Golimumab for the treatment of rheumatoid arthritis after failure of previous disease-modifying antirheumatic drugs: A Single Technology Appraisal

Report by the ERG for the second committee meeting

November 2010

Executive summary

The ERG have reviewed the manufacturer's submission to assess the quality of the evidence provided in response to the Committee's request for further information on the clinical and cost-effectiveness of golimumab, in section 1.4 of the ACD. The Committee requested further information on the following five points and the ERG's assessment of the evidence submitted is briefly summarised below each point.

1. Incorporation of ACR70 data in the economic model
    - An updated model was submitted and results were presented for the DMARD experienced population
    - The ERG found that the submitted model is not internally consistent and therefore the ICERs provided cannot be considered to be valid.
    - The results presented by the manufacturer do not use the outputs of the mixed treatment comparison and rely instead on the meta-analysis results for TNF-α inhibitor vs. placebo
    - Not all the efficacy evidence applied in the model matches that described in previous manufacturer submissions, so it is unclear how some of the estimates have been generated
    - No results were presented for the TNF-experienced population

2. Provision of SF-36 data from the GO-FORWARD and GO-AFTER trials and a sensitivity analysis in which these data are included in the economic model using SF-6D and/or mapping approaches to EQ-5D
    - The SF-36 data from the GO-FORWARD trial have been provided and these show significant improvements ($p<0.05$) in six out of eight domains
    - SF-6D utility estimates have been derived from the SF-36 data and these show differences in the mean utility between treatment arms and between groups with differing ACR responses
    - An updated model was submitted and results were presented for the DMARD experienced population
    - The model incorporating ACR70 was used as the starting point for this analysis and therefore all points made above regarding the validity of that model apply equally to these results
    - The methods used to incorporate the SF-6D utility values are not clearly reported, and appear to be inappropriate
    - The utility values used in the manufacturer's model do not match those presented in their supporting document

- No results were presented for the TNF-experienced population, although SF-36 data were not available from the GO-AFTER trial, therefore any estimates in this population would have relied on indirect evidence from the GO-FORWARD trial

3. Data including the proportion of people who will receive 100 mg golimumab (that is, people who weigh more than 100 kg and whose disease has not responded after three or four doses) and inclusion of this proportion in the economic model
    - No additional data or economic analysis was presented as a Patient Access Scheme (PAS) scheme has been proposed to the Department of Health which would, if accepted, result in the same costs being accrued regardless of whether a 50mg or 100mg dose is needed

4. A sensitivity analysis in which disease progression on palliative treatment is reflected as an increase in HAQ score of 0.06 per year
    - For the DMARD experienced population, sensitivity analysis results were provided for both the revised model incorporating ACR70 and for an earlier version of the model not incorporating ACR70.
    - The results using the model incorporating ACR70 are not valid due to the concerns raised above regarding the internal consistency of this model
    - The method used to incorporate the Kay study into the earlier model was not considered to be valid by the ERG as it failed to maintain study randomisation
    - For the TNF-experienced population, the results provided in Table 8 appear to come from an analysis using a palliative care progression rate of 0.06, although the rate applied isn't stated by the manufacturer in their submission.

5. Cost-effectiveness results for the population in 1.3 [of the ACD] for golimumab compared with adalimumab, etanercept, infliximab, abatacept and tocilizumab.
    - Results were presented for golimumab compared to tocilizumab but no comparison was provided against adalimumab, etanercept, infliximab or abatacept
    - No details were given of the costs or efficacy evidence applied for tocilizumab and no model was provided to support the results.
    - The ERG was therefore not able to assess the validity of the cost-effectiveness estimates reported for tocilizumab

In summary, the analyses presented by the manufacturer are considered by the ERG not to have adequately addressed the Committees' requests. Reliable ICERs are not available from the model incorporating the ACR70 health state and insufficient information is provided on the methods used for the comparison against tocilizumab and for the analysis in which utilities were derived from the SF-36 data.

Following concerns raised by several commentators, the ERG also conducted a sensitivity analysis to examine whether excluding monotherapy studies or the TEMPO study from the mixed treatment comparison had any impact on the relative cost-effectiveness of the anti-TNF drugs. Excluding monotherapy studies did not substantially alter the basecase results with certolizumab remaining the most cost-effective intervention. Excluding the TEMPO study improved the cost-effectiveness of etanercept such that it was no longer dominated by certolizumab.

ERG critique of the manufacturer's post-ACD submission

In response to the Committee's request for further information under point 1.4 of the ACD the manufacturer has submitted further evidence addressing each of the five requests made by the committee. Their submission included two revised economic models and SF-36 data from the GO-FORWARD trial. The ERG's critique of the evidence submitted is summarised below under each of the requests made by the committee. For several of the analyses presented, the manufacturer has not provided sufficient information regarding the methods used in the analysis and supporting models have not been provided. The ERG has attempted to establish the different model assumptions used to populate the results tables by trying to match the results against those generated by models previously submitted by the manufacturer under various assumptions. Our findings are summarised in Table 1 which provides a brief overview of the different assumptions that the ERG believe have been used to generate the results provided in the submission.

Table 1: Summary of different scenarios presented in results tables within the submission

Table	Population	ACR health states	Kay study included	MTC[a] or meta-analysis	Reference costs	Utility	HAQ in palliative care	Baseline HAQ
1	MTX experienced	20/50	No	MTC	Original	HAQ to EQ-5D	0.09	1.41
2*	MTX experienced	20/50/70	Yes	Meta-analysis	Updated	HAQ to EQ-5D	0.09	1.37
3	MTX experienced	20/50/70	Yes	Meta-analysis	Updated	HAQ to EQ-5D	0.06	1.37
4	MTX experienced	20/50/70	Yes	Meta-analysis	Updated	HAQ to SF-6D	NA	NA
5	MTX experienced	20/50/70	Yes	Meta-analysis	Updated	HAQ to SF-6D	NA	NA
6	MTX experienced	20/50/70	Yes	Meta-analysis	Updated	HAQ to SF-6D	0.06	NA
7	MTX experienced	20/50	Yes	MTC	Updated	HAQ to EQ-5D	0.06	1.41
8	Anti-TNF experienced	20/50	NA	MTC	Updated	HAQ to EQ-5D	0.06	1.41

*NB the adalimumab and golimumab rows in Table 2 of the manufacturer submission are labelled incorrectly. The row labelled golimumab shows the results for adalimumab and vice versa.

[a]MTC = mixed treatment comparison, MTX = methotrexate

1. Incorporation of ACR70 data in the economic model

The manufacturer has provided a revised economic model which incorporates an ACR70 health state alongside the previously incorporated ACR20 and ACR50 health states. Cost-effectiveness results for the DMARD experienced population are presented in Tables 2 and 3 of the submission. No results are presented for the anti-TNF experienced population. The ERG's validation of the submitted model has focused on the Markov sheets which track the progress of patients between health states as these have been substantially revised to incorporate the new ACR70 state. The ERG has also examined the efficacy evidence applied in the analysis as it was necessary to apply additional evidence in the model to calculate the probability of achieving an ACR70 response. The revised model has a similar structure to the previous model, although some changes have been made to the methods used in previous analyses submitted by the manufacturer. Table 1 of the submission shows the results from the original manufacturer basecase analysis and the differences between this analysis and the revised analyses have been summarised in Table 1 above.

1.1 Efficacy evidence applied in the ACR70 model

The results in Table 3 of the manufacturer submission match the submitted model (Golimumab model_ACR70Final_JF_4NOV2010 2010-11-11 STC HAQ 3.xlsm) when the "meta-analysis" option rather than then "indirect comparison" option is selected. This means that the individual meta-analysis for each intervention compared with placebo is being applied rather than the data from the mixed treatment comparison. This is inconsistent with the approach taken in the previous submissions, and that used to generate the results in Table 1, in which the mixed treatment comparison (using a random effects model) was used to populate the model. Given the availability of a mixed treatment comparison, it is unclear why this has not been used as it is more methodologically rigorous than applying results from several meta-analyses conducted independently for each intervention.

In addition, the meta-analysis data used to generate the ICERs in Table 3 does not appear to have been consistently selected. For ACR20 and ACR70, the relative risks are taken from the fixed effects meta-analysis, but for ACR50, the data doesn't match either the fixed or random effects data presented in the submission. For some interventions (golimumab, adalimumab, infliximab) the 95% confidence interval matches those presented for the fixed effects meta-analysis (tables 43, 36, 45 respectively of the original manufacturer submission), but even in these cases, the risk ratio used to estimate the mean of the log-normal distribution doesn't match the data presented in the submission. These problems with the ACR 50 efficacy evidence were present in previous versions of the model, but weren't identified previously by the ERG as results were not presented using this data. In addition, the ACR70 data for etanercept applied in the model does not match that provided previously by the manufacturer (as summarised on Table 9 of the ERG report).

The model does contain mixed treatment comparison inputs for the ACR70 state, which are used when the "indirect comparison" option is selected in the model. However, the data for the ACR70 state are identical to those for the ACR50 state and they do not match the evidence previously submitted by the manufacturer (Table 14 of manufacturer's response to the first clarification

request). In addition, the formulae used to calculate the transition probabilities from the mixed treatment comparison relative risks are incorrect as patients achieving ACR50 are subtracted not once, but twice from the patients achieving ACR 20, when calculating the proportion achieving an ACR between 20 and 50. Furthermore, the indirect comparison uses the absolute numbers reported in the trials to estimate the transition probabilities for both the golimumab and placebo arms, but it does this by summing together the event rates from the GO-FORWARD and Kay studies, which effectively breaks study randomisation. The correct approach would have been to estimate a relative risk for golimumab compared to methotrexate by meta-analysis of the two studies. This could then be applied within the model. In summary, even if the appropriate relative risks from the mixed treatment comparison were entered for the ACR70 outcome, the model does not appear to be set up to correctly estimate ICERs based on the mixed treatment comparison data.

The submitted model which incorporates ACR70 does not appear to have been set-up to work for the TNF-experienced population. The model only uses data from the indirect comparison for the TNF-experienced population. The indirect comparison data used in the model for ACR70 does not match the evidence submitted on ACR70 for rituximab and golimumab and is identical to the data for ACR50. Therefore no estimate is available on the cost-effectiveness of golimumab in the TNF-experienced population when incorporating the ACR70 health state.

1.2 Errors in the implementation of the Markov model

The submitted model is not internally consistent in that it does not use identical methods to evaluate each of the drugs. The model contains a Markov sheet for each drug which shows the progress of patients through the health states over time and the costs and QALYs accrued from time spent in the various health states. However, there are differences between the Markov sheets used for the different drugs which are not explained and which appear to be errors. For example, in the Markov sheet for etanercept, there are blank cells under the ACR70 utility column for patients receiving the 6^{th} treatment which are not blank in any of the Markov sheets for the other drugs. Also in the etanercept Markov sheet, under the section which traces the costs accrued during time spent in the various health states, there is no ACR70 health state for patients receiving the 2^{nd} treatment in the sequence, but this health state appears in the Markov sheet for the other drugs. In the section of the Markov sheet used to calculate the drop outs from each treatment and the HAQ decrements after switching treatment, inconsistencies were observed in five out of the six Markov sheets. The effect of these differences is that the Markov sheets produce difference costs and QALY estimates even when the data inputs are altered so that each sheet is effectively modelling the same drug and should produce equivalent costs and QALYs. The ICERs in Table 2 and 3 cannot therefore be relied upon as the errors identified will affect the estimates of incremental costs and QALYs between the interventions.

The ERG were not able to validate every aspect of the submitted model within the timeframe available, however, it was noted that the manufacturer has failed to correct the errors identified previously by the ERG, and corrected in their exploratory analysis, concerning the infliximab costs applied in the death state and the modelling of HAQ decrements for certolizumab.

1.3 Reporting errors and uncertainties

The results in Table 2 of the submission appear to be mislabelled. We can see by comparing Table 2 and 3 that changing the HAQ progression rate for palliative care affects only the QALY estimates and not the cost estimates, as would be expected given the structure of the model. However, it also appears that the costs for golimumab and adalimumab are swapped over between these two tables. Having examined the submitted model to establish which of the tables agrees with the data generated by the model, the ERG believe that the rows in Table 3 are correctly labelled and that the row labels for Table 2 need to be switched for golimumab and adalimumab as shown in Table 2 below. The ICER for etanercept vs. methotrexate has also been corrected in Table 2 below as the figure presented in the Table 2 of the submission did not follow from the costs and QALYs presented.

Table 2: Revised base case analysis incorporating ACR70 data (Table 2 of submission with row labels and etanercept ICER corrected by ERG)

	Total Costs (£)	Total QALYs	ICER (£) vs. baseline (methotrexate)
Methotrexate	38,175	5.261	-
Certolizumab	77,348	6.252	39,529
Infliximab	78,527	6.447	34,024
Golimumab	70,514	6.323	30,451
Adalimumab	74,201	6.554	27,862
Etanercept	83,472	6.900	26,637

The manufacturer has not clearly labeled the tables to describe exactly how the results have been generated. The results presented in Table 1 match those originally submitted by the manufacturer before they were asked to update the unit costs used in the model. Whilst the model submitted to support the revised ICERs provided in Table 2 and 3 appear to include the updated unit costs. The revised model also incorporates data from the Kay study which was not included in the manufacturer's basecase analysis. It also uses a HAQ scores at baseline in the methotrexate experienced population of 1.37, rather than the score of 1.41 previously applied, and no explanation is given for this change.

2. Provision of the SF-36 data and incorporation of this data in an economic sensitivity analysis

2.1 SF-36/SF-6D Data

The ERG welcomes the SF-36 data that has been provided by the manufacturer. They are correct that SF-36 data is only available from the GO-FORWARD study, and not from the GO-AFTER study. The ERG cross-checked this with the Clinical trials registry (www.clinicaltrials.gov) and are pleased that the appropriate Health Related Quality of Life data has been finally released. Table 3 below

summarises the SF-36 data for the Physical Component summary scores at baseline, week 14 and week 24.

Table 3: GO-FORWARD SF-36 Data (Physical component summary scores)

		Placebo + MTX	Golimumab 100mg + Placebo	Golimumab + MTX		
				50mg	100mg	Combined
Baseline	n	132	131	89	88	177
	Mean (SD)	31.63 (8.298)	30.93 (8.462)	30.45 (8.373)	29.93 (8.036)	30.19 (8.188)
Week 14 change from baseline	n	127	127	85	85	170
	Mean (SD)	2.39 (7.798)	4.72 (8.782)	8.02 (7.170)	7.41 (8.044)	7.71 (7.603)
	p-value	-	0.033	<0.001	<0.001	<0.001
Week 24 change from baseline	n	125	125	88	86	174
	Mean (SD)	2.54 (8.055)	4.74 (8.844)	8.28 (8.327)	7.01 (7.796)	7.65 (8.071)
	p-value	-	0.070	<0.001	<0.001	<0.001

The data shows that, for 50mg and 100mg GOL + MTX (golimumab plus methotrexate) there is a significant improvement in the SF-36 physical component summary score, at both week 14 and 24 compared to baseline. The mental health component summary is not provided. However the changes in individual domains are provided in the response by manufacturer and both 50mg and 100mg GOL + MTX arms see a significant ($p<0.05$) change across all physical health domains at both timepoints. Some mental and social functioning domains however do not see a significant improvement (Role-emotional week 14 and 24, social functioning for 50mg GOL + MTX at week 14 and 24, mental health for 50mg GOL + MTX at week 14).

On balance, the ERG feel that the manufacturer has provided conclusive evidence that GOL + MTX has a significant impact on the physical component of health related quality of life for patients who are DMARD-experienced, although the impact of GOL + MTX on the mental component of health related quality of life is less conclusive.

The manufacturer converts the SF-36 data into SF-6D utility values using a Bayesian non-parametric conversion algorithm (Kharroubi *et al.* 2007). This algorithm provides both a parametric and a posterior estimate of the mean utility at week 0 and week 24 (see Table 4 below).

Table 4: SF-6D Values

	Parametric mean utility			Posterior mean utility		
	Week 0	Week 24	Change	Week 0	Week 24	Change
50mg GOL +MTX	0.56	0.63	0.07	0.54	0.60	0.06
Placebo + MTX	0.56	0.57	0.01	0.54	0.55	0.01

The parametric approach is the random effects model used by Brazier et al (2002) in their original valuation of the SF-6D from SF-36. The posterior approach uses a non-parametric Bayesian hierarchical method that allows the model function to take any form. The results are similar, although the parametric method provides a slightly larger utility gain compared to the posterior method (0.07 vs 0.06). Kharroubi *et al.* argue that their posterior approach is more appropriate, and the manufacturer generates results in their document (Table 4 and Table 5) that use both the

parametric and posterior estimates. These results appear to be from an analysis using the trial evidence to directly compare golimumab to methotrexate, however no model or methods are provided.

Table 5: SF-6D Values for ACR20 and ACR50 responders

	Parametric mean utility			Posterior mean utility		
ACR50	Week 0	Week 24	Change	Week 0	Week 24	Change
Non-response	0.56	0.57	0.01	0.54	0.54	0.00
Response	0.58	0.70	0.12	0.56	0.67	0.10
ACR20	Week 0	Week 24	Change	Week 0	Week 24	Change
Non-response	0.55	0.55	0.00	0.53	0.53	0.00
Response	0.57	0.67	0.10	0.56	0.64	0.08

The data in Table 5 show that ACR response is a predictor of change in utility, and also the results show a difference in magnitude of change between the ACR20 and ACR50 responder groups.

2.2 SF-6D Data incorporated in manufacturer's model

The manufacturer attempts to apply the SF-6D data into their DMARD-experienced population model. This is subsequent to their attempt to add the ACR70 health state. Due to the errors made when incorporating the ACR70 health state, the ERG is unable to properly validate the incorporation of the SF-6D data.

The manufacturer provides a model incorporating the ACR70 state and maintaining the original HAQ to EQ-5D utility function, and they also provide a model incorporating the ACR70 state and including the SF-6D utility values.

The manufacturer uses SF-6D utility values for baseline and for week 24 non-response, ACR20 and ACR50 response. They claim that SF-36 data is not available for the ACR70 responders in the GO-FORWARD trial and so assume ACR70 utilities are equal to the ACR50 values. The values applied in the model and summarised in Table 6 below do not match any of the values provided in the manufacturer's documentation (see Table 5 above).

Table 6: Model values – taken from Excel file

		MTX experienced GOL values Utility Value (sd)	MTX experienced MTX values Utility Value (sd)
Baseline SF-6D score		0.574 (0.103)	0.560 (0.103)
Week 24	Non-response	0.591 (0.114)	0.518 (0.123)
	ACR20	0.637 (0.094)	0.593 (0.086)
	ACR50	0.732 (0.129)	0.733 (0.153)
	ACR70	0.732 (0.129)	0.658 (0.153)

The manufacturer has estimated the MTX values by applying "the ratio of HAQ scores", presumably from the golimumab HAQ scores. A justification for this method is not provided. No other details are given and the ERG cannot replicate the utility values generated. These values do not have face

validity, with patients not achieving an ACR20 response having a substantially lower utility at week 24 compared to baseline (0.518 – 0.560). Also the utility from an ACR70 response is lower (0.658) compared to the utility from an ACR50 response (0.733).

The SF-6D MTX values could have been estimated by converting the SF-36 values for the MTX + Placebo values using the SF-6D algorithm. These values are provided in the spreadsheet attached to the manufacturer's response.

The manufacturer has assigned normal distributions to the utility values. Normal distributions are generally not appropriate for utility values because the maximum value of a utility value is 1 but normal distributions can sample above this. For example, the golimumab ACR70 utility value will sample above 1 for approximately 2% of PSA runs.

As mentioned previously, the manufacturer provides results for golimumab compared to methotrexate using both the posterior and parametric SF-6D estimates (Tables 4 and 5 of the submission), however the ERG was unable to replicate these results using the model provided.

The manufacturer provides a set of results for all TNF-α inhibitors and methotrexate (Table 6 of the submission). These results are replicated below with a little more detail for the incremental analysis. However, they should not be considered valid due to the lack of internal consistency found in the ACR 70 model and the uncertainties in the methods used to derive the utility values for the methotrexate arm.

Table 7: ICERs derived from Sheffield algorithm SF-6D mapped estimates (replicated from Table 6 of the submission with additional information on incremental analysis)

Technology	Total Costs (£)	Total QALYs	Incremental Analysis – comparison made to next least effective non-dominated strategy ICER (£/QALY)	ICER (£/QALY) versus Methotrexate§
Methotrexate	36,327	7.227	-	-
Certolizumab	75,499	8.062	Dominated by Golimumab	35,115
Golimumab	68,666	8.086	28,904	28,904
Infliximab	76,678	8.207	Dominated by Adalimumab	35,533
Adalimumab	72,352	8.278	19,197	31,451
Etanercept	81,627	8.501	41,592	38,511

§ Indicates cost-effectiveness when all other biologics contraindicated

The manufacturer has not reviewed for SF-36/SF-6D values for all TNF-α inhibitors, and instead has applied the golimumab estimates for change in SF-6D score by ACR response. The results indicate that golimumab dominates certolizumab, but is not as effective as adalimumab and etanercept. It is

unclear why the results in Table 6 for golimumab and methotrexate differ from those in Tables 4 and 5 as there is a lack of detail given on the methods used.

3. Data on the use of a 100mg dose and its implications for the economic analysis

No additional analysis was presented. The manufacturer state that a patient access scheme (PAS) scheme has been proposed to the Department of Health which would, if accepted, result in the same costs being accrued regardless of whether a 50mg or 100mg dose is needed. The manufacturer also states that there is "there is no direct clinical data supporting dose escalation for patients weighing >100kg" and that they will not be marketing golimumab in a manner which advocates dose escalation. However, dose escalation was described in the Table 7 of the manufacturer's original submission and also appears within the SPC.

4. Palliative care HAQ progression rate sensitivity analysis

For the DMARD experienced population, sensitivity analysis results were provided for both the revised model incorporating ACR70 (Table 3) and for an earlier version of the model not incorporating ACR70 (Table 7). The results in Table 3 are not valid due to the concerns raised under section 1 above regarding the internal consistency of the model incorporating ACR70.

The description of the scenario presented in Table 7 is inadequate making it difficult to be sure exactly which scenario the results relate to. We were unable to replicate the results in Table 7 by updating the original manufacturer model to use a palliative care progression rate of 0.06. The costs, but not the QALYs, in Table 7 appear to agree with an analysis, submitted in response to the second clarification request, in which the study data from Kay was incorporated in the model. This analysis was not considered valid by the ERG as the event rates from the Kay study and the GO-FORWARD study were simply summed rather than being properly synthesized using meta-analysis. However, we found that we were able to replicate the results in Table 7 when using the manufacturer's analysis incorporating the Kay study as the starting point and by incorporating updated unit costs and by applying the revised HAQ progression rate for palliative care. The results in Table 7 are therefore not considered by the ERG to be valid due to the method used to synthesise evidence from the Kay and GO-FORWARD study. The ERG note that they should not be compared directly with those presented in Table 1, as the ERG believes that neither the efficacy evidence, nor the unit costs applied are equivalent. They can also not be compared with the results in Table 3 as this analysis uses efficacy evidence from the separate meta-analyses for each drug rather than the mixed treatment comparison.

For the TNF-experienced population, the results provided in Table 8 appear to come from an analysis using a palliative care progression rate of 0.06, although the rate applied isn't stated by the manufacturer in their submission. As there was no description given of the methods used to generate the results in Table 8 and no model was submitted in support of this analysis, the ERG have attempted to determine what methods have been applied by updating previous versions of the model and comparing the results against those presented. The costs, but not the QALY estimates, in Table 8 appear to match those in Table 55 of the manufacturer's response to the first clarification

request. This suggests that they have incorporated updated unit costs. The ERG were able to generate results matching those in Table 8 for rituximab, methotrexate and golimumab, by taking the manufacturer's model incorporating revised unit cost and applying a HAQ progression score of 0 for rituximab and 0.06 for palliative care. It should also be noted that this analysis used the original dosing regime for rituximab in which it is assumed to be re-administered every 6 months and the results are not based on a model which incorporates ACR70.

5. Comparison against adalimumab, etanercept, infliximab, abatacept and tocilizumab

Results are presented in Table 8 of the manufacturer's submission for golimumab, rituximab and tocilizumab, but no comparison has been provided against adalimumab, etanercept, infliximab or abatacept. No model was submitted in support of the analysis presented in Table 8. Basic information is lacking on the methods used to incorporate tocilizumab, such as the cost and efficacy data applied. Whilst the available tocilizumab studies were summarized in the original manufacturer submission, they were excluded from the efficacy meta-analysis section so we have no indication of the relative risks that may have been applied in the model. Tocilizumab was also excluded from the economic evaluation in the original submission, so no estimate is given of its administration costs. Without further information on the methods used to generate the results in Table 8, it is not possible to accept these estimates as valid.

Economic sensitivity analyses on study inclusion criteria for the MTC

Several commentators raised concerns regarding the selection of evidence to inform the mixed treatment comparison (MTC), particularly with respect to the inclusion of the TEMPO study and various studies which used anti-TNF drugs without concomitant methotrexate (monotherapy studies). The manufacturer's original submission provided sensitivity analyses excluding these studies from the MTC and showed that this did not alter the broad conclusions regarding golimumab's efficacy relative to other anti-TNF drugs. However, the estimates generated in these sensitivity analyses were not applied in the economic model to establish whether there was any impact on the relative cost-effectiveness. The ERG have therefore conducted sensitivity analyses to establish whether excluding these studies has any impact on the relative cost-effectiveness of the different anti-TNF drugs. MTC results were not available for a scenario in which both the monotherapy arms and the TEMPO study were excluded. Separate analyses were therefore conducted using the MTC results when only the TEMPO study (based on table 59 and 60 of the original submission) or only the monotherapy studies (based on tables 61 and 62 of the original submission) are excluded.

These analyses were done using the model adapted by the ERG to generate their preferred scenario, as described in section 6 of the ERG report, and results are presented using the mean outputs from the probabilistic model. The results of these sensitivity analyses are given in Tables 8 and 9 below with the ERG's basecase scenario provided in Table 7 for reference. The Tables show mean costs and QALYs from the probabilistic sensitivity analysis. It can be seen that etanercept becomes more cost-effective when the TEMPO study is excluded as it is no longer dominated by certolizumab. When the monotherapy studies are excluded the results do not differ substantially from the basecase with certolizumab remaining the most cost-effective intervention. Golimumab is dominated by adalimumab and extendedly dominated by certolizumab.

Table 7: ERG basecase results

Technology	Total Costs (£)	Total QALYs	Incremental Analysis – comparison made to next least effective non-dominated strategy ICER (£/QALY)	ICER (£/QALY) versus Methotrexate§
Methotrexate	39,701	4.622	-	-
Infliximab	67,528	5.792	Extendedly dominated by certolizumab	23,774
Golimumab	71,530	5.889	Extendedly dominated by certolizumab	25,123
Adalimumab	72,824	5.968	Extendedly dominated by certolizumab	24,604
Etanercept	80,096	6.307	Dominated by certolizumab	23,966
Certolizumab	79,185	6.518	20,828	20,828

§ Indicates cost-effectiveness when all other biologics contraindicated

Table 8: ERG basecase results when TEMPO study excluded

Technology	Total Costs (£)	Total QALYs	Incremental Analysis – comparison made to next least effective non-dominated strategy ICER (£/QALY)	ICER (£/QALY) versus Methotrexate§
Methotrexate	39,694	4.612		
Infliximab	67,069	5.766	Extendedly dominated by certolizumab	23,727
Adalimumab	71,074	5.883	Extendedly dominated by certolizumab	24,703
Golimumab	71,295	5.892	Extendedly dominated by certolizumab	24,695
Certolizumab	80,797	6.598	20,705	20,705
Etanercept	91,576	6.996	27,058	21,767

§ Indicates cost-effectiveness when all other biologics contraindicated

Table 9: ERG basecase results when monotherapy studies excluded

Technology	Total Costs (£)	Total QALYs	Incremental Analysis – comparison made to next least effective non-dominated strategy ICER (£/QALY)	ICER (£/QALY) versus Methotrexate§
Methotrexate	39,523	4.587	-	
Infliximab	67,458	5.735	Extendedly dominated by certolizumab	24,327
Golimumab	71,729	5.848	Extendedly dominated by certolizumab	25,529
Adalimumab	71,568	5.852	Extendedly dominated by certolizumab	25,336
Etanercept	76,222	6.055	Extendedly dominated by certolizumab	24,990
Certolizumab	77,318	6.334	21,634	21,634

§ Indicates cost-effectiveness when all other biologics contraindicated

References

Brazier JE, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. Journal of Health Economics 2002; 21:271–292

Kharroubi SA, Brazier JE, Roberts J, O´Hagan A. Modelling SF-6D health state preference data using a nonparametric Bayesian method Journal of Health Economics. Journal of Health Economics 2007; 26:597-612