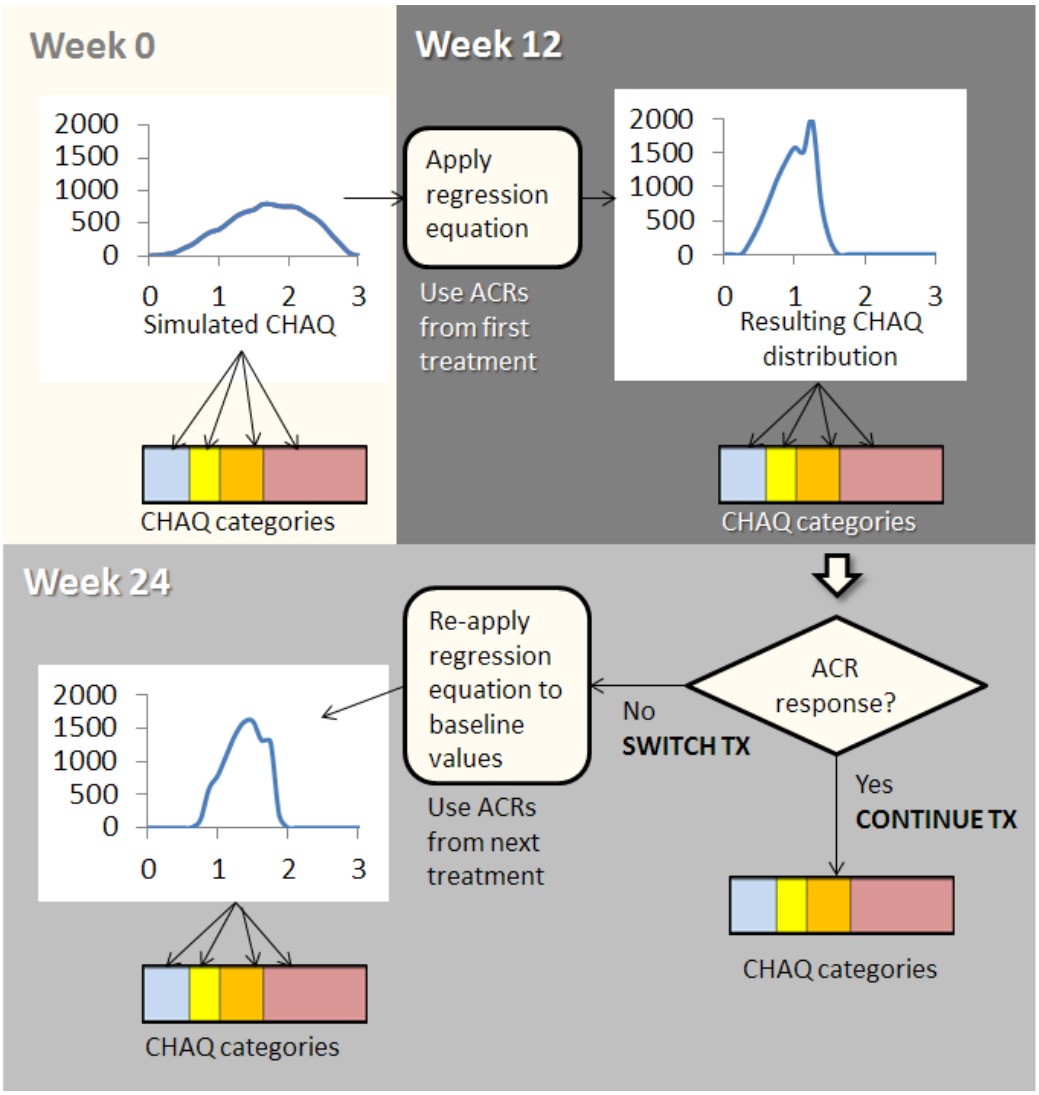
F-test for ACR terms combined:
F(4, 105) = 1, p<0.001

Model flow

All subjects begin with the same, simulated baseline distribution of CHAQ scores – this

was modeled on the baseline CHAQ distribution from the TENDER study. Subjects' allocation to CHAQ category health states is thus identical across all treatment arms at baseline. At week 12, the regression equation is applied to the baseline CHAQ values. The resulting values form a distribution and differ by allocated treatment group, because tocilizumab, anakinra, and infliximab were all associated with different ACR responses. ACR values used are the same as those used in our previous base case except that 'no fever' has been included in the indirect comparison with anakinra as per the Appraisal Committee's request, and the population has been limited to MTX-IR patients. A simplified view of the new model flow is shown in Figure 3. Only weeks 0 to 24 are shown, but as our base case compares treatment sequences which involve a maximum of two active drugs, this level of detail was felt to be sufficient to illustrate the principle.

Figure 3. New model flow shown for weeks 0 to 24



Movement between health states whilst on treatment

The new model behaves similarly to the old with respect to treatment switching and discontinuation. When ACR 'non-response' is encountered, the treatment is switched and health states are modified to reflect the efficacy of the new treatment. If ACR response is sustained, no re-calculation is applied to the CHAQ distribution or health

states related to this.

The Appraisal Committee has requested that our revised economic model should allow "movement between categories of ACR response while on treatment". We have no clinical data to inform an estimate of the likelihood of transition between ACR categories whilst on treatment. But we are mindful of the Committee's concerns about the original model's omission of health state variation whilst on a treatment. We are also mindful that in the absence of clinical evidence, any assumptions about such variation will be arbitrary and the effect on model behaviour can be difficult to interpret. The ERG made a similar observation in their assessment report.

To explore this issue further, we prepared trace diagrams of the ACR response categories expected in the model, and observed how these changed when sets of simple rules about transition between ACR categories was introduced.

Figure 4. ACR response categories over time whilst on treatment - 10% improvement or 10% deterioration per cycle

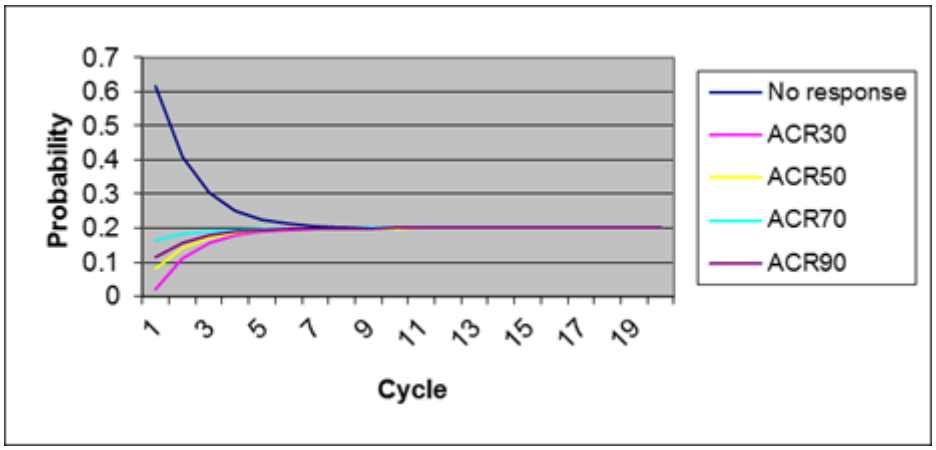


Figure 5. ACR response categories over time whilst on treatment - 10% improvement, 0% deterioration per cycle

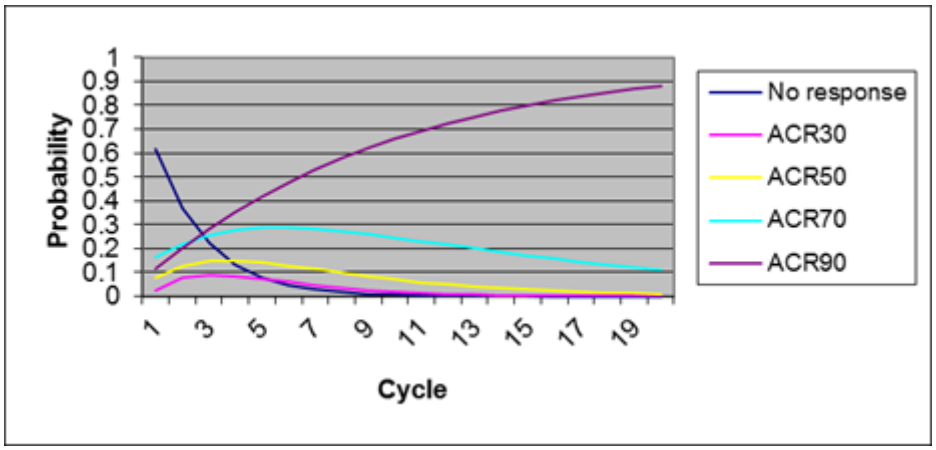


Figure 6. ACR response categories over time whilst on treatment - 10% improvement, 2% deterioration per cycle

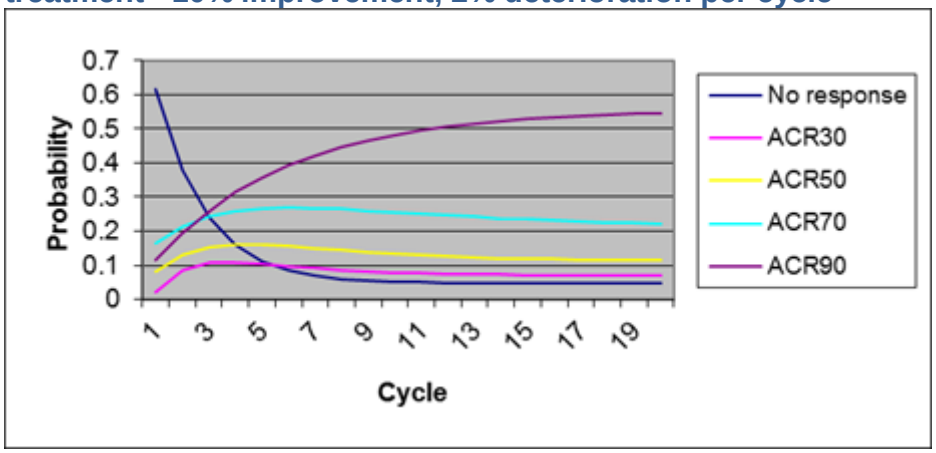
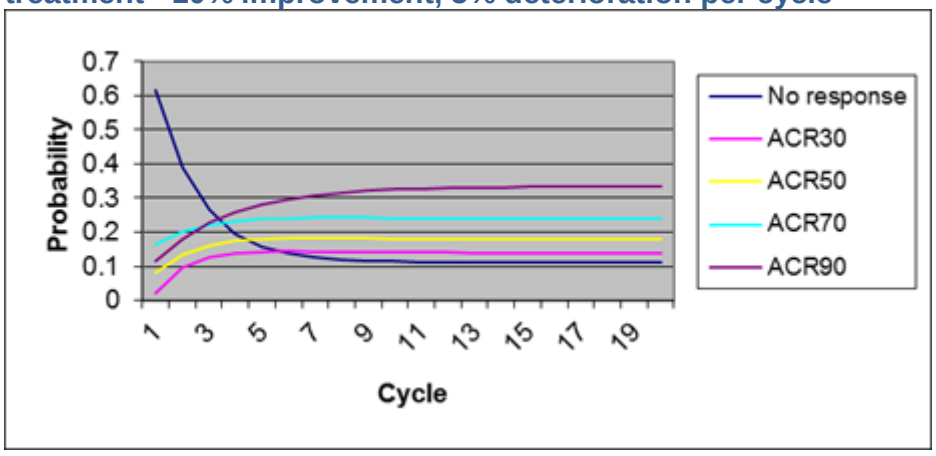


Figure 7. ACR response categories over time whilst on treatment - 10% improvement, 5% deterioration per cycle



As the figures show, each adjustment caused marked changes in the proportions of patients assumed to remain in each ACR category. When an equal probability of going 'up' or 'down' on ACR category was assumed at each cycle on treatment, subjects rapidly tended toward a perfectly uniform distribution of ACR category. When an adjustment stipulated a greater likelihood of high ACR responses, ACR 90 tended to dominate. These effects often eclipsed any effects associated with the treatments

themselves. We therefore considered it inappropriate to revise our base case to include any of these scenarios. Our new base case maintains a zero probability of transitioning between ACR categories in any patients who have responded to, and remain on, a particular treatment. Should it still be considered necessary, ICERs could be calculated for a selection of possible approaches. However, as our diagnostic analyses suggest that the modeled treatment efficacy reacts quite dramatically to these adjustments, we can only recommend caution around any interpretation of ICERs produced using this method.

Definition and costing of CHAQ categories

In relation to the Committee's request to devise new health states based on categories of CHAQ, we interviewed clinicians about how sensible categories might be drawn up. Respondents came from the same pool of subjects as previously contacted (for our initial submission) about resource use. We interviewed three clinicians on the subject of defining categories of CHAQ, all by personal communication (PC) on 11/8/2011:

1. Prof Pat Woo (Great Ormond Street Hospital)
2. Dr Gavin Cleary (Alder Hey, Liverpool)
3. Dr A Ramanan (Bristol Royal Hospital for Children)

Each clinician was asked to comment on the following questions:

1. **Are there clinical cut-off points for CHAQ in terms of say, mild, moderate, severe or uncontrolled sJIA disease?**
2. **If so, are you aware of the expected resource use for these categories (in particular, days of hospital in-patient stay)?**

In the clinician's responses the following key themes emerged and were unanimous:

- There are no clinically relevant cut-off points of CHAQ in sJIA. CHAQ measures disease activity in JIA; the only sJIA-specific disease activity measurement tools are those proposed by Batthish and Ramanan (see references Batthish et al 2005, Ramanan et al 2005) papers, as well as JIA ACR response measures.
- CHAQ (for general JIA) only has scores of 0 (that represent no disability) and 3 (to represent very severe disability). CHAQ is not used regularly as the sole indicator of disease in the UK.
- JIA ACR (+/-fever) response is the gold standard validated tool as it measures not just CHAQ but other composites that are most relevant for systemic disease.
- CHAQ is interpreted differently by different physicians, and because sJIA is a systemic disease, other factors like active joints, fever, serositis, hepatosplenomegaly, erythrocyte sedimentation rate (ESR) are more important indicators of disease severity.
- No data or experience is available to inform the estimation of costs by CHAQ.

In the absence of further clinical guidance or any clear cues from the TENDER dataset, we have selected CHAQ cutpoints based on a suggestion made in the ERG report which forms part of the ACD consultation. We recognize that the ERG specified these values purely hypothetically, and that CHAQ distribution by treatment arm is expected to drive QALY estimates and ICERs. However because we were unable to locate any evidence-based method, we felt the cut-points suggested by ERG represented a reasonable starting point. Our new base case health states are shown in Table 3.

Table 3. CHAQ cut-points used to define health states in new base case

CHAQ-defined health state	Cut-points used in base case
'Controlled'	≥0 to ≤0.75
'Mild'	>0.75 to ≤1
'Moderate'	>1 to ≤1.5
'Severe'	>1.5 to ≤3.000

Costs arising from resource use other than active treatment

As our clinician interviews did not yield any guidance in relation to the estimation of costs per health state let alone their definition according to CHAQ, we have adapted our existing methodology in a way which we hope will address some of the Appraisal Committee's concerns around our original use of costs by ACR response category.

In the ACD, the Committee heard that the costs associated with health states in the original Roche model were reflective of UK clinical practice. However, a concern was expressed with regard to the large increase between the 'no response' and 'ACR 30' health states.

In the revised model, we have used the costs from the original model's 'no response' and 'ACR 90 health states', to represent categories of highest and lowest burden respectively. For intermediate categories of disease we have linearly interpolated these two extreme values to produce a smooth transition of costs. This adjustment is based on an assumption of continuity rather than any new clinical evidence. We consider this approach more conservative than our original set-up, as we had previously found that costs decreased dramatically as soon as patients achieved disease control, owing to the reduced incidence of hospital in-patient stays. The new costs are shown in

Table 4.

Two points to note regarding this interpolation procedure:

- for both utility and costs, we have corrected the values per health state to reflect benefit relevant to 12-week cycles, rather than three-month cycles
- for both utility and costs we have adjusted the cost per cycle to line up with a 'mid-point' CHAQ value rather than a value expected at the extremes of the CHAQ categories

Table 4. Allocation of costs to CHAQ category

Health state	CHAQ category (range)	Assumed CHAQ mid-point for linear interpolation	Cost assumed per cycle (£) (linear interpolation)	Original ACR health state cost
'Severe'	≥0 to ≤0.75	2.3125	£3,360.47	£3640.51 ('no response')
'Moderate'	>0.75 to ≤1	1.3125	£1,804.30	
'Mild'	>1 to ≤1.5	0.9375	£1,220.73	
'Controlled'	>1.5 to ≤3.000	0.375	£345.38	£374.16 ('ACR 90')

Estimation of utility per CHAQ category

For each CHAQ category, we allocated a mean utility which corresponds, by way of our existing mapping algorithm, to the mid-point CHAQ in that category.

Table 5. Allocation of utility scores to CHAQ categories

Health State	CHAQ category (range)	Mid-point CHAQ	Corresponding Utility level
'Controlled'	≥0 to ≤0.75	0.0625	0.7689
'Mild'	>0.75 to ≤1	0.5000	0.6554
'Moderate'	>1 to ≤1.5	1.2500	0.5550
'Severe'	>1.5 to ≤3.000	2.3125	0.1913

For the comparison of tocilizumab and anakinra a revised base case using the primary outcome (ACR30 response and no fever).

As requested, our efficacy inputs have been adjusted to reflect this. These can be found in the economic model inputs. However, please note that the 'no fever' outcome was not captured in the infliximab trial and thus our indirect comparison of tocilizumab with infliximab still operates on the ACR30 response parameter excluding the 'no fever' outcome.

A revised manufacturer's base case with the assumption that treatment starts at age 5 years

The model's new assumed starting age is 5 years.

A fully incremental cost-effectiveness analysis of tocilizumab compared with anakinra and infliximab

As requested, we provide a fully incremental analysis for all possible treatment sequences in the results. Please note that the 'no fever' outcome was not captured in the infliximab trial and thus our indirect comparison of tocilizumab with infliximab still operates on the ACR30 response parameter excluding the 'no fever' outcome.

Sensitivity analysis on the revised base cases that includes:

- **the uncertainty around the adjustment factor derived from the etanercept study used to take account of the other juvenile idiopathic arthritis subgroups in the infliximab study**
- **a 'stopping rule' for tocilizumab after 2 years of treatment**
- **a decreased frequency of administration of tocilizumab after 6 months to a 4-weekly regimen**

Our modeling results take these suggestions on board.

Scenario analyses that include a fully incremental cost-effectiveness analysis of tocilizumab compared with anakinra and infliximab with:

- **infliximab as first-line treatment followed by tocilizumab**
- **infliximab as first-line treatment followed by anakinra**

Noted. These have been included in our analysis. Please note that the 'no fever' outcome was not captured in the infliximab trial and thus our indirect comparison of tocilizumab with infliximab still operates on the ACR30 response parameter excluding the 'no fever' outcome.

Results

Revised base case

Because of the diversity of treatment sequences to be considered in this analysis, we have presented cost and QALY estimates for infliximab-only, anakinra-only, tocilizumab→infliximab, tocilizumab→anakinra all together in

Table 6, with incremental cost effectiveness ratios (ICERs) for comparisons with anakinra and infliximab as requested. Overall, the ICERs for tocilizumab treatment sequences compared to sequences without indicate that tocilizumab is likely to be a cost-effective addition to standard care involving either anakinra alone or infliximab alone.

Table 6. Cost and QALYs for base case and alternative base case analyses

Treatment sequence	Total cost	Total QALYs	Inc. Cost	Inc. QALYs	ICER	Comparison
I	£114,593.33	2.7709				
A	£136,871.46	3.3700				
TI	£151,439.61	4.1262	£36,846.28	1.3553	£27,186.60	vs. I
TA	£165,321.05	4.3310	£28,449.59	0.9609	£29,606.23	vs. A

NOTES: I = infliximab, A = anakinra, T = tocilizumab

Fully incremental analysis results

For the full incremental analysis, we have presented costs and QALYs for all treatment sequences relevant to the appraisal consultation. Because of the nature of the combination treatments of interest in the ACD, a fully incremental analysis would be of less value given the Appraisal Committee's request. Therefore we have given ICERs which we believe meet the information request in section 1.2 of the ACD rather than ones which strictly conform to an incremental analysis.

Please note that the results show that it was necessary to use efficacy statistics excluding the 'no fever' outcome when the treatment sequence included infliximab as monotherapy or in a sequence involving tocilizumab.

Table 7. Incremental analysis results

Treatment sequence	Total cost	Total QALYs	Inc. Cost	Inc. QALYs	ICER	Comparison	Outcome used
I	£114,593.33	2.7709					ACR-NoFever
IA	£127,802.55	3.6062					ACR+NoFever*
A	£136,871.46	3.3700	£9,068.91	-0.2361	Dominated	vs. IA	ACR+NoFever
IT	£139,674.57	3.3848	£25,081.24	0.6139	£40,855.96**	vs. I	ACR-NoFever
TI	£151,439.61	4.1262	£36,846.28	1.3553	£27,186.60	vs. I	ACR-NoFever
TA	£165,321.05	4.3310	£37,518.50	0.7248	£51,765.09	vs. IA	ACR+NoFever

NOTES: I = infliximab, A = anakinra, T = tocilizumab; *This sequence assumes that ACR response rates of I = A; **Extendedly dominated by TI

Base case sensitivity analyses

Varying adjustment factor from indirect comparison with infliximab

As described, we present the base case results with a set of modifications to the adjustment factor used in our indirect comparison with infliximab. These can be found in Table 8 and Table 9. As can be seen the ICERs are increased or decreased by around £3,000 per QALY as the adjustment factor is increased or decreased respectively.

Table 8. Results assuming increase of the adjustment factors by 30%

Strategy	Total cost	Total QALYs	Inc. Cost	Inc. QALYs	ICER	Comparison	Outcome used
I	£112,068.97	2.9051					ACR-NoFever
A	£136,871.46	3.3700					ACR+NoFever
TI	£150,250.86	4.1894	£38,181.89	1.2843	£29,729.84	vs. I	ACR-NoFever
TA	£165,321.05	4.3310	£28,449.59	0.9609	£29,606.23	vs. A	ACR+NoFever

NOTES: I = infliximab, A = anakinra, T = tocilizumab

Table 9. Assume decrease of the adjustment factors by 30%

Strategy	Total cost	Total QALYs	Inc. Cost	Inc. QALYs	ICER	Comparison	Outcome used
I	£117,031.11	2.6404					ACR-NoFever
A	£136,871.46	3.3700					ACR+NoFever
TI	£152,587.59	4.0647	£35,556.48	1.4243	£24,963.43	vs. I	ACR-NoFever
TA	£165,321.05	4.3310	£28,449.59	0.9609	£29,606.23	vs. A	ACR+NoFever

NOTES: I = infliximab, A = anakinra, T = tocilizumab

Stopping or dose-reduction scenarios

In Table 10 and Table 11, the base case results are re-presented with stopping rules where tocilizumab is no longer given after either two years or where its dosing frequency is halved after six months. In both scenarios, tocilizumab sequences either dominate those without, or in the case of tocilizumab (half dose)→infliximab compared to infliximab, produced an ICER less than £1,000 per QALY gained. Total costs are lower in the tocilizumab strategies in Table 10 because:

- the shorter tocilizumab treatment duration has a strong effect on lifetime cost
- in the TI and TA strategies, infliximab and anakinra are given for a shorter duration

Table 10. Results assuming stopping rule after two years

Strategy	Total cost	Total QALYs	Inc. Cost	Inc. QALYs	ICER	Comparison	Outcome used
I	£114,593.33	2.7709					ACR-NoFever
A	£136,871.46	3.3700					ACR+NoFever
TI	£90,314.50	4.1262	-£24,278.82	1.3553	Dominant	vs. I	ACR-NoFever
TA	£104,195.94	4.3310	-£32,675.52	0.9609	Dominant	vs. A	ACR+NoFever

Table 11. Results assuming administration every four weeks after the first six months

Strategy	Total cost	Total QALYs	Inc. Cost	Inc. QALYs	ICER	Comparison	Outcome used
I	£114,593.33	2.7709					ACR-NoFever
A	£136,871.46	3.3700					ACR+NoFever
TI	£115,493.34	4.1262	£900.02	1.3553	£664.07	vs. I	ACR-NoFever
TA	£129,374.78	4.3310	-£7,496.68	0.9609	Dominant	vs. A	ACR+NoFever

Part B. Additional clinical data requested in ACD 1.3

Radiographic evidence of progression of joint damage for patients receiving tocilizumab in sJIA

The TENDER study:

Radiographic results from the phase III study, TENDER are not yet available. The X-rays in the TENDER study have not yet been read and scored by the collaborative group, Paediatric Rheumatology International Trials Organization (PRINTO). [redacted]

Other published radiographic studies:

The radiologic effect on large joints during tocilizumab treatment in children with sJIA has been investigated in a case series with 7 patients (Inaba et al. 2011, 2007). The large joints studied were, shoulders, elbows, hips, knees and ankles and evaluation included joint space narrowing, cyst formation, erosion and localized growth abnormalities. A modified Larsen method was employed with a grading scale from 0 (normal) to 5 (mutilating changes) for each joint. By measuring 10 joints the Larsen score therefore ranged from 0-50. Radiographs were compared to a healthy child of the same bone age and results were agreed upon by an experienced orthopaedist and paediatric rheumatologist.

The patient characteristics are shown in Table 12.

Table 12. Inaba 2011 Patient characteristics

Patient	Gender	Age at disease onset	Age at tocilizumab treatment
1	M	2	9.2
2	M	2.1	10.4
3	F	2.9	9.9
4	M	8.2	13.6
5	F	6.9	9.1
6	M	3.9	10
7	M	2.8	3.4
Average	F:2; M:5	4.1	9.4

Patients received 8mg/kg of tocilizumab every 2 weeks. The average follow up months (range 55-57). At the final follow up radiographic abnormalities had all decreased with the exception of growth abnormalities. The results are shown in

Table 13.

Table 13. Inaba 2011 Results

	Before tocilizumab treatment	After tocilizumab treatment	P value
Juxta-articular osteoporosis	85.7%	15.9%	not available
Soft tissue swelling	45.7%	5.7%	not available
Joint space narrowing	42.9%	30.7%	<0.05
Sub-chondral bone cysts	12.9%	1.4%	<0.01
Erosions	21.4%	0%	<0.01
Localized growth abnormalities	15.7	36.4	<0.01
Active joints	4.6	0.3	not available
Larsen score	17.6	10.1	<0.05

There were improvements in radiographic changes in 57% of joints, worsening in 13% and no change in 30%.

Joint space narrowing was measured in the hand by the Poznanski method. The Poznanski score did not change after tocilizumab treatment, however it correlated significantly with the total Larsen score ($p<0.05$).

The authors concluded that there were marked radiographic improvements of damaged large joints in tocilizumab-treated sJIA patients. Radiographic findings of joint destruction, such as joint space narrowing, subchondral bone cysts and erosion, all diminished after tocilizumab treatment, and osteoporosis showed marked amelioration. Moreover, damaged joints were remodelled, and improvement of radiographic findings was maintained for about 5 years.

The limitations of this study are small sample size and radiographic deterioration in some joints (13%), despite stabilization of systemic inflammatory responses. Further studies with a larger number of patients are needed.

The same authors have subsequently presented results for 20 patients (Inaba et al. 2009). All patients received 8mg/kg every 2 weeks. Hips, knees, ankles, shoulders and elbows were evaluated for soft tissue swelling, juxta-articular osteoporosis, epiphyseal irregularity, joint space narrowing (JSN), cyst formation, erosion, and localised growth disturbance, along with minimal joint space width.

The mean duration of treatment and follow-up was 41 months (11-82 months).

Table 14. Inaba 2011 - joint results

	Before tocilizumab treatment	Final follow-up
Mean active joints	3.3 (0-12)	0.3 (0-4)
Juxta-articular osteoporosis	83.7%	29.8%
Soft tissue swelling	44.9%	9.1%

All radiologic abnormalities improved at the final follow-up, except localised growth disturbances. At the final follow up, localised growth disturbances at 27.9%, mild juxta-articular osteoporosis at 29.8% and mild joint space narrowing at 19.7% were the most frequent abnormalities.

The authors concluded that tocilizumab demonstrated radiographic improvement of the large joints in sJIA.

A study from Japan included 46 patients receiving 8mg/kg every 2 weeks for systemic-onset JIA (Kaneko et al. 2009). The study objective was to investigate clinical and laboratory features of systemic inflammation during tocilizumab treatment in these patients. Patients were assessed for clinical response by tender and swollen joints, serum markers including metalloproteinase 3 (MMP-3) and inflammatory cytokine levels, and also radiographic joint damage.

Table 15. Kaneko 2009 Baseline characteristics

Characteristic	**Value**
Mean age at disease onset	4 (1-9) years
Disease duration pre tocilizumab	1.8 (0.3-5.8) years
Mean follow up	3 (1-4.7) years
Mean steroid dose pre treatment	1.0mg/kg
Mean steroid dose post treatment	0.3mg/kg

Markers of systemic inflammation and numbers of tender/swollen joint counts were markedly improved following tocilizumab treatment. However progression of joint damage was observed in weight-bearing joints such as hip (85%) and knee (57%), along

with growth disturbances and osteopenia. The radiographic progression was not seen in the small joints.

Increased corticosteroids, addition of methotrexate and shortening the interval of tocilizumab treatment were effective for decreasing MMP-3 and IL-6 levels. The authors conclude that persistent high levels of MMP-3 were observed in a few patients despite improvement of systemic symptoms and serum inflammatory markers. These high levels correlated with IL-6 levels and lead to weight-bearing joint damage.

A small study in Japan has investigated the effect of tocilizumab on serum cartilage oligomeric matrix protein (COMP) levels in 11 patients with sJIA (Nakajima 2009).

COMP concentration is a biomarker of growth cartilage turnover in the growth phase of children. In this study COMP levels were examined during active disease phase and remission following tocilizumab treatment in 11 sJIA patients, and compared to 201 healthy controls. Serum bone alkaline phosphate (BAP) levels were also measured as a marker of bone constructing/remodelling.

Via the healthy controls it was shown that serum COMP levels were not related to age. During the active disease phase in sJIA patients, mean serum COMP levels were decreased as compared to control. During treatment with tocilizumab levels were significantly higher as compared to the active phase ($p<0.05$).

MMP-3 levels and BAP levels also showed improvement during tocilizumab treatment. MMP-3 levels decreased as compared to the active phase ($p<0.01$) and BAP concentrations increased ($p<0.01$) during the remission phase. These results suggest that the suppressed growth cartilage turnover of sJIA was improved during treatment with tocilizumab. The increased levels of BAP also suggest that growth cartilage turnover and bone turnover are improved due to inhibition of inflammation by tocilizumab.

Figure 8. Serum cartilage oligomeric matrix protein (COMP) concentration in the control and tocilizumab groups

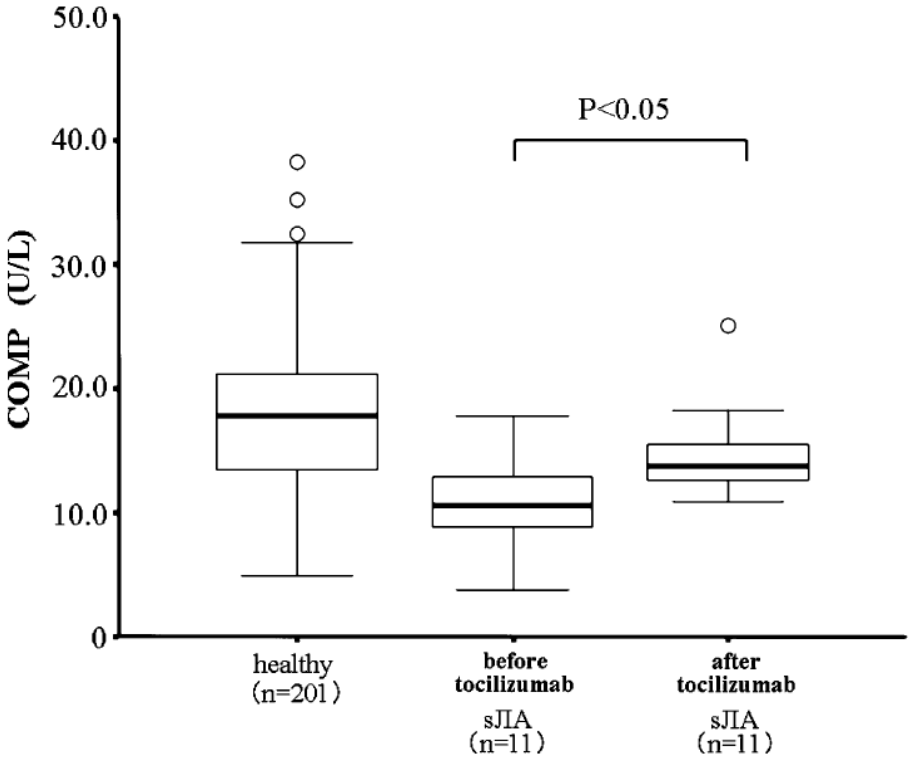
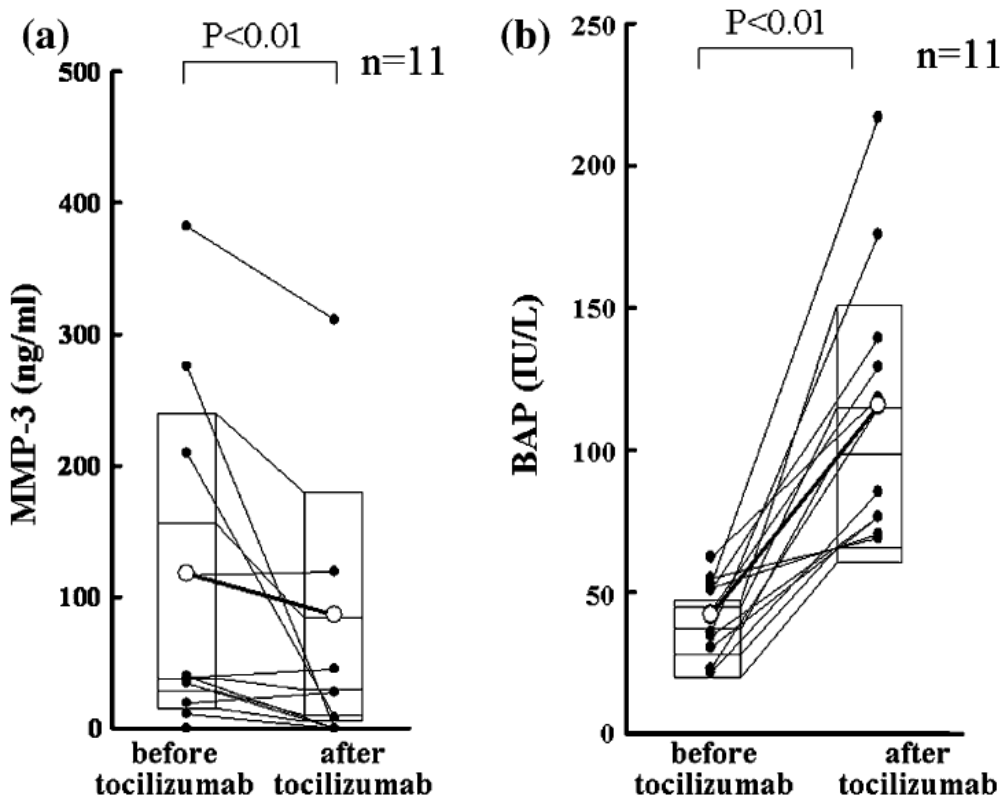


Figure 9. a) Serum matrix metalloproteinase-3 (MMP-3) concentrations during the active disease and remission phases in the systemic juvenile idiopathic arthritis (sJIA) patients. b) Serum bone alkaline phosphatase (BAP) concentrations during the active disease



The authors conclude that tocilizumab treatment improves both growth plate cartilage turnover and bone turnover in sJIA patients.

Long-term follow-up data from the trials being conducted in Japan

Published data:

Phase II extension

As previously described, a dose escalation study in which 11 patients with sJIA were treated with tocilizumab was completed (Yokota et al. 2005a). All patients were treated every 2 weeks, at a starting dose of 2mg/kg. If CRP levels were >1.5mg/dl on day 5 or more after the infusion, the dose was escalated to 4mg/kg. If this dose did not stabilise CRP levels the dose was escalated to 8mg/kg. Patients then entered an open-label 3 year extension phase receiving 2-8mg/kg every 2 weeks (Yokota et al. ACR 2005b). Fixed NSAID, cyclosporine and methotrexate doses were permitted, but other biologics and parental steroids were not. The primary endpoint was JIA ACR30.

Treatment was received for 10-35 months in total. During the extension phase the tocilizumab dose was adjusted to maintain low disease activity and inflammatory markers, while tapering off steroids. 10/11 patients achieved JIA ACR 70 improvement

and 10 patients successfully reduced steroid dose. IL-6 levels after the initial administration abruptly and paradoxically increased. With maintenance doses IL-6 levels 1) gradually decreased, 2) continuously fluctuated or 3) spontaneously fluctuated with no changes of CRP or ESR. While patients were suffering an episode of infection IL-6 and CRP levels increased. There were no deaths or malignancies, but one patient withdrew due to duodenum perforation at 10 months. The authors suggest this could be due to long-term steroid and NSAID use. The most serious AEs were pneumonia in 2 patients.

Phase II/III extension

This study included a total of 128 patients, recruited from phase II and III studies, with additional patients recruited directly into the extension phase (Yokota et al. 2008). Patients received tocilizumab 8mg/kg every 2 weeks, with efficacy measured every 12 weeks using JIA ACR scores.

Table 16. Yokota 2008 Baseline characteristics

Characteristic	Value
Median age	9 years
Gender	43% male
Mean disease duration	4 years
Median corticosteroid dose	0.5mg/kg/day
Median duration of tocilizumab treatment	78 weeks

Table 17. Yokota 2008 Efficacy results

	Week 48 (n=78)	Week 96 (n=58)	Week 144 (n=41)
JIA ACR30	94%	100%	100%
JIA ACR50	88%	98%	100%
JIA ACR70	81%	93%	90%

Table 18. Yokota 2008 Safety results

Adverse events	120 (94%) patients
AEs per 100 PY	787.1
Withdrawals: **due to AE** **due to anti-tocilizumab antibodies** **lack of efficacy**	14 8 5 1
Serious AEs (per 100 PY)	37.2
Serious infections (per 100 PY)	14.5
Most frequent AEs: **Gastroenteritis (per 100 PY)** **Pneumonia (per 100 PY)**	 3.8 3.4

The AEs which lead to treatment discontinuation were macrophage activation syndrome (MAS), anaphylactoid reaction (2 patients), cardiac amyloidosis, duodenal perforation, gastrointestinal hemorrhage and infusion reaction (2 patients). There were 2 deaths,

one due to MAS and one cardiac amyloidosis.

At the analysis time point there were 4 patients in remission receiving no drug therapy.

Phase II/III extension

Results of the phase II and III extension studies have also been presented. This included 67 patients (11 from phase II and 56 from phase III) (Yokota et al 2009a).

Table 19. Yokota 2009 Baseline characteristics

Characteristic	Value
Median age	8 years (9-19)
Gender	43% male
Median disease duration	3.8 years (0.4-16.2)
Median duration of tocilizumab treatment	185 weeks

Of the 67 patients, 53 entered the fourth year of continuous treatment (79%). The most frequently observed non-serious AEs were nasopharyngitis, upper respiratory tract infection and gastroenteritis. There were no opportunistic infections, malignancies, autoimmune diseases or death.

Table 20. Yokota 2009 Safety results

Withdrawals: **due to SAE** **due to anti-tocilizumab antibodies** **lack of efficacy**	9 4 4 1
Serious AEs (per 100 PY)	35.5
Serious infections (per 100 PY)	13.6

At enrolment all patients were receiving corticosteroids. At week 168, 77% of patients had reduced their doses by at least 50%.

Table 21. Yokota 2009 Efficacy results

	Week 96 (n=58)	Week 168 (n=51)
JIA ACR30	100%	96%
JIA ACR50	98%	96%
JIA ACR70	93%	88%
JIA ACR90	64%	73%

There were 8 patients who discontinued tocilizumab treatment due to sustained remission. Six of these patients subsequently flared and resumed tocilizumab treatment. These 8 patients were selected from a possible 24 as candidates for complete withdrawal from tocilizumab therapy. These patients from the long-term extension had successfully withdrawn corticosteroid therapy and had low IL-6 levels for 1 year or longer (Yokota et al. 2009b).

Discontinuation was 2.6-4.9 years after initiation of tocilizumab treatment. Of the 6 patients who restarted tocilizumab therapy, resumption was between 46-571 days. One of these patients restarted with both tocilizumab and steroid, but the other 5 resumed only tocilizumab. These 5 patients achieved JIA ACR90 within 7-66 days.

Table 22. Yokota 2009 Features of patients in clinical remission

	24 possible drug withdrawal patients	8 actual drug withdrawal patients
Age	2-18	3-10
Gender	42% boys	50% boys
Duration of sJIA	0.4-8.2 years	0.5-8.2 years
CRP	0.3-29 mg/dl	0.3-15 mg/dl
Steroid	3.8-40 mg/body	3.8-30 mg/body

During the tocilizumab-free period there were 22 AEs in the 8 patients. These were mostly mild in severity and resolved without medical intervention. There were 2 patients who suffered upper respiratory infections which were suspected to lead to sJIA flare.
Further investigations have looked at clinical remission with tocilizumab in sJIA patients (Yokota et al. 2010). The 67 patients from phase II and III studies were observed for a median of 3.5 years. There were 59/67 (88%) patients who achieved inactive disease and 45/67 (67%) achieved clinical remission. A state of inactive disease during clinical remission was reached at a median of 7 months (1-37), and continued for a median of 32 months (8-68). Of the 8 patients who halted treatment due to remission, 3 maintained inactive disease with no treatment for 12 months.

There was no association between gender, age of JIA onset, age at tocilizumab treatment initiation, corticosteroid dose, CRP, IL-6, active joint, physician's global assessment (VAS), or the Childhood Health Assessment Questionnaire (CHAQ) at tocilizumab treatment initiation, and patients achieving remission.

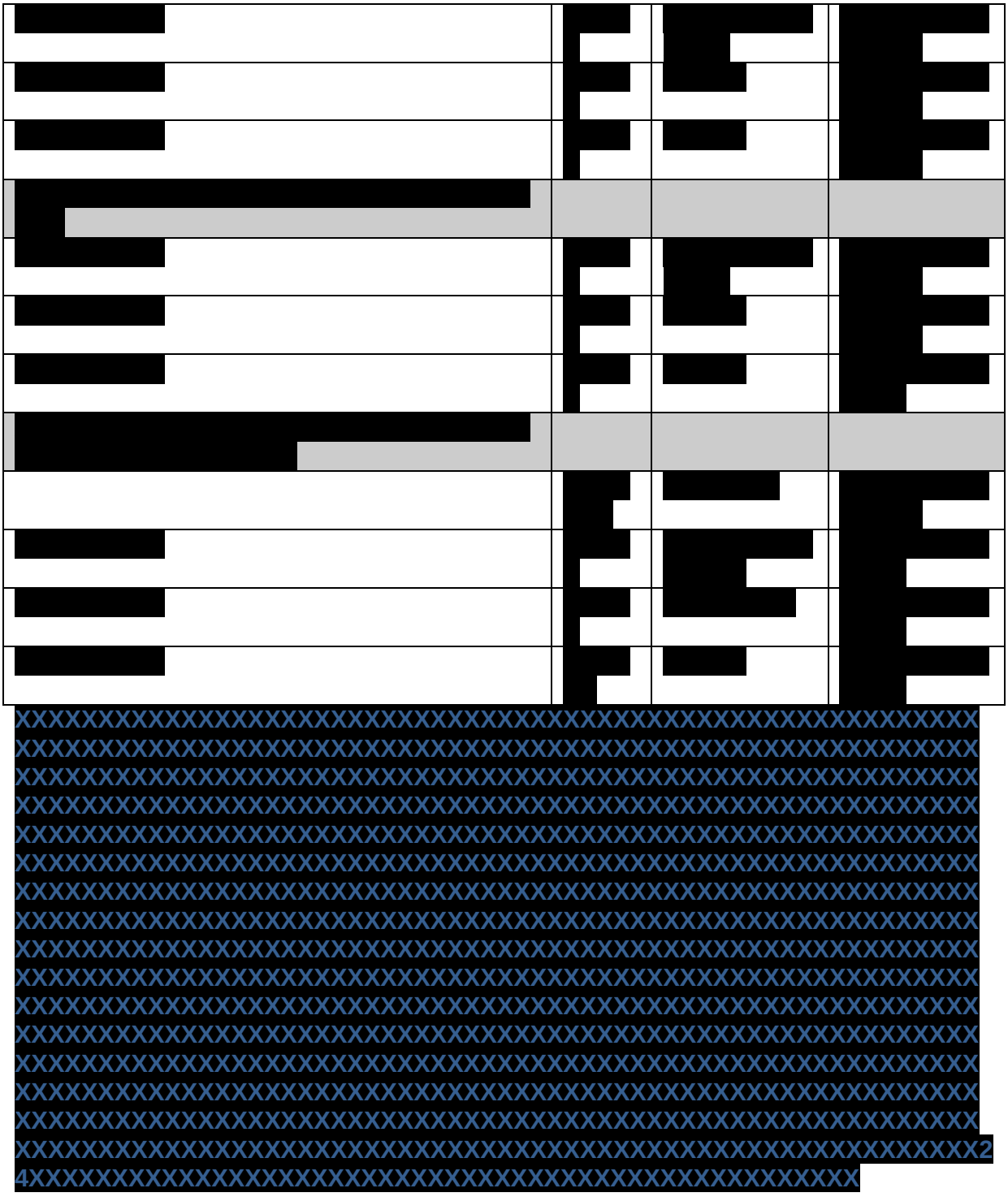
Patients with disease duration of more than 5 years accounted for 52% (13/25) of patients. Of the patients in clinical remission, 76% (32/42) had disease duration less than 5 years (p<0.05). This implies that early treatment with tocilizumab often results in clinical remission.

Unpublished/internal data:

As per the original submission, CSRs were supplied for MRA317JP and MRA324JP. The following is a summary of these results.

[redacted]

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



[redacted]

How CHAQ responses were elicited from children aged under 5 years in tocilizumab trials

The parents of such patients completed the CHAQ results. At the time of randomization 21/112 (19%) were less than 5 years of age.

References for clinical responses:

Inaba et al. Radiographic improvement of damaged large joints in children with systemic juvenile idiopathic arthritis following tocilizumab treatment. Ann Rheum Dis 2011 70: 1693-1695

Inaba et al. Radiologic Improvement Of Damaged Large Joints In Children With Systemic Juvenile Idiopathic Arthritis Following Treatment With Tocilizumab, Anti-Il 6 Receptor Monoclonal Antibody. Ann Rheum Dis 2007;66(Suppl II):550

Inaba et al. Radiologic Evaluation Of Large Joints During Tocilizumab Treatment In Children With Systemic Juvenile Idiopathic Arthritis Open label - single arm Radiographic endpoints Ann Rheum Dis 2009;68(Suppl3):720

Nakajima et al. Improvement of reduced serum cartilage oligomeric matrix protein levels in systemic juvenile idiopathic arthritis patients treated with the anti-interleukin-6 receptor monoclonal antibody

tocilizumab. Mod Rheumatol (2009) 19:42–46

Kaneko et al. Discrepancy Between Progression Of Joint Damage And Improvement Of Systemic Inflammation In Patients With Systemic-Onset Juvenile Idiopathic Arthritis Treated With Tocilizumab. Ann Rheum Dis 2009;68(Suppl3):719

Yokota et al. (a) Therapeutic Efficacy of Humanized Recombinant Anti–Interleukin-6 Receptor Antibody in Children With Systemic-Onset Juvenile Idiopathic Arthritis. Arthritis Rheum 2005;52(3):818-825

Yokota et al. (b) Long-term treatment of systemic onset juvenile idiopathic arthritis (SO-JIA) with humanized anti IL-6 receptor monoclonal antibody, tocilizumab (Actmera). Arthritis Rheum 2005;52(9 Supp 1):S725

Yokota et al. Long-term Safety And Efficacy Of Tocilizumab In Patients With Systemic Juvenile Idiopathic Arthritis (JIA) Under The Extension And Long-term Trials LTE. Arthritis Rheum 2008; 58(9 Suppl. 1):S631-S632

Yokota et al. (a) Safety and efficacy of up to three years of continuous tocilizumab therapy in children with systemic-onset juvenile idiopathic arthritis LTE. Ann Rheum Dis 2009;68(Suppl3):715

Yokota et al. (b) Drug-Free Remission of Patients with Systemic Juvenile Idiopathic Arthritis Receiving Tocilizumab for Treatment LTE drug withdrawal. Arthritis Rheum 2009

Yokota et al. Clinical Remission In Children With Systemic Juvenile Idiopathic Arthritis Receiving Tocilizumab Treatment – Analysis From Phase II And Phase III Extension Trials LTE. Ann Rheum Dis 2010;69(Suppl3):627

Part C. Response to consultation questions

1. **Has all of the relevant evidence been taken into account?**

We feel that the clinical evidence has not been interpreted fully in light of tocilizumab's license. Please see our comments to specific points on the ACD below.

2. **Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?**

Generally the summary of clinical evidence was sound. However, we have very specific comments regarding the interpretation by the ERG and Appraisal Committee of our economic modeling which we have set below specific comments from the ACD below.

3. **Are the provisional recommendations sound and a suitable basis for guidance to the NHS?**

No, we feel that the provisional recommendations may change upon receipt of our revised base case [redacted]

4. **Are there any aspects of the recommendations that need particular consideration to ensure we avoid unlawful discrimination against any group of people on the grounds of gender, race, disability, age, sexual orientation, religion or belief?**

We have not identified any issues which have the potential to lead to unlawful discrimination, but without consulting legal specialists this is a difficult question to

answer. We can only emphasise that sJIA does affect a particular age group. Therefore any recommendation which impinges directly on this age group does require particular attention. We are in no doubt that the Institute fully understands these issues and the need for careful consideration, given its extensive experience with appraisals of medicines which are primarily given to elderly patients.

5. **Are there any equality -related issues that need special consideration and are not covered in the appraisal consultation document?**

We have not identified any such issues.

3.24 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 patient's condition (change in CHAQ score based on ACR response) instead of the absolute condition of the patient. The ERG further noted that the change in a patient's 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 patient's 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.

No justification is given as to why the model health states are not mutually exclusive. In page 55 of the ERG report it is stated that the model assumes a homogenous cohort; which contradicts the above criticism.

We feel that our choice of ACR as a model parameter was justified given current clinical practice; ACR is considered by clinicians to be the most meaningful clinical outcome. In fact, we encountered resistance to the notion of establishing CHAQ thresholds, as no clinician would be able to use these thresholds to truly assess a patient's level of illness, much less decide whether to switch or continue treatment. ACR response may explicitly refer to a proportionate improvement level, but it in clinical practice its interpretation appears to be broader, incorporating a certain set of expectations about a patient's likely state once ACR30,50,70 or 90 have been achieved. That is why our original model strove to use ACR response level as an indicator for expected CHAQ score change, rather than as an explicitly linked 'health state' in its own right.

We also would like to point out that our model mechanism was based on that used in adult RA to allocate patients to health states. The only difference is that in adult RA a mechanism is built to calculate different CHAQ values over time (improvement/deterioration) based on treatment line. Indeed our original model's health state allocation was identical (NOT just 'similar' as has been assumed) to that used in adult RA.

3.25 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 (move to the next treatment line) or die. The ERG thought that, given the nature of the disease, this assumption was unlikely to be correct.

Roche has evidence that the proportion of "high" responders (ACR 70-90) increases over time following the first 12 weeks. However, this evidence is available only for tocilizumab. Therefore, we have adapted our model to allow some movement over time (albeit based on an arbitrary assumption), but we have taken a conservative approach in not attempting to model this sustained response among "high" responders which was observed in TENDER trial patients after the main experimental phase of the study.

3.30 The ERG questioned the cost estimates for health states were defined by expert opinion because they present a cost for non-responders (£3300) that is more than six times higher than the cost for an ACR30 response (£500), whereas a ACR90 response is associated with only a 30% decrease in cost (to £350) compared with an ACR30 response.

The key driver of the quoted difference in cost is hospitalization rate. Patients that do not respond to treatment are predominantly treated in the hospital. Given the severe symptoms associated with uncontrolled, non-responding disease (which may include fever and skin rash), and also considering that all patients are of young age, it seems plausible to suggest that extensive hospitalisation would occur while patients experience a disease flare. It is also expected that non-responders would experience a number of disease flares in a given year. Clinical experts suggested that for non-responders the length of hospitalisation could far exceed three weeks and that the average patient could stay in hospital for as much as three months in a year. Our model took a conservative approach to this estimate and considered the lowest value suggested by clinical experts (equating to 3-4 weeks a year) rather than the highest.

3.33 The ERG conducted some exploratory analyses using the following assumptions and modifications:

- **The starting age is 7 years (based on the observed average of 6 years, plus 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 are adjusted to reflect the methotrexate non-responder population (95% of the whole population).**
- **The relative risk for 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 probabilistic sensitivity analysis are adjusted to include additional uncertainty around the relative risks and around the adjustment factor.

We noted these experimental changes by the ERG. We also noted that apart from the change to model starting age, all other modifications essayed had a negligible impact on the model results. We would like to point out that in our original submission, we considered both linear and exponential models for withdrawal risk and did provide a range of possible distribution types and widths for probabilistic sensitivity analysis.

3.35 In the second scenario, the ERG explored the effect of the manufacturer's assumption that after the initial response patients stay in their current health state, withdraw and move to next line, or die. The ERG assumed instead 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 optimistic.

There is no evidence that a patient's condition improves (by 10% or other probability) across all treatments. There is some evidence that this occurs under tocilizumab, however if this improvement was assumed to apply only to tocilizumab only the analysis would "unfairly" favour treatment sequences involving Roche's product. We also noted that the ERG assumed both improvement and deterioration would occur between ACR categories. It is unclear to us whether it was the overall improvement or overall deterioration caused by these adjustments which impacted the results. Our diagnostic analysis of the ACR responses predicted when 'deterioration' or 'improvement' is introduced to the model suggest that this approach will generate unpredictable and difficult-to-interpret model results.

4.4 The Committee heard that there is variation in the use of tocilizumab in the UK, but that in general tocilizumab is currently used for patients whose condition does not respond to methotrexate, and following either infliximab or anakinra.

We are surprised to read this summary of the discussions at the Appraisal Committee meeting. Our notes suggest that clinicians named etanercept as a commonly-used first line option, with tocilizumab sometimes given after it in case of inadequate response. We noted that infliximab and anakinra were both also mentioned as treatment options, with anakinra considered less efficacious. We agree that the discussions on the day reflected the great variation in clinical use of tocilizumab in sJIA, but cannot agree with the statement that tocilizumab is 'in general' used in patients who have already received anakinra or infliximab. We would urge the Committee to seek further clarification ahead of the next meeting. We would also point out that the infliximab Summary of Product Characteristics (SPC) does not include an indication for sJIA, and specifically states that evidence for use in children is not available. A similar SPC is found for anakinra; this product is not recommended for use in children or adolescents under 18 years of age

with sJIA.

4.6 The Committee considered the evidence presented for the two populations defined in the scope and the different views of the population definitions from the manufacturer and the ERG. For the population of patients whose systemic JIA had failed to respond to NSAIDs and systemic corticosteroids, the Committee noted that only 5% of the TENDER trial population were methotrexate naive. The Committee also noted that the manufacturer had used a post-hoc analysis to compare patients receiving tocilizumab with those patients in the placebo group receiving methotrexate and that this was not methodologically acceptable. The Committee therefore concluded that there was no evidence to allow them to further consider the clinical or cost effectiveness of tocilizumab compared with methotrexate.

The TENDER study addresses Populations 1 and 2, as stated on page 39 of the manufacturer submission (MS):

"As such, by viewing only the inclusion criteria, the TENDER population matches population 1: children and young people 2 years and older with systemic JIA which has not responded adequately to prior NSAID(s) and systemic corticosteroids.

"However, on closer analysis of patients' treatment histories on joining TENDER, the study most accurately reflects population 2: children and young people 2 years and older with systemic JIA which has not responded adequately to prior NSAID(s), systemic corticosteroids and methotrexate.". This was also clarified in Roche's Response to Clarification Questions (A2, p2).

4.10 The Committee considered the economic model submitted by the manufacturer for the cost-effectiveness analysis. The Committee accepted the ERG's comments that the manufacturer's economic model did not adhere to the conventions of a Markov model. The Committee noted that the model structure meant that the health states were not homogeneous, and did not allow transition from one ACR response to another in the same line of treatment. The Committee considered that after each treatment cycle, patients could move from one health state to another, which was represented by a change in CHAQ score. An improvement in the health state did not necessarily translate to a specific improvement in ACR response. The Committee disagreed with the manufacturer's assumption of an absolute change in CHAQ score but a relative change in ACR response. The Committee heard from the clinical specialists that clinicians used the ACR response to quantify the improvement in health outcome of patients and that there would be continuous improvement in the ACR response of patients with systemic JIA who were on tocilizumab beyond 12 weeks. The Committee concluded that the manufacturer's economic model did not accurately represent the natural history of systemic JIA and its response to treatment.

Please see comments on 3.24.

4.12 The Committee considered the costs for tocilizumab used in economic model. The Committee noted that the costs of treatment were a composite of cost

of medication and cost of administering the medication. The Committee understood that, in practice, clinicians decreased the frequency of administration of tocilizumab, and in some instances stopping treatment altogether after 18 months. The Committee heard from the clinical specialists that the costs for the health states in the model were a reasonable reflection of clinical practice in the UK. The Committee also noted that potential cost savings could result from reductions in orthopaedic surgery for future joint damage and in bone marrow transplant and stem cell procedures. These factors had not been taken into account in the model.

We would like to emphasise the clinicans' consensus that the health state costs we had used were reasonable, as this conclusion addresses the ERG's concern at the disparity observed in costs between 'ACR non-responder' and 'responding' ACR categories.

Due to a lack of evidence, the cost savings from reductions in orthopaedic surgery for future joint damage and in bone marrow transplant and stem cell procedures are not taken into account in the model.

4.13 The Committee noted that the manufacturer's base-case ICER for the comparison of tocilizumab with anakinra was £23,200 per QALY gained. However, given the Committee's concerns around the model structure, it could not accept this as a reliable ICER. The Committee noted that the primary endpoint of the TENDER trial (ACR30 response and no fever) had not been the outcome used in the base-case model, but was included in the sensitivity analysis. Given that both the TENDER and ANAJIS trials included the outcome ACR30 response and no fever, the Committee concluded that this should have been considered in the base case.

In our original model, we preferentially used the ACR-Nofever outcome for comparison with other biologics, in cases where only this outcome was available from the clinical trials.

4.14 The Committee considered the sequencing of the comparators in the model. The Committee noted that even though the manufacturer had used infliximab in the indirect comparison analyses, it had not been used in the sequencing. The Committee heard from the clinical specialists that infliximab is often used because of the option to adjust the dose and the potential for an improvement in the condition of patients with active systemic JIA who receive higher doses.

We are somewhat surprised to read that the Committee considers that infliximab is 'often' used, if this consideration is based only on the discussions held at the Appraisal Committee Meeting. Our notes suggest that whilst infliximab was named as a treatment option with benefits as stated in 4.14, it was not by any means regarded as a standard of care.

Infliximab was not included our original model's choices because it is not recommended for the treatment of children or adolescents with JIA due to insufficient evidence

[Infliximab Summary of Product Characteristics accessed 20/03/2011].

Only one TNF study could be used in the comparison with TENDER for ACR response. This was NCT00036374 (infliximab study). Our original model used the response rates from NCT00036374 to inform a "class effect" estimate for all anti-TNF medicines rather than for infliximab only. The single infliximab trial was selected from the available evidence because it was the only trial with a comparable trial design to TENDER.

From the infliximab SPC:
The safety and efficacy of Remicade in children and adolescents younger than 18 years in the indications juvenile idiopathic arthritis, psoriatic arthritis and ankylosing spondylitis have not been established. No data are available.

4.15 The Committee considered the starting age of 2 years in the manufacturer's economic model. The Committee noted from the exploratory analyses conducted by the ERG that as the starting age of the child increases, so does the ICER. The Committee discussed the starting age of treatment with tocilizumab and understood that the peak age of onset of systemic JIA is about 2 years, with the mean age being around 5 years.

This concern has been noted and we have responded by providing a revised economic model.

4.16 In summary, because the Committee did not have any clinical evidence on the comparison of tocilizumab with methotrexate, it concluded that tocilizumab is not recommended for the treatment of systemic JIA in children and young people aged 2 years and older whose condition has responded inadequately to NSAIDs and systemic corticosteroids. Because the Committee did not have appropriate cost effectiveness data it was minded not to recommend tocilizumab for the treatment of systemic JIA in children and young people aged 2 years and older whose condition has responded inadequately to NSAIDs, systemic corticosteroids and methotrexate. The Committee recommended that the manufacturer be asked to submit a revised cost-effectiveness analysis for patients whose systemic JIA has not responded to methotrexate. This should include a revised economic model that allows transitions between ACR response categories within a treatment line and either restructures the health states defined by CHAQ scores and uses treatment response to define transition probabilities or alternatively the model should be a patient-level simulation to define the cost and utility values of a health state depending on both the starting CHAQ score and the change in CHAQ score in relation to the ACR response. The manufacturer's revised base case should include the assumption that treatment starts at age 5 years. For the comparison of tocilizumab and anakinra a revised base case using the primary outcome (ACR30 response and no fever). The Committee requested a fully incremental cost-effectiveness analysis of tocilizumab compared with anakinra and infliximab. Sensitivity analysis should be carried out on the revised base cases, including: the uncertainty around the adjustment factor derived from the etanercept study used to take account of the other types of JIA subgroups in the

infliximab study; a 'stopping rule' for tocilizumab after 2 years of treatment; and a decreased frequency of administration of tocilizumab after 6 months to a 4-weekly regimen. Scenario analyses should be carried out that include a fully incremental cost-effectiveness analysis of tocilizumab compared with anakinra and infliximab with: infliximab as first-line treatment followed by tocilizumab; and infliximab as first-line treatment followed by anakinra. The Committee requested further information on: radiographic evidence of progression of joint damage for patients receiving tocilizumab; long-term follow-up data from the trials being conducted in Japan; and how CHAQ responses were elicited from children aged under 5 years.

We feel that the Appraisal Committee has made an unfair decision regarding the population of patients who have responded inadequately to DMARDs and systemic corticosteroids. The TENDER trial's comparison with methotrexate reflects clinical reality, in which methotrexate use is endemic and the clinical question of interest is whether methotrexate can be dispensed with in future and tocilizumab given without the need for methotrexate. The European regulatory authority accepted this approach and considered the safety and efficacy data from TENDER sufficient for the granting of a license in this spirit. By contrast, the ERG's emphasis on obtaining a comparison of tocilizumab with methotrexate in 'methotrexate-naive' patients reflects a misunderstanding of clinical practice. The decision about treating with methotrexate does not represent an important clinical juncture at which a doctor will choose to either prescribe a biologic or methotrexate, bearing in mind cost-effectiveness at this time. We believe that the ERG and Appraisal Committee have been too quick to reject evidence in this setting, potentially driving a needless and clinically unrealistic limitation to patient access. This potential limitation is underscored by the apparently strong clinical and cost-effectiveness case for tocilizumab compared to methotrexate alone.

END OF DOCUMENT