COMPARISONS OF SUBMITTED MODELS FOR THE APPRAISAL OF ETANERCEPT, ADALIMUMAB AND INFLIXIMAB FOR ANKYLOSING SPONDYLITIS

REPORT BY THE NICE DECISION SUPPORT UNIT.

13 February 2007

Nick Bansback, Centre for Health Evaluation and Outcome Sciences, University of British Columbia
Allan Wailoo, Health Economics and Decision Science, ScHARR, University of Sheffield

1. INTRODUCTION

All three manufacturers submitted cost effectiveness models to NICE as part of this appraisal. Schering (infliximab) submitted a model in DATA Treeage, whilst Wyeth (etanercept) and Abbott (adalimumab) both submitted Excel based models. The assessment group (LRiG) undertook detailed critical appraisal of these submissions in the assessment report and identified possible errors in each case. The manufacturers claim to have corrected some of these errors but it is not part of the appraisal process for corrected versions of these models to be submitted.

The DSU has attempted to replicate results where manufacturers claim to have made corrections to models and submitted revised results. We have made amendments to the submitted models but the DSU has not conducted a full review of any of the models, including the assessment group model.

We next apply a common set of assumptions and parameter values, where feasible, to each of the corrected manufacturer models. These assumptions and parameters were those decided by the committee to be most plausible in their discussions of the preliminary recommendations and referred to in section 4.3.9 of the ACD.

We finally apply the drug and monitoring costs of each TNF inhibitor to each of the manufacturers models. For this, we make the assumption that all benefits and

characteristics within each model are the same for each treatment, and only the drug and monitoring cost differs.

2. REPLICATION OF STATED CORRECTIONS TO THE MANUFACTURER MODELS

A) SCHERING PLOUGH MODEL

A1 No BASFI progression after withdrawal from TNF-alpha.

The LRiG assessment report identified that there is no BASFI progression in the Schering model when patients withdraw from infliximab. This is the case both for patients that withdraw early i.e. within 24 weeks for the Assert model and within 12 weeks for the Braun model, as well as for patients that withdraw from infliximab beyond that point.

Figures 1 and 2 highlight the magnitude of this error in terms of BASFI. Figure 1 shows the extent of the BASFI gain in early withdrawers (45% and 50% of the entire treatment arm in the Braun and Asset models respectively). The overestimate is less but still substantial in those that withdraw later (figure 2).

Since BASFI is a major driver of both utility and cost in the model, it is reasonable to expect the impact on ICERs to be substantial. In the base case analysis Schering report ICERs of £18k and £19k for Braun and Assert respectively. Table 1 shows how these figures rise to £28k and £27k after Schering implemented corrections to BASFI progression (Schering Plough Ltd. Response letter to NICE, 10th July 2006, page 7). The version of the model amended by DSU generates much higher ICERs than this (£46k and £42k for Braun and Assert respectively).

The results generated by the DSU amended model are similar to those reported by LRiG. This is true both in terms of the overall ICERS and the incremental costs and benefits considered in isolation.

Using the probability distributions in the models submitted by Schering produced cost effectiveness acceptability curves for the DSU corrected models as shown in figure 3. For both ASSERT and BRAUN the probability of infliximab being cost effective at £40k per QALY is below 0.02.

Figure 1: BASFI progression error in Schering model – early withdrawers

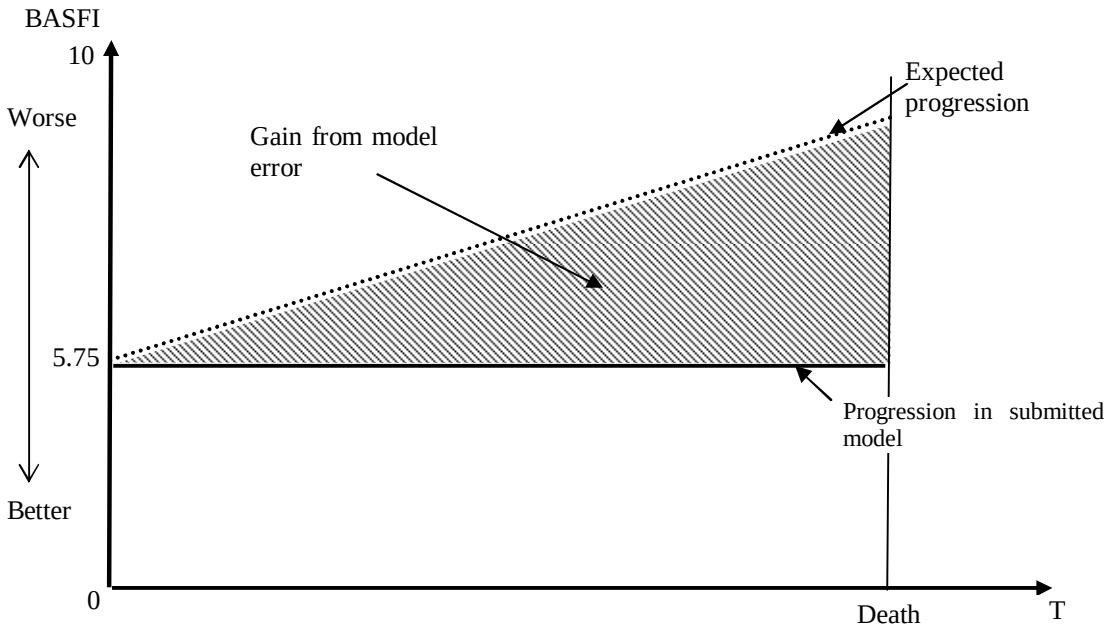


Figure 2: BASFI progression error in Schering model – late withdrawers

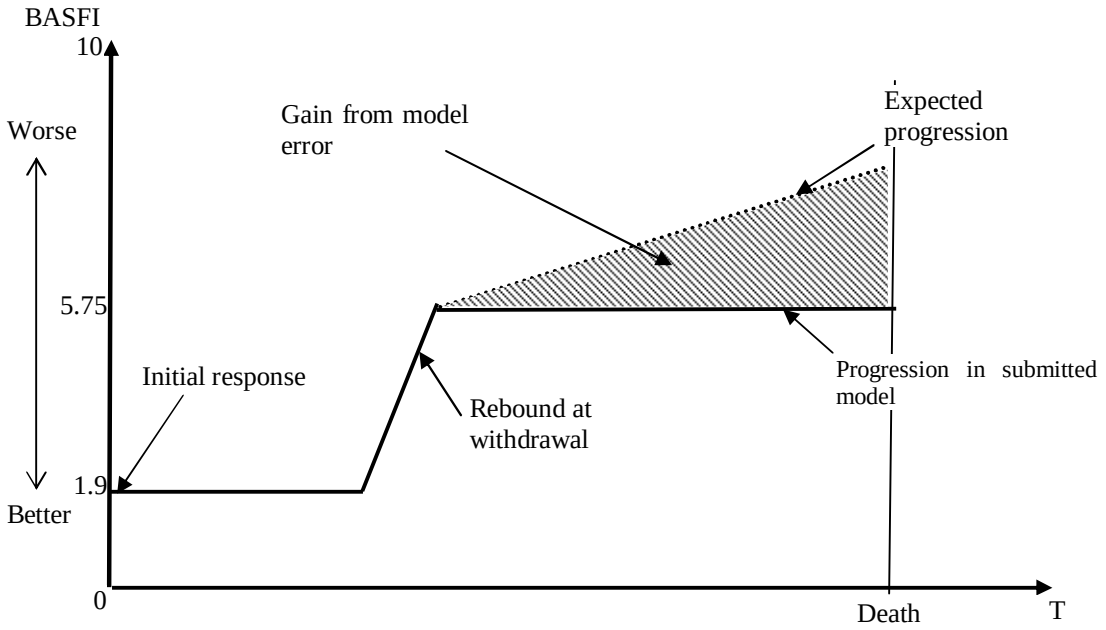
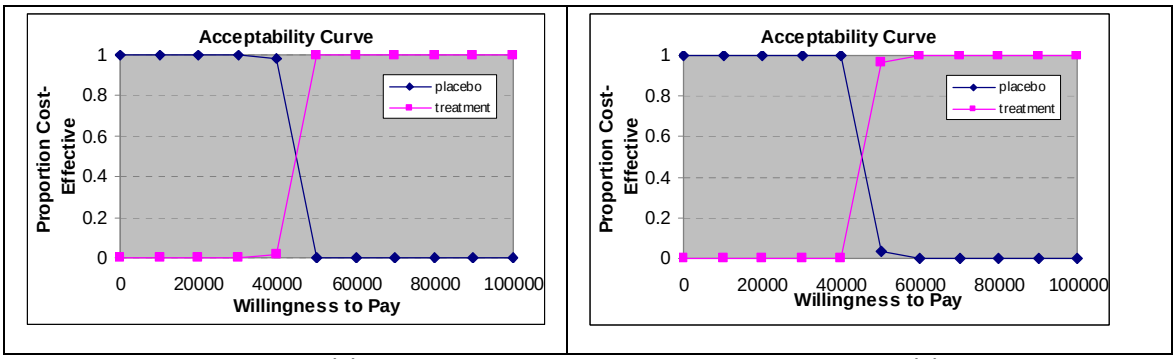


Table 1: Comparison of DSU amendments to Schering model, LRIG replication and Schering results

	Costs (£ 000's)	Incremental Costs	QALYs	Incremental QALYs	ICER
ASSERT model					
DSU correction					
placebo	70		10.25		
treatment	110	40	11.21	0.96	***41,959***
LRIG Schering replication		40		0.976	***40,889***
Schering corrected					***26,751***
BRAUN model					
DSU correction					
placebo	65		10.57		
treatment	106	42	11.48	0.91	***45,659***
LRIG Schering replication		44		0.88	***50,380***
Schering corrected					***28,332***

Figure 3: Cost effectiveness acceptability curves for DSU revised Schering models



Assert model Braun model

The final corrected result for infliximab vs conventional care is 42k per QALY (Assert model) or 46k per QALY (Braun model)

B) WYETH MODEL

In the base case analysis, Wyeth reported an ICER of £13k per QALY. Several criticisms of the model were given in the LRiG review which we have attempted to implement in the submitted model.

B1 The cost of adverse events was not incorporated

When the cost of adverse events was incorporated into the model, the ICER increased to £13,301

B2 The gender parameter was the inverse

When male and females were swapped in the model, the ICER increased to £13,297

B3 Some samples of age were below 18

When the minimum age was set at 18, the ICER decreased to £12,866

B4 The life expectancy calculations were incorrect

The calculations appear to be correctly performed.

B5 BASDAI and BASFI progression rates were not realistic

When the BASDAI progression is set to 0 and BASFI progression is set to 0.7, the ICER increases to £15,891

B6 The AS cost relationship was not correctly implemented

When the revised cost relationship was implemented, the ICER increased to £16,696

B7 BASDAI and BASFI initial values

The initial BASDAI and BASFI values (baseline to 24 week) were criticised for not correctly incorporating the correlation between measures, and that patients with higher baseline values had different magnitude of changes to patients with lower baseline values. The manufacturer describes that the model was revised to correct for this but insufficient detail is given to replicate this. The manufacturer reports that, when also including the correction described above, the ICER increased to £22,704. Since the sensitivity analysis on BASDAI and BASFI change shows these parameters to not be hugely sensitive, it seems reasonable that the ICER would increase by this amount, but is impossible to verify the exact number.

The final corrected result (B1-B6)for etanercept vs conventional care is 20k per QALY

NB This does not include the correction B7 as explained above

C) ABBOTT MODEL

In the base case analysis Abbott report an ICER of £23k per QALY. Three main criticisms of the model were given in the LRiG review. Abbott maintained the validity of their model and did not submit revised ICERs. The two criticisms are addressed below and their impact on the ICER explored.

C1 The model re specifies patient characteristics at 48 weeks onwards.

The analysis submitted by Abbott follows individual patients from the actual adalimumab studies for the first 48 weeks. Beyond this time, average characteristics are used to predict the long term progression of patients. This appears to mix patient and cohort level parameters and could potentially bias the results. Abbott argue that the patient characteristics on each arm are not identical and thus with a long extrapolation could lead to misleading results. While both arguments have plausible rationales it is not clear whether the assumptions would bias treatment or not and, given the design of the model, it is not possible to test without major revision of the code. However, it appears unlikely that the magnitude of different will be large and when some of the factors are varied in the sensitivity analysis, the result barely changes (£23,330 from £23,029)

C2 Patients rebound to placebo average BASDAI and BASFI after withdrawal from adalimumab

Again, since the model changes from a patient level to a proxy cohort level model from 48 weeks onwards, it was not possible to rebound BASDAI and BASFI values back to the exact baseline values. Instead, they rebound to average values from the non treatment arm. While this is not the same assumption about rebound that has been used in the past, it does mean that there is no further benefit beyond the withdrawal of TNF therapy and is therefore conservative.

The final corrected result for adalimumab vs conventional care is 23k per QALY

3. IMPLEMENTATION OF COMMON PARAMETER VALUES

The common parameters implemented were drawn from section 4.3.9 of the ACD. These are:

1. Patients not on anti-TNFα treatment do not experience improvements in their condition
    - As with all manufacturer models
2. BASFI progression prevented whilst on anti-TNFα treatment
    - As with all manufacturer models
3. BASFI progression rate 0.07 per annum
    - Schering model – 0.07
    - Wyeth original model – 0.03, correct – 0.07
    - Abbott model – 0.05
4. Annual anti-TNFα withdrawal rate of 7% per annum
    - Wyeth and Abbott models – 10% per annum
    - Schering models – 15% per annum
5. Baseline BASDAI/BASFI averages: 6.5/5.6
    - Wyeth model - 6.1/5.9
    - Schering model - 6.41 / 5.75 and 6.3 / 5.4
6. Assessment Group base case utility model (the same as the Schering-Plough utility model)
    - Wyeth model – age not included
    - Abbott model – age not included
7. Assessment Group base case assumptions for cost parameters
        - Schering models – changes limited to drug costs. For infliximab this is 4 vials per infusion (mean 3.875 Assert, 3.685 Braun) and £267 per infusion (£88.56)).
        - Wyeth corrected model – 383.75*exp(0.19225*BASFI)
        - Abbott - 795.45 + 680.20 BASFI.
        - Both amended to 1585.30*exp(0.1832*BASFI)
8. Assessment Group base case assumptions for BASDAI progression

- Schering and Wyeth corrected models – Patients on treatment experience a fall in BASDAI that returns to baseline on withdrawal of infliximab. No progression for placebo arm.
- Abbott model – Patients on treatment experience a fall in BASDAI that returns to the level of BASDAI experienced by the conventional treatment arm at the point of withdrawal

In addition, we have explored the implementation of a 20 year time horizon, as in the LRiG base case.

- Schering models – 70 years
- Abbott and Wyeth models - lifetime

Table 2: ICERs for corrected manufacturer models using common parameter values

Parameter	Schering Model (ICER per QALY)		Wyeth Model (ICER per QALY)	Abbott Model (ICER per QALY)
	Assert	Braun		
Implementing common parameters				
Original	41959	45659	19645	23097
1	41959	45659	19645	23097
2	41959	45659	19645	23097
3	41959	45659	19645	22837
4	40507	43723	14650	22619
5	Not implemented. Base values 6.41 / 5.75	6.3 / 5.4	19203	Not implemented Base value 6.3/5.4
6	41959	45659	22910	25082
7	48819	55917	17440	17039
8	41959	45659	19645	23097
All above	47112	53914	19383	17427
All above 20yr horizon	50485	58148	19889	Not implemented
Replacing individual drug and monitoring costs into each model				
Etanercept vs placebo	22136	25805	19889	17427
Adalimumab vs placebo	22136	25805	19889	17427
Infliximab vs placebo	47112	53914	38934	42854
Comparing costs and benefits in a multiway comparison				
Etanercept/ Adalimumab	22136 / 25805		19889	17427
Infliximab	Dominated		Dominated	Dominated
Placebo	-		-	-

The results are presented in table 2. Many of the common parameters are present in the Schering base case model. Changes to other parameters make only slight

differences to the ICERs. It was not possible to implement changes to baseline BASFI/ BASDAI, although the base case values are only slightly different and therefore unlikely to have any significant impact.
Similarly, with both the Abbott and Wyeth models, no substantial changes in the ICERS are observed from implemented parameter changes.

When the drug and monitoring costs for each of the drugs were entered into the three models, consistent results are seen. The ICERs for infliximab are £47k, £39k and £43k in the Schering, Wyeth and Abbott models respectively. Adalimumab and etanercept generate ICERs of between £22k and £26k in the Schering models, slightly higher than the results in the Wyeth and Abbott models.

CONCLUSION

The only major problem we found was with the way Schering corrected the BASFI progression rate in their model to estimate of the cost-effectiveness of infliximab.

When DSU corrected the model, the ICER increased to 42k per QALY (Assert model) or 46k per QALY (Braun model). This compares to ICERs of £28k and £27k per QALY after Schering implemented their own corrections.

Once this correction was appropriately made, all three manufacturer models give relatively consistent results.

These are that in comparison to conventional care, infliximab has an ICER of above 40k per QALY, and both etanercept and adalimumab have ICERs of around 20k per QALY. The principal driver of these differences seems to be the drug costs.