Abbott's response to the consultation on the York CRD/CHE Technology Assessment Report: Etanercept, Infliximab and Adalimumab for the Treatment of Psoriatic Arthritis

Executive Summary

Abbott welcomes the opportunity to comment on the Technology Assessment Report (TAR) for the appraisal of adalimumab, etanercept, and infliximab, for the treatment of psoriatic arthritis.

Abbott notes the base case analysis from modelling conducted by the Assessment Group, which suggests that adalimumab, etanercept and infliximab are cost effective options when compared with palliative care assuming a threshold of £30,000 per QALY, and that only adalimumab and etanercept are cost-effective at a threshold of £20,000 per QALY.

Abbott is surprised that the report draws such strong inferences from the base case analysis, in particular that adalimumab is extendedly dominated by etanercept, without emphasising that this conclusion is very sensitive to a number of assumptions enumerated elsewhere in the report. The cost-effectiveness of adalimumab and etanercept versus palliative care are very similar in the base case and the report indicates that there is considerable uncertainty around many of the model inputs. In light of these uncertainties and the very similar cost-effectiveness results for adalimumab and etanercept, we would suggest that a strong conclusion of extended dominance of one treatment over the other is misplaced and would ask that this be drawn to the attention of the Appraisal Committee when they consider the report. Abbott also considers that the results of the probabilistic sensitivity analyses do not appropriately reflect the uncertainty that etanercept will be a more cost effective therapy option than adalimumab. This is partly because the mixed treatment comparison results generated by the Assessment Group give a much greater HAQ improvement for etanercept than adalimumab. Furthermore, Abbott considers that the low ICERs for infliximab versus palliative care are highly questionable as they are predicated on both a greater HAQ response for patients receiving infliximab and an average patient weight of 70kg. Neither of these assumptions seems to be easily supported.

In the model developed by the Assessment Group, the decision to allow the change in HAQ to depend on both PsARC response and the biologic treatment is a key model assumption with sensitivity analyses indicating that adalimumab is no longer dominated by etanercept when the same change in HAQ is applied for all PsARC responders, regardless of treatment. This sensitivity analysis also indicates that infliximab has a high ICER versus palliative care or versus the other anti-TNF agents when HAQ response is modelled as being the same for all PsARC responders.

There is a high degree of uncertainty in the Assessment Report conclusion that etanercept is the most cost-effective anti-TNF for the treatment of PsA. The effectiveness estimate is driven by two 12-week trials, one containing 30 patients receiving etanercept vs. another trial of 51 patients receiving adalimumab. This uncertainty is compounded by the possibility that the patient populations included in each of these two trials are not the same. Furthermore, data from the BSRBR mimicking routine clinical practice in the UK in a much greater number of patients, suggest that the three anti-TNFs are similarly effective in treating the arthritis component of the disease.

Abbott welcomes the inclusion of the benefits of the different treatments on the skin component of the disease in the modelling conducted for this appraisal. The Assessment Group state that "*the assessment of effectiveness in Section 5.2.2 did not find any appreciable differences in the biologics' response rates for joint disease or psoriasis between approximately 12 weeks compared with 24 weeks*." As a result of this, the Assessment Group used 12 week efficacy data to inform the clinical-effectiveness estimates in their model. Given the strong inference the Assessment Group make in their conclusions about the most cost-effective drug, it is worth noting

that there are appreciable differences in the PASI response rates between weeks 12 and 24 for adalimumab. Further, it is important to recognise that improvements in psoriasis with adalimumab, when 12 week data are used, have been underestimated in the Assessment Group model. Therefore, given that the ICERs for adalimumab and etanercept vs. palliative care are similar, the improved PASI data at week 24 for adalimumab could have an impact on the conclusion made by the Assessment Group that etanercept is the most cost effective treatment option.

1. Presentation of the results in the Technology Assessment Report

The Assessment Group conclude that adalimumab is extendedly dominated by etanercept in the base case analysis. However, Abbott feels that it is important to consider the cost-effectiveness of each of the drugs compared to palliative care. As can be seen in the Table 1.1 below, the base case analysis suggests that all of the three anti-TNF inhibitors are cost-effective treatment options when compared with palliative care assuming a threshold of £30,000 per QALY, and that both adalimumab and etanercept are cost-effective at a threshold of £20,000 per QALY.

Table 1.1: ICER vs. palliative care for all three anti-TNFs from the Assessment Group's model

Strategy	QALY	Cost	ICER vs. palliative care
Palliative care	5.241	42205	-
Adalimumab	6.642	66408	£17,275.52
Etanercept	7.115	72172	£15,990.93
Infliximab	7.43	89107	£21,426.22

It is also important to note that the cost-effectiveness of adalimumab and etanercept versus palliative care are very similar, with ICERs of £17,274 and £15,990 respectively. The report also indicates that there is considerable uncertainty around many of the model inputs. In light of these uncertainties and the very similar cost-effectiveness results for adalimumab and etanercept, it is surprising that the conclusion of the TAR makes such strong inferences regarding which treatment is the most cost-effective treatment option in patients with PsA.

2. HAQ change by PsARC responder/non-responder

The univariate sensitivity analysis conducted by the Assessment Group indicates that the differences across treatments in HAQ change by PsARC responder/non-responder is a key driver of the relative cost-effectiveness of the anti-TNF therapies. Since these values are not available in the public domain, these data were provided by each of the manufacturers on request from the Assessment Group. The Assessment Group then used these values in their mixed treatment comparison (MTC).

Due to the importance of this input, Abbott feels that it is crucial to ensure that the data used to inform the HAQ change are thoroughly examined. During the course of this review, three issues were identified which could have a significant bearing on the cost-effectiveness results: the MTC inputs, the baseline HAQ data and the assumptions around different HAQ changes for different treatments.

2.1 MTC Inputs

Table 10.5.2 in the Assessment Group report (reproduced below as Table 2.1.1) shows the MTC inputs for each of the drugs.

Table 2.1.1: HAQ change by PsARC response – MTC inputs (Table 10.5.2 of Assessment Group Report)

HAQ given PsARC response		standard error	HAQ given NO PsARC response		standard error
Placebo	-0.258	0.006	Placebo	-0.002	0.042
Etanercept	-0.635	0.062	Etanercept	-0.196	0.072
Placebo	-0.27	0.14	Placebo	0.02	0.05
Infliximab	-0.65	0.09	Infliximab	-0.2	0.09
Placebo	-0.16	0.096	Placebo	0.07	0.042
Infliximab	-0.58	0.057	Infliximab	-0.11	0.06
Placebo	-0.3	0.077	Placebo	0	0.037
Adalimumab	-0.5	0.041	Adalimumab	-0.1	0.053
Placebo	-0.2	0.0429	Placebo	0.1*	0.0429
Adalimumab	-0.4	0.056	Adalimumab	-0.1	0.056

*Note: this should be -0.1, however this appears to be a typo in the table only as the correct value appears to have been used in the MTC code.

As can be seen in Table 2.1.1, the HAQ changes for etanercept are provided to 3 decimal places, to 2 decimal places for infliximab, and to 1 decimal place for adalimumab. The results of the MTC were reproduced using the WinBUGS code provided in Appendix 10.5.6 of the Assessment Group report. These results are shown in Table 2.1.2, and are similar to those shown in the Assessment Group report.

Table 2.1.1: Results of MTC using Assessment Group inputs

	Patients who responded to treatment			Patients who did not respond to treatment		
	Mean	Credible intervals		Mean	Credible intervals	
		2.50%	97.5%		2.50%	97.50%
Placebo	-0.2371	-0.3140	-0.1557	0	0	0
Etanercept	-0.6249	-0.7998	-0.4577	-0.1876	-0.3756	-0.0024
Infliximab	-0.6331	-0.7738	-0.5024	-0.1686	-0.3110	-0.0313
Adalimumab	-0.4446	-0.5515	-0.3207	-0.0870	-0.2060	0.0328

In order to assess the impact of rounding to different decimal places, the MTC was re-run with the HAQ changes rounded to 2 decimal places for each drug. Additionally, the standard errors for adalimumab appear to have been incorrectly calculated. It seems that instead of dividing SD by the square root of the number of patients in each cell, SD was divided by the square root of the total number of patient in each treatment arm. This error was also corrected when re-running the MTC. The results of the re-run MTC are shown in Table 2.1.3.

Table 2.1.2: Results of MTC using inputs to 2 decimal places for all drugs

	Patients who responded to treatment			Patients who did not respond to treatment		
	Mean	Credible intervals		Mean	Credible intervals	
		2.50%	97.5%		2.50%	97.50%
Placebo	-0.2579	-0.3342	-0.1760	0	0	0
Etanercept	-0.6423	-0.8073	-0.4759	-0.2017	-0.3823	-0.0172
Infliximab	-0.6394	-0.7610	-0.5118	-0.1681	-0.3078	-0.0370
Adalimumab	-0.4952	-0.6089	-0.3832	-0.1443	-0.2628	-0.0131

A comparison of the results shows that the apparently inconsequential issue of rounding has a large impact on the mean HAQ improvements for both adalimumab responders and non-responders, with an increase of 0.0506, and 0.0573 respectively. Given that the differences in the

change in HAQ between treatments is a key driver of the results, this improvement in HAQ will result in an increase in QALYs for adalimumab thus changing the cost-effectiveness results.

Since the results of the MTC are so sensitive to the number of decimal places reported for the change in HAQ, Abbott has provided these data to 4 decimal places in Appendix 1. Since Abbott does not have access to the response rates from the etanercept and infliximab clinical trials to this level of accuracy, we were unable to determine the exact impact this change will have on the cost-effectiveness results. Abbott therefore suggests that the Assessment Group request the data to this level of detail from the other manufacturers and uses these data to re-run the MTC and the cost-effectiveness analysis.

2.2 Baseline HAQ

Since a patient's baseline outcome is likely to be strongly correlated with their change in outcome over the follow-up period, the concept of baseline adjustment with respect to the analysis of outcome measures is well understood in the clinical trial literature (e.g. Pocock *et al* 2002[1]) and the health economic literature (Manca *et al* 2005[2]). In the trials of PsA included in the MTC, it is clear there are differences in baseline HAQ (Table 2.2.1).

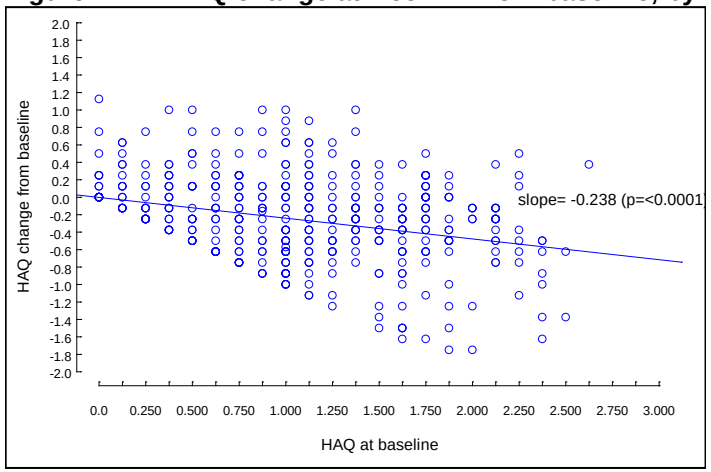
Table 2.2.1: Baseline HAQ for all trials

	Etanercept				Infliximab				Adalimumab			
	Mease 2000[79]		Mease 2004[53, 98, 100, 106, 108, 111]		IMPACT[80-82, 90, 97, 110, 112, 114-116, 118, 119]		IMPACT 2[83, 91, 92, 96, 99, 107, 113, 117]		ADEPT[52, 89, 93, 94, 101-105]		Genovese 2007[84]	
HAQ (0-3) Mean (SD)	1.3 (0.9, 1.6)*	1.2 (0.8, 1.6)*	1.1 (-)*	1.1 (-)*	1.2 (0.7)	1.2 (0.7)	1.1 (0.6)	1.1 (0.6)	1.0 (0.6)	1.0 (0.7)	0.9 (0.5)	1.0 (0.7)

Source: Assessment Group Report

Data from both the ADEPT and M02-570 trials indicate that there is a clear relationship between baseline HAQ and HAQ change (Figure 2.2.1 and Figure 2.2.1).

Figure 2.2.1: HAQ change at week 12 from baseline, by HAQ at baseline (ADEPT)



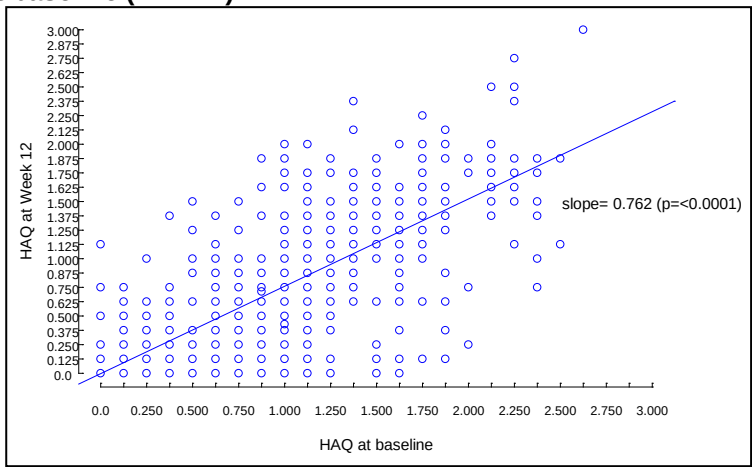
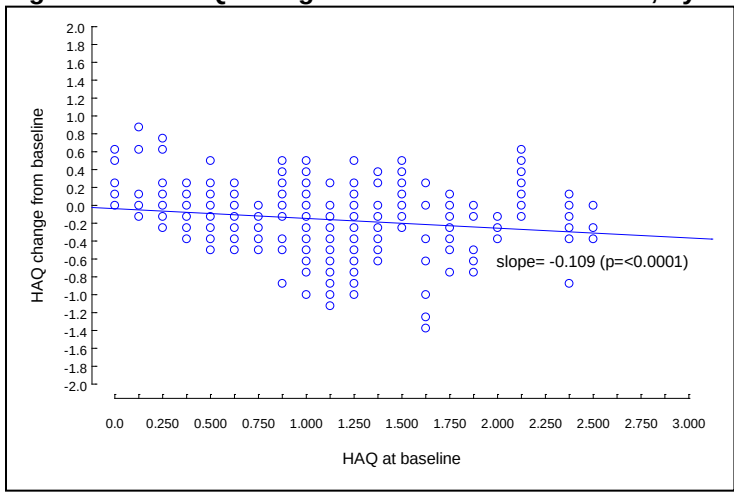
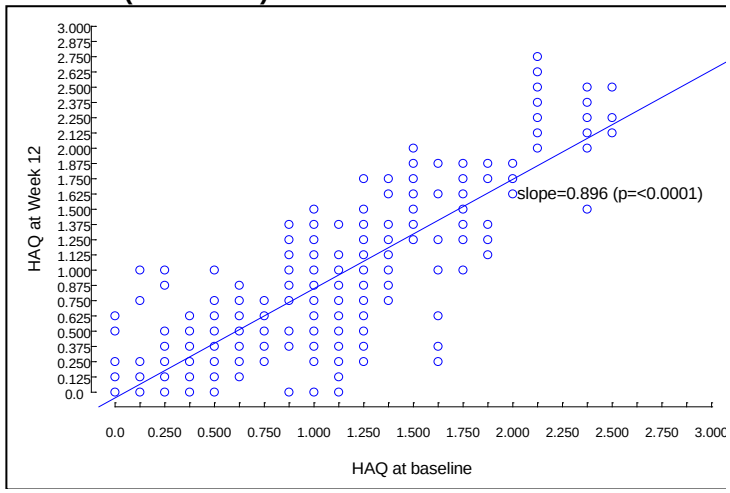


Figure 2.2.1: HAQ change at week 12 from baseline, by HAQ at baseline (Genovese)





Surprisingly the Assessment Group did not adjust for baseline HAQ when considering HAQ change in the MTC. A failure to adjust for baseline HAQ when considering HAQ change in the MTC will therefore bias the results of the analysis.

In order to assess the impact of baseline HAQ on the change in HAQ, the MTC code was adapted to include baseline HAQ as follows:

$$\mu_{PNRi} = baseline + \beta\,(HAQ.base)_{PNRij}$$
$$\mu_{PRi} = \mu_{PNRi} + \delta.diff_{PRi} + \beta\,(HAQ.base)_{PRij}$$
$$\mu_{TNRi} = \mu_{PNRi} + \delta.diff_{TNRij} + \beta\,(HAQ.base)_{TNRij}$$
$$\mu_{TRi} = \mu_{PNRi} + \delta.diff_{TRij} + \beta\,(HAQ.base)_{TRij}$$

A common slope was assumed across studies and treatments. Since Abbott only has access to the baseline HAQ for PsARC responders and non-responders from the adalimumab trials, it was not possible to conduct this analysis for all comparators.

The results of the MTC with no adjustment for baseline HAQ are presented in Table 2.2.2.

Table 2.2.2: MTC results for HAQ improvement with no adjustment for baseline HAQ

	Patients who responded to treatment			Patients who did not respond to treatment		
	Mean	Credible intervals		Mean	Credible intervals	
		2.50%	97.5%		2.50%	97.50%
Placebo	-0.2316	-0.4027	-0.0603	0	0	0
Adalimumab	-0.4709	-0.6278	-0.2918	-0.1255	-0.2951	0.0438

*MTC inputs rounded to 2 decimal places

After adjusting for baseline HAQ, the mean HAQ improvement increases for both adalimumab responders and non-responders (Table 2.2.3). In line with the base case assumptions outlined in the Assessment Group Report, this analysis assumes a baseline HAQ of 1.05.

Table 2.2.3: MTC results for HAQ improvement after adjusting for baseline mean HAQ

	Patients who responded to treatment			Patients who did not respond to treatment		
	Mean	Credible intervals		Mean	Credible intervals	
		2.50%	97.5%		2.50%	97.50%
Placebo	-0.2255	-0.4756	0.0411	0	0	0
Adalimumab	-0.4949	-0.8540	-0.1712	-0.1298	-0.3793	0.1296

*MTC inputs rounded to 2 decimal places
*Assuming baseline HAQ = 1.05

As can be seen in Table 2.2.1, this baseline HAQ level is higher than was observed in the adalimumab trials, but lower than the baseline HAQ in either the etanercept or infliximab trials. It is therefore expected that in contrast to the impact on the mean HAQ improvement for adalimumab, adjusting for baseline HAQ would result in a decrease in the mean HAQ improvement for etanercept and infliximab.

The results are shown in Table 2.2.3. The results show a small increase in the mean HAQ from adalimumab (from -0.4709 to -0.4949) but it would be expected that the other drugs would show a decrease in their HAQ improvements (since their baseline HAQ is higher than the mean).

It should be noted that this approach was not adopted in the initial submission by Abbott as only the assessment group has access to data on baseline data for HAQ and PASI by responders and

non responders for each of the studies. Consequently, the analysis provided by Abbott relied on ACR response.

Abbott requests that the Assessment Group re-run the analyses using more precise input values for each of the anti-TNF therapies, and adjusting for baseline HAQ as it appears that these changes will have a significant impact on the results.

2.3 Assumption of different change in HAQ by treatment

On page 114 of the Assessment Report, it states that:

"It is uncertain whether the change in HAQ is the same for all PsARC treatment responders, or depends on the particular biologic treatment followed. In the opinion of our clinical advisor, either scenario could be plausible (Ian Bruce, personal communication) In the base-case model, we allow the change in HAQ for treatment responders to depend on PsARC response and the biologic treatment, and consider the alternative scenario as a sensitivity analysis".

No justification for the decision to allow the change in HAQ to depend on both PsARC response and the biologic treatment is provided. However, this is a key model assumption with sensitivity analyses indicating that adalimumab is no longer dominated by etanercept when the same change in HAQ is applied for all PsARC responders, regardless of treatment. This analysis is presented as scenario 22 in the Assessment Group report, with the results provided in Table 6.6 of the Assessment Group Report.

Comparing the results of this sensitivity analysis against the base case analysis, it is clear that the decision to allow the change in HAQ to depend on both PsARC response and the biologic treatment rather than just PsARC response gives lower effectiveness estimates for adalimumab. As outlined in section 2, Abbott considers that it is highly unlikely that adalimumab would have a lower effectiveness on treating the arthritis component of the disease than etanercept or infliximab. The total costs and QALYs for each strategy are shown in Table 2.3.1. It can be seen that while the assumption of different HAQ change by PsARC response for each treatment increases the QALYs and reduces the costs for both etanercept and infliximab, the opposite is true for adalimumab.

Table 2.3.1: Comparison of Total costs and QALYs for each strategy when HAQ change is and isn't dependent on PsARC response status and choice of biologic

	Base case		Scenario 22	
Strategy	**QALY**	**Cost**	**QALY**	**Cost**
N	5.241	42205	5.241	42205
A	6.642	66408	6.766	66226
E	7.115	72172	7.07	72239
I	7.43	89107	7.347	89230

It therefore appears that had the alternative assumption that all biologics have the same change in HAQ at 3 months for a PsARC responder been made, the conclusion of the Assessment Report may have been different as adalimumab would no longer have been dominated by etanercept in the base case analysis, and many sensitivity analyses would have indicated that adalimumab is in fact the most cost-effective strategy. Such a situation is highly likely since this clinical assumption appears to have been entirely arbitrary.

3. Evidence used in the estimates of clinical effectiveness of the anti-TNFs

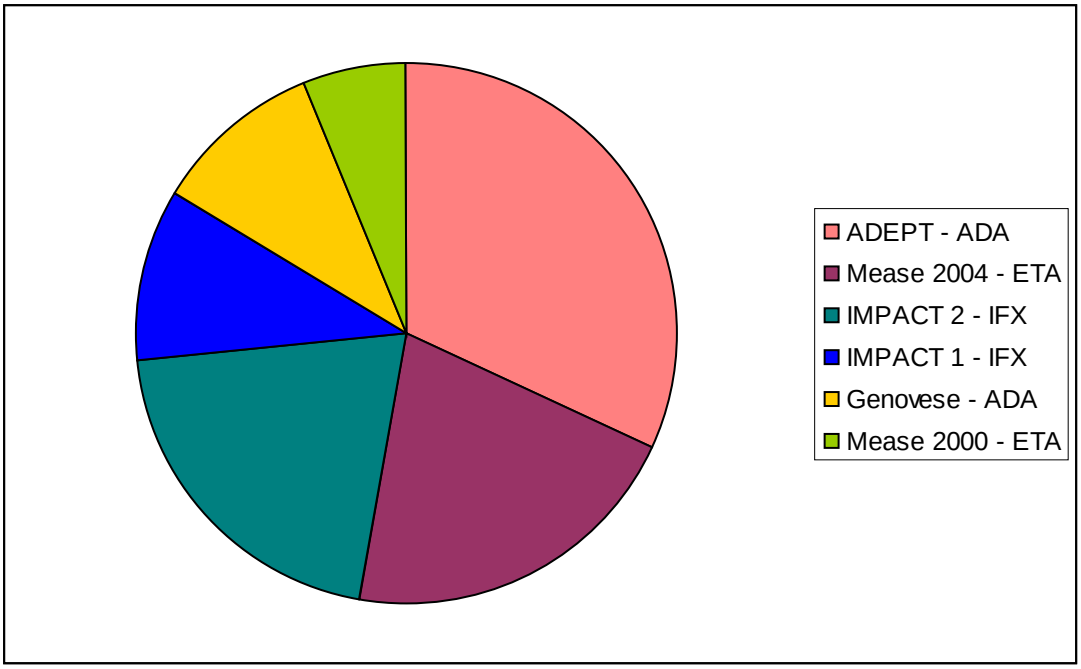
3.1 Impact of smaller anti-TNF RCTs on estimates of efficacy

On page 21 of the Technology Assessment Report (TAR), it states that: "*The response in joint disease (PsARC and ACR) is greater with etanercept than with adalimumab, whereas the*

response in skin disease (PASI) is greater with adalimumab than with etanercept, though these differences are not statistically significant." This statement is based solely on 12 week pooled data from the randomised controlled trials (RCTs) for the two anti-TNFs. Although there were no statistically significant differences between adalimumab and etanercept with regards to joint response, the conclusions from the Assessment group are worded such that etanercept is considered to be the most cost-effective anti-TNF for patients with PsA. Based on the RCT evidence, and given that the annual drug cost of adalimumab and etanercept is equivalent, Abbott considers that this conclusion cannot be robustly supported by the data.

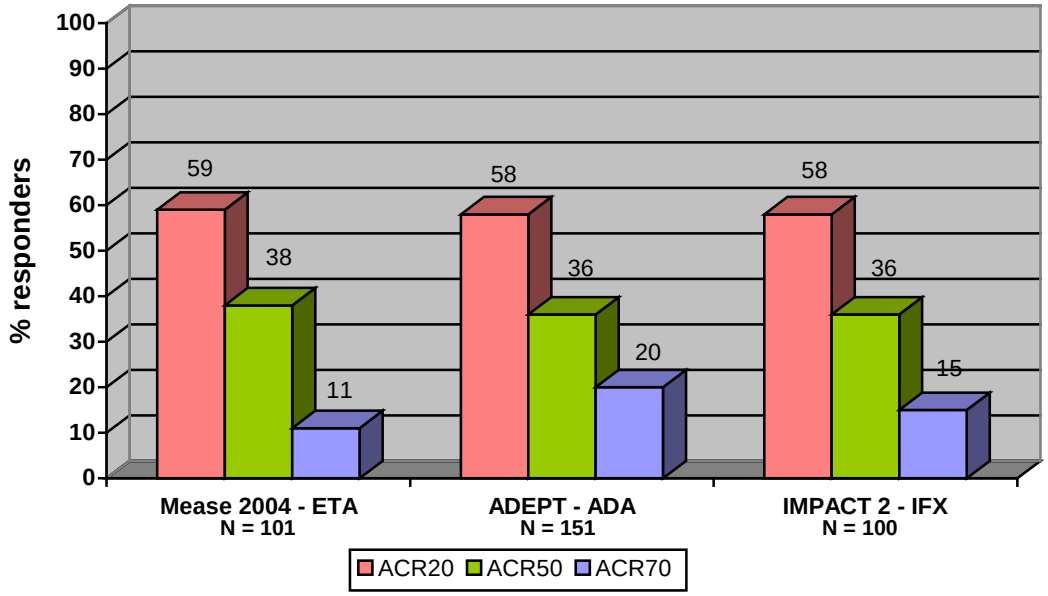
Given that the Assessment Group has only considered the 12 week RCT evidence when determining the effectiveness of the agents, effectively disregarding any open-label extension data and observational data from routine clinical practice in the UK (BSRBR data); Abbott considers it necessary to examine the RCTs included in the effectiveness analysis and highlight the impact the smaller RCTs have on the estimates of efficacy for the different anti-TNFs when the data are pooled. Figure 3.1.1 illustrates the proportion of patients comprising the RCT evidence base for the anti-TNFs in PsA. The largest trial is the ADEPT study (n=313) and the smallest trial is the Mease 2000 etanercept study (n=60).

Figure 3.1.1: Proportion of study participants in the six RCTs of the anti-TNFs in PsA



If the ACR response levels for the active arms of the larger trials for the anti-TNFs are examined, it is clear to see from Figure 3.1.2 that the ACR20 and ACR50 response levels are similar for all the anti-TNFs, and the ACR70 responses for adalimumab are in fact much better than for the other two agents. If there were indeed differences in the joint efficacy between adalimumab and etanercept i.e. etanercept was better than adalimumab, then it would be expected that this would be apparent from the comparison of these studies, which it is not.

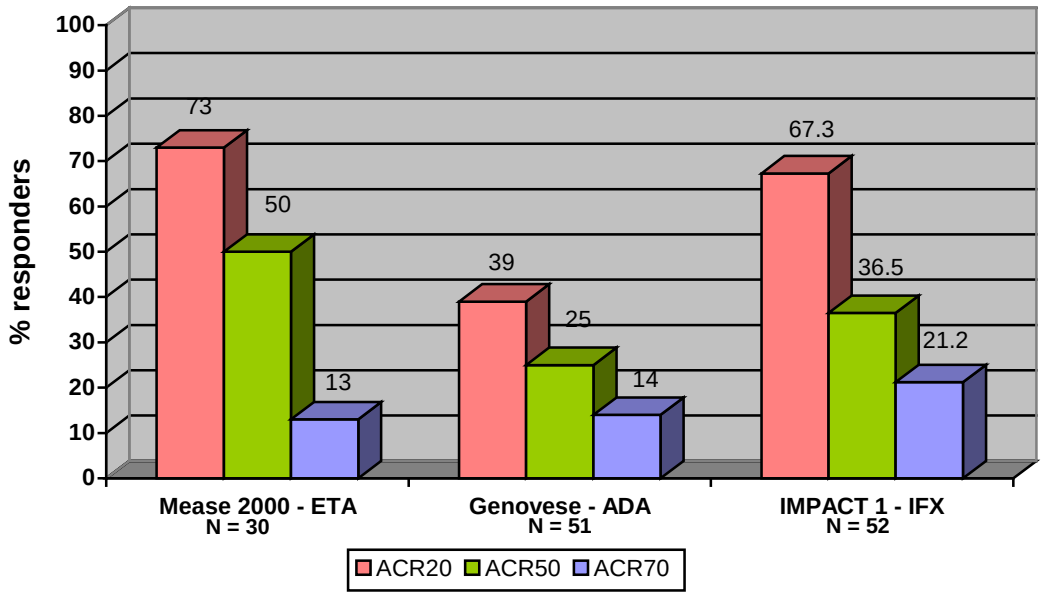
Figure 3.1.2: Percentage of ACR responders at Week 12 for the active arms of the three large anti-TNF trials



Data taken from the Assessment Report. Note: ACR data for the active arms only. Placebo responses for the trials were similar.

However, if the ACR responders for the smaller anti-TNF RCTs are crudely compared, there are some notable differences in the percentage of ACR20 and ACR50 responders between adalimumab and etanercept (Figure 3.1.3).

Figure 3.1.3: Percentage of ACR responders at Week 12 for the active arms of the three smaller anti-TNF trials



Data taken from the Assessment Report. Note: ACR data for the active arms only. Placebo responses for the trials were similar.

These differences in efficacy are based on trials with very small n numbers, where any differences due to chance will have a considerable impact on the results and the sample population means are more likely to differ from the true population mean. Indeed, in some published meta-analyses, trials with fewer than 50 participants per arm have been excluded because the numbers are too small. Abbott is not suggesting that the smaller trials should be excluded from the analysis or that pooling the trials is not statistically appropriate. It is just that when the trials are pooled in an attempt to give a more 'robust' estimate of the effectiveness of each anti-TNF, because there are only two RCTs per anti-TNF, rather than giving a more robust efficacy estimate, pooling the trials instead creates differences in the ACR response rates between the agents, which are not apparent in the larger trials and are therefore based solely on the smaller trials.

Abbott considers that there is a high degree of uncertainty in the Assessment Group's conclusions about the most cost-effective anti-TNF for the treatment of PsA based on differences in joint efficacy arising from a trial containing 30 patients receiving etanercept vs. 51 patients receiving adalimumab for only 12 weeks. Particularly when the patient populations included in each of these two trials are likely not the same (see section 3.3); and when there are data from the BSRBR mimicking routine clinical practice in the UK in a much greater number of patients, which suggest that the three anti-TNFs have similar efficacy in treating the arthritis component of the disease (see section 3.2).

3.2 Importance of registry data mimicking routine clinical practice to the clinical evidence base

Given that there are no head to head trials of the three anti-TNFs, and the efficacy evidence from randomised controlled trial data for each anti-TNF agent is limited to two studies for each drug, it is also important to consider effectiveness data available from observational data sources. There is limited discussion in the assessment report of the larger evidence base for effectiveness of the three drugs based on observational data. The conclusions regarding comparative effectiveness of the drugs are based on the mixed treatment comparison (MTC) data, which are also used to populate the economic model:

"An indirect comparison of the three drugs indicated that infliximab is associated with the highest probability of response on joint and skin outcomes. The response in joint disease appeared greater with etanercept than with adalimumab, whereas the skin response appeared greater with adalimumab than with etanercept, though these differences are not statistically significant. In those patients who achieve a PsARC response to treatment the highest mean reductions in HAQ are seen with infliximab and etanercept." Page 153 of the TAR.

Table 3.2.1 outlines the change in HAQ estimated for adalimumab, etanercept and infliximab in the MTC from the assessment report:

Table 3.2.1: Mean HAQ change estimated for the three anti-TNFs in the Assessment Group's MTC

Drug	HAQ improvement for PsARC responders	Credible intervals	
	Mean	2.5%	97.5%
Etanercept	-0.624	-0.815	-0.438
Infliximab	-0.653	-0.796	-0.509
Adalimumab	-0.423	-0.539	-0.296

Given the magnitude of the differences in HAQ improvement calculated from the MTC of trial data, it would be expected that these differences would also be apparent in the data for effectiveness in clinical practice. The three main sources of observational data available for the effectiveness of adalimumab, etanercept and infliximab in PsA are the BSRBR in the UK[3], the SSATG registry in Sweden[4] and the NOR-DMARD registry in Norway[5]. As noted in the TAR, observational data from other countries are available, but only have data for long term drug

survival, which is less preferable to measuring effectiveness directly. It should, however, be noted that these observational studies show consistently higher discontinuation rates for infliximab than for adalimumab or etanercept. However, there are many reasons why patients may discontinue a drug in clinical practice so assessment using validated outcome measures is more appropriate for assessing effectiveness. As with all observational studies, they are subject to potential bias and confounding. Nonetheless, these registry data sources indicate that adalimumab is not associated with a lower effectiveness in terms of the arthritis component of the disease.

Data from the BSRBR are available in the publication by Saad *et al.* showing improvements in disease activity as measured using the DAS28. These data are shown in Table 3.2.2.

Table 3.2.2: DAS28 and EULAR responses for the three anti-TNFs, data from the BSRBR

	Etanercept (n=333)	Infliximab (n=171)	Adalimumab (n=92)
6-month follow-up (n=480[a])			
Baseline DAS28, mean (S.D.)	6.1 (1.2)	6.3 (1.1)	6.0 (1.0)
6-month DAS28, mean (S.D.)	3.3 (1.4)	3.9 (1.6)	3.3 (1.4)
Mean difference in DAS28 (S.D)	2.8 (1.6)	2.3 (1.7)	2.7 (1.4)
EULAR response			
Good, n (%)	109 (43)	35 (24)	36 (43)
Moderate, n (%)	92 (37)	55 (38)	37 (45)
None, n (%)	51 (20)	55 (38)	10 (12)

[a] Number of patients with complete data on DAS-28 at follow-up

Saad *et al.* note that there were no significant differences in EULAR response rates at 6 (P = 0.679), 12 (P = 0.904) and 18 (P = 0.583) months between the three anti-TNF therapies. These data do not support a higher probability of response with infliximab or etanercept than adalimumab, although it is important to also consider potential biases and confounding factors in this study. The authors noted that there were no significant statistical differences between the three anti-TNF cohorts in age (P = 0.325), sex (P = 0.581) or disease duration (P = 0.384). There was also no significant statistical difference in DAS at baseline among patients receiving the three anti-TNF therapies.

Given the similar baseline characteristics of the adalimumab and etanercept PsA cohorts in the BSRBR and the comparable mean DAS28 clinical improvements observed for these two drugs, Abbott considers that the most reasonable conclusion that can be drawn from the BSRBR data is that adalimumab is not less effective at treating arthritis symptoms. If the magnitude of HAQ improvement predicted for etanercept compared to adalimumab from the MTC were to be observed in clinical practice, it would be expected that this substantially greater efficacy would outweigh any unobservable biases and confounding factors which led to patients with adalimumab having a similar level of response to patients with etanercept. Although smaller in size, similar results were noted in observational studies from Southern Sweden (SSATG registry), which also did not observe significant differences in arthritis response rates between adalimumab, etanercept and infliximab.

Abbott considers that it is more likely that the similar EULAR response rates observed in these two observational studies are due to the similar effectiveness of the drugs on arthritis symptoms rather than bias and confounding factors leading to a lower observed effectiveness with etanercept and infliximab bringing their response rates in line with the improvements observed with adalimumab. Abbott agrees with the statement from Heiberg *et al*. in the discussion section of their publication, summarising the results from other published trials[6]:

"*Although no head-to-head comparisons have been performed between the different TNF-blocking agents, similar magnitude of clinical response has been observed across trials with the different agents with respect to joint symptoms, whereas improvements in skin manifestations seem to be somewhat greater with the monoclonal antibodies*."

Furthermore, evidence of comparable joint effectiveness amongst the three anti-TNFs has been found in rheumatoid arthritis. Nixon *et al.* evaluated the use of mixed treatment comparisons and meta-regression to perform indirect comparisons to estimate the efficacy of biologic treatments in rheumatoid arthritis[7]. The authors found that including study level characteristics of mean baseline disease duration and mean baseline HAQ had a substantial effect on the estimated log odds ratio of an ACR50 event. Results showed that the three TNF antagonists (adalimumab, etanercept and infliximab) appeared to have comparable effectiveness. Furthermore, the between-treatment variability was reduced and the authors reported that including these study covariables accounted for 72% of the between-treatment heterogeneity. This is supported by extensive literature from country specific registries (UK, Denmark, Sweden, The Netherlands, and Spain) which show the anti-TNFs have comparable efficacy in rheumatoid arthritis.

3.3 Different types of patient are being compared in the smaller RCTs

Abbott considers that the differences in joint efficacy reported in the Assessment Group's MTC between adalimumab and etanercept are based solely on differences arising from the smaller RCTs; the impact of these small trials on the cost-effectiveness estimates has already been discussed in Section 3.1. However, an important point to note when comparing these smaller trials is that two different patient populations are being compared indirectly. This could be another contributing factor to the notable differences in arthritis efficacy, in addition to the effect of chance due to the small sample size.

In the Mease 2000 paper, the authors did not use Moll and Wright or equivalent criteria to diagnose a patient as actually having PsA. It is therefore likely that there were a proportion of RA patients in the trial. Conversely, both adalimumab trials specified analysis of PsA sub-type using the Moll and Wright criteria and no RA patients were included. Furthermore, patients were only required to have failed treatment with NSAIDs and not DMARDs in the Mease 2000 etanercept study; whereas patients were required to have failed DMARDs in the Genovese adalimumab study. As a result, given the very low numbers in the etanercept study (n=30 in the treatment arm), even if a very small proportion of RA patients were included in the trial (i.e. 3 or 4), if they were DMARD naïve, one would expect that on average patients would achieve good results based on the evidence. There is plenty of literature in RA which shows that DMARD naïve RA patients receiving an anti-TNF have superior ACR response rates compared to patients who have already failed DMARDs before receiving an anti-TNF. Conversely, PsA patients in the Genovese study had already failed DMARDs prior to starting adalimumab and are therefore a more refractory patient group than the Mease 2000 etanercept study.

In addition, the Mease 2000 study does not give any information about the type of PsA patients in the trial e.g. symmetric polyarthritis, asymmetric polyarthritis, asymmetric oligoarthritis, arthritis mutilans, etc; whereas the adalimumab RCTs do. As a result, it is unclear how many polyarthritic or oligoarthritic patients there are in the etanercept study. This is important, because asymmetric oligoarthritis patients are less likely to show an ACR20 response as they have fewer joints to improve statistically. In the Genovese study around 14% of patients receiving adalimumab were asymmetric oligoarthritis PsA patients. If the small etanercept study included predominantly polyarthritic patients then there would be more scope for improvement for the patients in this trial, thus increasing the possible number of ACR20 or ACR50 responders.

Abbott understands that it is very difficult to account for these differences in sub-types in the modelling, particularly when the studies are so small. However, it is important to highlight that the differences in joint efficacy between adalimumab and etanercept that drive the Assessment Group's conclusions are based on pooled response rates from different sub-types of PsA patient. Therefore, the Assessment Group's conclusions that etanercept is more efficacious in treating the arthritic component of PsA than adalimumab should be treated with caution.

3.4 Use of 12 week PASI response data to extrapolate longer-term effectiveness underestimates the efficacy of adalimumab

On page 114 of the TAR, the Assessment Group state that *"the assessment of effectiveness in Section 5.2.2 did not find any appreciable differences in the biologics' response rates for joint disease or psoriasis between approximately 12 weeks compared with 24 weeks."* As a result of this, the Assessment group used 12-week efficacy data to inform the clinical-effectiveness estimates in their model. Given the strong inference the Assessment Group make in their conclusions about the most cost-effective drug, Abbott considers it necessary to point out that there are appreciable differences in the PASI response rates between weeks 12 and 24 (Table 3.4.1). Although the skin component of PsA is not the biggest driver in the modelling, it is still important to recognise that improvements in psoriasis with adalimumab, when 12 week data are used, have been underestimated in the Assessment Group model. Furthermore, given that the ICERs for adalimumab and etanercept vs. standard care are similar, the improved PASI data at week 24 could have an impact on the conclusions made by the Assessment Group, particularly in PsA patients with moderate to severe psoriasis.

Table 3.4.1: Comparison of Week 12 and Week 24 PASI response rates from ADEPT

	Week 12			Week 24		
	Adalimumab N=69	**Placebo N=69**	**RR**	**Adalimumab N=69**	**Placebo N=69**	**RR**
PASI 50, %	72%	15%	**5.0**	75%	12%	**6.5**
PASI 75, %	49%	4%	**11.3**	59%	1%	**42.0**
PASI 90, %	30%	0%	**41.0**	42%	0%	**58.0**

Note where there are 0 patients who experience an event, a value of 0.5 is used for the relative risk calculations because of the impossibility of dividing through by zero.

3.5 Adalimumab long term radiographic data demonstrates that it reduces the progression of peripheral joint damage in PsA patients

On page 65 of the TAR, the Assessment Group state that: "*Radiographic data from a single controlled trial for adalimumab in PsA demonstrate a beneficial effect on progression of joint disease at 24 weeks. This is a very short time over which to identify a statistically significant effect of therapy and indicates a rapid onset of action of adalimumab. Data from uncontrolled follow-up are inadequate to determine whether any potential delay in disease progression persists at 1-2 years follow-up.*" In addition, on page 155 of the TAR it states: "*Given the fact that the treatment effect on the joint disease is more accurately reflected by the more objective radiographic measure, radiographic long-term data could provide more generalisable estimates of the biologic treatment effect.*"

Abbott can understand why the Assessment group has only used results from the randomised controlled parts of the anti-TNF trials to avoid any potential biases arising from either open-label data or observational studies, although Abbott considers that these data are important in the clinical effectiveness analyses. This is particularly true of outcome measures that evaluate radiographic progression. As the TAR acknowledges, radiographic measures are more objective and are therefore a better reflection of the estimates of biologic treatment effect. Given that the radiographs are blinded, then the results should be measured consistently and objectively regardless of treatment arm. This is supported by evidence submitted to the EMEA in 2008, which led to a change in the wording of the adalimumab licence to include the following: "Humira has been shown to reduce the rate of progression of peripheral joint damage as measured by X-ray in patients with polyarticular symmetrical subtypes of the disease and to improve physical function." The evidence to support this change was based on the 24-week randomised double-blinded data and also on the 2 year open-label data. Therefore, Abbott believes that the data for 1-2 years follow-up are adequate to support the premise that adalimumab inhibits long-term radiographic progression.

3.6 Limitations of the MTC in both the Abbott and Assessment Group MTC

In the MTC for both the Abbott and Assessment Group model, the estimated probability of an ACR70 response for adalimumab is lower than etanercept. Yet, crude comparison of the reported ACR70 data from the trials show that adalimumab has a better ACR70 response than etanercept (Table 3.6.1). Given that the ACR70 response level is a much harder level of response to achieve than either the ACR20 or ACR50, it supports the premise that the joint efficacy data for adalimumab and etanercept are similar.

Table 3.6.1: Comparison of the Week 12 ACR70 response rates for adalimumab and etanercept

Study	Active drug	Placebo	Relative Risk
ADEPT – adalimumab	30/151 (20%)	1/162 (1%)	32.2 (4.4-233.1)
Genovese – adalimumab	7/51 (14%)	0/49 (0%)	13.45 (0.8-230.7)
Mease 2004 – etanercept	11/101 (11%)	0/104 (0%)	22.6 (1.3-380.4)
Mease 2000 - etanercept	4/30 (13%)	0/30 (0%)	8.00 (0.4-144.8)

Note where there are 0 patients who experience an event, a value of 0.5 is used for the relative risk calculations because of the impossibility of dividing through by zero.

Both the Assessment Group and Abbott models link the probability of achieving an ACR70 response to the probability of first achieving an ACR20 response, then achieving an ACR50 response, etc which is a logical approach. However, due to the unexpectedly low ACR20 response rates for adalimumab as a result of the Genovese study, the probability of achieving an ACR70 response is predicted to be lower than etanercept in the MTC because of the hurdle-like approach used to construct the MTC. This low ACR20 response for adalimumab effectively caps the proportion of patients able to achieve an ACR70 response in the MTC.

3.7 Skin improvements were assessed in the Genovese study using Target Lesion Score (TLS)

On page 65 of the TAR, the Assessment Group state that: "*There is limited evidence from a single RCT that adalimumab treatment has a beneficial effect on the psoriasis component of the disease in patients with PsA."* This statement is incorrect as the Genovese study did examine the psoriasis component of the disease using the Physicians Global Assessment of disease and also the Target Lesion Score. Unfortunately the PASI was not used as an outcome measure so it is not possible to include the data in the modelling, however the data do show statistically significant improvements in both PGA and TLS in patients receiving adalimumab compared to patients receiving placebo. At Week 12, the mean target lesion score had decreased from baseline by 3.7 units for adalimumab patients compared with 0.3 units for placebo patients ($p < 0.001$). At Week 12, the physician global assessment for psoriasis was "Clear" or "Minimal" for significantly more adalimumab patients (40.6%, 13/32) than placebo patients (6.7%, 2/30) ($p = 0.002$)[8]. Furthermore, from Week 12 to Week 24, target lesion scores decreased by 4.4 and 0.8 for patients from the placebo and adalimumab arms, respectively, resulting in total improvements from baseline of 4.7 and 4.5. From Week 12 to Week 24, the percentages of patients who had achieved physician global assessments of "Clear" or "Minimal" increased by 43% (from 6.7% to 50.0%) for placebo patients treated with open-label adalimumab, and by 16% (from 40.6% to 56.3%) for patients in the adalimumab arm [8].

3.8 Impact of the psoriasis component of PsA on quality of life

In the Assessment Group model, the psoriasis component of PsA is not given as much weight in patients with moderate to severe skin disease as the arthritis component. Abbott understands that a proportion of PsA patients will not have moderate or severe psoriasis with their arthritic symptoms. However, it is important to acknowledge the impact psoriasis has on quality of life of those patients who do have moderate-to-severe skin disease. Symptoms of the skin component

of PsA occur as visible manifestations that can also cause physical discomfort. The circumscribed, thickened, scaly plaques often cause itching, irritation, and redness, or more severely, physical pain, skin soreness, bleeding from lesions, fatigue and insomnia[9]. Furthermore, the impact of severe psoriasis on health-related quality of life is considered to be similar to that of other major medical conditions including diabetes, heart disease, and cancer[10,11]. Compared to placebo-treated patients in psoriasis clinical trials, adalimumab-treated patients demonstrated significant improvements not only in dermatology-specific quality of life measures (DLQI), but also in general health-related quality of life measures (SF-36) and work productivity measures (WPAI-SHP)[12]. The utility of interrogating the psoriasis rather than the PsA database is that these changes can be ascribed primarily to the effect of adalimumab on skin disease, so these data reinforce that (a) psoriatic skin disease [in PsA or psoriasis] is associated with impairment in general health-related quality of life and work productivity, and (b) that adalimumab is efficacious at mitigating these skin-associated impairments. Abbott considers that improvements in the skin manifestations of PsA should be given greater weight in the consideration of the cost-effectiveness for each intervention in PsA patients with moderate-to-severe psoriasis.

4. Uncertainty in other model inputs

4.1 Probabilistic Sensitivity Analysis

As discussed previously, the base case analysis indicates that the cost-effectiveness of adalimumab and etanercept versus palliative care are very similar, with ICERs of £17,274 and £15,990 respectively. Furthermore, the results of the mixed treatment comparison indicate that there is significant overlap in the credible intervals for response – in particular for ACR and PsARC response rates. These results are presented in Tables 5.14, 5.17 and 5.18 of the Assessment Report and have been summarised in Table 4.1.1 below.

Table 4.1.1: Credible intervals from the Assessment Group Mixed Treatment Comparison

	Adalimumab – credible intervals		Etanercept – credible intervals	
	2.50%	97.50%	2.50%	97.50%
ACR 20	0.429	0.686	0.459	0.750
ACR 50	0.209	0.438	0.231	0.516
ACR 70	0.077	0.205	0.087	0.260
PASI 50	0.552	0.881	0.236	0.592
PASI 75	0.275	0.693	0.085	0.313
PASI 90	0.120	0.452	0.032	0.145
PsARC response	0.444	0.713	0.566	0.832

It is therefore surprising that the probabilistic sensitivity analysis shows that there is very little uncertainty in which is the most cost-effective of these two treatments (p=0.524 for etanercept and p=0.044 for adalimumab).

Sensitivity analysis 22 suggests that this apparent lack of uncertainty is driven to a large extent by the HAQ change by PsARC responder/non-responder. In this analysis, it is assumed that HAQ change depends only on PsARC response rate, and does not differ between treatments, which clinical opinion suggests is an equally plausible assumption which produces quite different results (see section 2.3 for further discussion). Under this assumption, the probability that adalimumab is the most cost-effective treatment at a threshold of £20,000/QALY increases to 0.198 while the probability that etanercept is the most cost-effective treatment falls to 0.400. As discussed in section 2.1, Abbott anticipates that this trend will be continued if the MTC is amended to use more precise input values for each of the anti-TNF therapies, and adjustments are made for baseline HAQ.

4.2 Subgroup analyses

The Assessment Group model is a cohort model, and therefore assumes a homogeneous mix of patients. Although the base case patient characteristics were selected based on expert opinion as to the most common patient type observed in clinical practice, these characteristics by definition represent only a subgroup of the PsA patient population and are not reflective of the mix of PsA patients. In recognition of this limitation, the Assessment Group conducted some subgroup analyses using alternative patient characteristics (table 6.8 of the Assessment Group Report).

The Assessment Group has shown that the patient characteristics have a significant impact on the results of the cost-effectiveness analysis. In order to obtain an accurate picture of the expected cost-effectiveness of each of the treatments, it is therefore important to consider the results of these subgroup analyses alongside the base case results.

Furthermore, the base case analysis assumes that patients continue to receive treatment only if a PsARC response is achieved at 3 months in line with the BSR guidelines. However, the BAD guidelines state that a patient should also continue to receive treatment if a PASI 75 response is achieved. Abbott considers that it is reasonable to assume that patients with both skin and joint involvement would be managed by both a rheumatologist and a dermatologist, and that both of these guidelines would therefore apply. Scenario analyses conducted by the Assessment Group indicate that when using either the BSR or the BAD stopping rule, adalimumab is the most cost-effective treatment when using a threshold of £20,000/QALY.

4.3 Sensitivity analysis varying baseline HAQ

As discussed in section 2.2, the change in HAQ score is a key model input which is modelled using a random-effects meta-analysis. The code for this analysis indicates that the change in HAQ score depends on the baseline HAQ. However, when the Assessment Group conducted the sensitivity analysis in which the baseline HAQ was increased from 1.05 as per the base case analysis to 1.8 (analysis 10), it appears that they failed to alter the change in HAQ score simultaneously. This sensitivity analysis is therefore incorrect and does not accurately reflect the expected cost-effectiveness in this population.

5. Infliximab Costs

Abbott notes that the Assessment Group assumes an average patient weight of 70kg based on the average weight of the UK population. In order to determine whether this weight is representative of the psoriatic arthritis population, Abbott used data from M02-570 and ADEPT trials for adalimumab, and the smaller etanercept trial (Mease, 2000) to calculate the average weight of moderate to severe PsA patients enrolled in clinical trials. Patient weight was not reported for either of the infliximab clinical trials, nor in the larger etanercept trial (Mease, 2004). The average weight from these three trials was calculated to be 87kg.

Since infliximab has a weight-based dosing schedule, patients weighing over 80kg would require one additional vial than patients weighing 70kg which will increase the costs associated with infliximab (Table 5.1).

Table 5.1: Infliximab drug costs

	Vials per dose	Doses	Cost per vial	**Total cost**
0-12 weeks				
Infliximab (70kg patient)	4	3	£419.62	**£5,035.44**
Infliximab (87kg patient)	5	3	£419.62	**£6,294.30**
12-24 weeks				
Infliximab (70kg patient)	4	2	£419.62	**£3,356.96**
Infliximab (87kg patient)	5	2	£419.62	**£4,196.20**
24 weeks +				
Infliximab (70kg patient)	4	1.625	£419.62	**£2,727.53**
Infliximab (87kg patient)	5	1.625	£419.62	**£3,409.41**

Furthermore, the Assessment Group assume a ½ day in-patient hospital cost for each infusion of infliximab at a cost of £144 per infusion. However, since an infliximab infusion is more likely to be a day case rather than an in-patient procedure, this would be a more appropriate cost to use. The NHS reference costs (2007/08) indicate that the cost would therefore be £462[13].

Abbott therefore feels that both the drug and administration costs of infliximab have been underestimated in the Assessment Group Report.

6. Issues raised with the model submitted by Abbott Laboratories

6.1 Mixed treatment comparison (MTC)

There were a number of questions surrounding the approach of the MTC used in the Abbott submission.

Abbott adopted the approach used by Woolacott *et al* (2006)[14] for the meta-analysis of the PASI50, 75 and 90 responses rates from the RCTs for comparing all treatments for moderate to severe psoriasis. In that study, the end-points were jointly modelled using an ordered probit model. The ordered probit model is to model a discrete dependent variable *y* which takes ordered multinomial outcomes and the model can be expressed in terms of an underlying latent variable *y**. In the context of psoriasis, *y** be interpreted as the patient's underlying percentage reduction in PASI score; the higher the value of *y**, the more likely the patient is to report a higher category of the PASI response (<50, 50-75, 75-90, >90). The model was implemented as a Bayesian hierarchical model.

The univariate ordered probit model was simply extended to a bivariate ordered probit model in the meta-analysis.

6.2 Withdrawal rates

6.2.1 Derivation of the Weibull parameters for anti-TNF therapies

Comment: Parameters for a Weibull distribution were derived using longitudinal data from three time points, and the data were assumed to be independent. This assumption is incorrect, because the same patients contribute data to the probability of survival at 2 years as 1 year. (p. 254)

The data were not assumed to be independent as the function was fitted to all data points simultaneously.

6.2.2 Derivation of the Weibull parameters for conventional DMARDs

Comment: Weibull distribution used. Unclear how this was specified as only 1 data point reported (Malesci et al., 1997i). (p. 138)

Abbott accepts that there are significant limitations with the survival function for DMARDs. In order to address these issues substantial sensitivity analyses were conducted around this parameter.

6.2.3 Choice of Weibull distribution

Comment: no justification was given for the choice of Weibull distributions rather than other parametric distributions. It may be that other distributions offered a better fit. Secondly, the 1 year rates from the BSBDR are likely to include non-responders to biologics in addition to those who withdraw due to loss of efficacy or adverse events after the initial 3-month period. As these initial withdrawals are already counted as non-responders, there is a degree of double counting. (p. 275)

Based on data from other rheumatological conditions, the prior belief was that a Weibull distribution was the most appropriate parametric distribution. This distribution was used in the model since there was not sufficient information to refute this belief.

6.2.4 Use of BSRBR

Comment: Withdrawals after 3-months due to adverse events and lack of efficacy were estimated from a single dataset (BSR register) in all of the industry models. There are other potential biologic registry datasets available which could have been synthesised. (p. 109)

Abbott accepts that there are limitations surrounding the use of a single dataset to determine withdrawal rates after 3 months. In order to address these issues substantial sensitivity analyses were conducted around this parameter.

6.3 Costs

6.3.1. Cost year and country

Comment: The year and country to which unit costs apply is stated with appropriate adjustments for inflation and/or currency conversion – not done (p. 255)

The cost and resource use sources are provided along with the year of publication for each of the cost inputs in the relevant section of the submission (3.4.6, 3.4.7 and 3.4.8). The information provided in these sections are summarised in Table 6.3.1.

Table 6.3.1: Cost year and country

	Cost source	**Resource use source**
Drug Costs	MIMS, 2009	MIMS 2009, or University of Toronto dataset (for DMARD mix)
Administration and Monitoring Costs	NHS Reference Costs 2007/08, or HTA report for previous NICE appraisal in RA – (Barton et al) inflated to 2008 costs using HCHS pay and prices index (Curtis et al., 2008)	British Society of Rheumatologists
Hospital Costs	NHS Reference Costs 2007/08,	Norfolk Rheumatoid Arthritis Registry (NOAR)

6.3.2 Estimation of uncertainty around NOAR data

Comment: As the NOAR data did not include any measure of uncertainty in the mean estimates of resource use, the estimates of the standard errors of mean costs in the Abbott submission cannot be valid. (p. 282)

This comment is correct, however, it would also be incorrect to assume that these values are known with full certainty. Reasonable errors were therefore assumed to more accurately portray the uncertainty.

6.4 Other

6.4.1 Model

Comment: Do not give adequate justification for why an individual sampling model is used (p261)

An individual sampling model is used to enable the incorporation of a sequence of treatments.

6.4.2 PASI transformation

Comment: In order to estimate the PASI the data was transformed by Log(PASI+0.5). The authors state this was done "to obtain normality". It is important to note that this log-transformation assumes that a 1% improvement in PASI will lead to a constant change in utility, regardless of the absolute change in PASI. For example, this regression assumes that a reduction in PASI score from 16 to 0 leads to the same change in HRQOL as a reduction in PASI score from 8 to 0. A linear regression on the other hand assumes that a reduction in PASI by 16 points gives twice the HRQOL benefit of a reduction in PASI by 8 points, regardless of the baseline. (p. 251)

This is a valid criticism. However, since the size of the PASI is very small, this issue has very little impact on results, and needs to be compared with the problems of using the PASI untransformed.

Appendix 1

				HAQ at Baseline		HAQ at Week 12		Changes in HAQ from baseline		
Trial	**PsARC**	**Treatment Group**	**n**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**SE**
ADEPT	NO	ADALIMUMAB	58	0.8199	0.6613	0.9397	0.6134	-0.1198	0.3997	0.0525
	NO	PLACEBO	120	1.0469	0.7332	1.0208	0.6665	0.0260	0.4012	0.0366
	YES	ADALIMUMAB	93	0.5013	0.5745	1.0013	0.6270	-0.5000	0.4295	0.0445
	YES	PLACEBO	42	0.6956	0.6149	1.0089	0.6815	-0.3134	0.4930	0.0761
Genovese	NO	ADALIMUMAB	25	0.9250	0.6465	1.0750	0.4974	-0.1500	0.4521	0.0904
	NO	PLACEBO	37	0.9291	0.7433	0.9865	0.7400	-0.0574	0.3222	0.0530
	YES	ADALIMUMAB	26	0.2548	0.2926	0.6779	0.4733	-0.4231	0.4124	0.0809
	YES	PLACEBO	12	0.8438	0.7578	1.0208	0.7739	-0.1771	0.2162	0.0624

References

[1] Pocock S, Assmann SE, Enos LE, Kasten LE. Subgroups analysis, covariate adjustment and baseline comparisons in clinical trials reporting: current practice and problems. Stat Med 2002; 21: 2917–2930.

[2] Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. Health Econ. 2005 May;14(5):487-96

[3] Saad AA, Ashcroft DM, Watson KD, Symmons DP, Noyce PR, Hyrich KL; on behalf of the BSRBR. Efficacy and safety of anti-TNF therapies in psoriatic arthritis: an observational study from the British Society for Rheumatology Biologics Register. Rheumatology (Oxford). 2010 Jan 7. [Epub ahead of print]

[4] Kristensen LE, Gulfe A, Saxne T, Geborek P. Efficacy and tolerability of anti-tumour necrosis factor therapy in psoriatic arthritis patients: results from the South Swedish arthritis Treatment Group register. Ann Rheum Dis 2008;67:364-369

[5] Heiberg MS, Kaufmann C, Rodevand E, Mikkelsen K, Koldingsnes W, Mowinckel P, Kvien TK. The comparative effectiveness of anti-TNF therapy and methotrexate in patients with psoriatic arthritis: 6- month results from a longitudinal, observational, multicenter study. Ann Rheum Dis. 2007 Jan 9

[6] Heiberg MS, Kaufmann C, Rodevand E, Mikkelsen K, Koldingsnes W, Mowinckel P, Kvien TK. The comparative effectiveness of anti-TNF therapy and methotrexate in patients with psoriatic arthritis: 6- month results from a longitudinal, observational, multicenter study. Ann Rheum Dis. 2007 Jan 9; [Epub ahead of print]

[7] Nixon RM, Bansback N, Brennan A. Using mixed treatment comparisons and meta-regression to perform indirect comparisons to estimate the efficacy of biologic treatments in rheumatoid arthritis. *Stat Med* 2007;**26**(6):1237-54.

[8] Genovese M, Mease P, Thomson G, Kivitz A, Perdok R, Weinberg M, Medich J, Sasso E. Safety and efficacy of adalimumab in treatment of patients with psoriatic arthritis who had failed disease modifying anti-rheumatic drug therapy. *The Journal of Rheumatology* 2007;**34**(5)1-11.

[9] Mease PJ, Menter MA. Quality-of-life issues in psoriasis and psoriatic arthritis: Outcome measures and therapies from a dermatological perspective. *J Am Acad Dermatol.* 2006;**54**:685-704.

[10] Rapp SR, Feldman SR, Exum ML, Fleischer AB,Jr, Reboussin DM. Psoriasis causes as much disability as other major medical diseases. *J Am Acad Dermatol.* 1999;**41**:401-407.

[11] Krueger G, Koo J, Lebwohl M, et al. The impact of psoriasis on quality of life: results of a 1998 National Psoriasis Foundation patient-membership survey. *Arch Dermatol.* 2001;**137**(3):280-284.

[12] Revicki DA, Willian MK, Menter A, Gordon KB, Kimball AB, Leonardi CL, Langley RG, Kimel M, Okun M. Impact of adalimumab treatment on patient-reported outcomes: results from a Phase III clinical trial in patients with moderate to severe plaque psoriasis. J. Dermatolog Treat. 2007; 18(6): 341-350.

[13] NHS Reference Costs 2007/08 Day Case HRG code: HD23C (Inflammatory Spine, Joint or Connective Tissue Disorders without cc)

[14] Woolacott N, Hawkins N, Mason A, et al. Etanercept and efalizamab for the treatment of psoriasis. Health Technology Assessment 2006; Vol.10: No. 46