USE OF TUMOUR NECROSIS FACTOR ALPHA (TNF A) INHIBITORS (ADALIMUMAB AND INFLIXIMAB) FOR CROHN'S DISEASE

REPORT BY THE DECISION SUPPORT UNIT

Allan Wailoo, Jon Tosh, Pippa Hemingway
School of Health and Related Research, University of Sheffield

10th June 2009

NICE Decision Support Unit
ScHARR, University of Sheffield
Email: [redacted]
Tel: [redacted]

CONTENTS

1. INTRODUCTION

The purpose of this report is to consider the reasons for differences between models submitted for the appraisal of the anti-TNFs, adalimumab (Abbott Laboratories Ltd, "Abbott") and infliximab (Schering-Plough, "SP"), for the treatment of severe active Crohn's disease. Two sets of models were submitted by the manufacturers as well as an independent assessment group model (the "Leeds model").

In our previous report[1], some of the key structural and parameter differences between these models and an independently submitted analysis were highlighted. It was not considered feasible to draw firm conclusions about the reasons for the differences between the models without further investigation and data review. This report outlines the additional work that has been undertaken.

First we provide details of the key differences in modelling approaches. We then consider one of the central criticisms of the Leeds model which was explored in detail by one of the manufacturers (Abbott): the use of transition probabilities from a single, cohort study and in particular the probability of transiting from remission to relapse for patients receiving standard care (the relapse rate). We report a review of the literature and contrast this with the value used in the Leeds model and that proposed by Abbott based on their own systematic review and synthesis. We highlight the results of implementing these changes on the estimated ICERs.

We then make additional changes to the Leeds model in order to make the approach more consistent with a) the SP modelling approach and b) the Abbott modelling approach. Conclusions can be found in section 5.

2. GENERAL MODELLING APPROACHES

In the following sections we concentrate on models presented for severe, adult Crohn's disease. Key features of the three models are presented in

Table 1.

Table 1: Key features of submitted models

Structural Issues	*Schering Plough*	*Abbott*	*Leeds*
Model type	Transition state cohort model	Trial based health state decision model	Transition state cohort model
Cycle length	2 weeks (cycle 1) 4 weeks (cycles 2-4) 8 weeks	Trial follow up	4 weeks
Key states	Remissions, relapse, surgery, death	Remission, moderate, severe, very severe	Remission, relapse, surgery
Duration	5 years	1 year basecase, lifetime (28 years) extrapolated analysis	1 year basecase
Source for Transition Relapse/Remission	Initial response - TARGAN	CHARM for induction and subsequent maintenance therapy	Silverstein for standard and episodic care
Transition probabilities	Time varying	Time varying	Fixed over time
Cost year	2005/06	2006	2005/06
Discount rate	3.5% costs and benefits	None in basecase, 3.5% costs and benefits	None
Mortality	General Population Mortality	No mortality in 1 year, all patients die at 60 in lifetime model	No mortality

2.1 THE LEEDS MODEL

Leeds present separate models for infliximab and adalimumab which are identical in structure. Comparisons between standard care, episodic treatment and maintenance treatment are made over a one year timeframe, using a cohort, state transition model with transition probabilities that are constant over time for each of the 4 week cycles. Model states reflect remission, relapse, surgery and post surgical remission. Perhaps importantly, there is no partial response: treatment benefits are only demonstrated in the model directly by its ability to distinguish relapse from full remission. Furthermore, there is no mortality effect in the model.

The model considers patients in the episodic and maintenance arms to be “on treatment” provided they do not experience two consecutive periods of relapse. Model

transitions are derived from a single study[2] which reports a retrospective cohort study of all patients diagnosed with Crohn's disease between 1970 and 1993, resident in Olmsted County, Minnesota. As stated in our previous report, we were unable to replicate the transition probabilities used in the Leeds model from the published paper.

The effect of treatment on these transition probabilities comes from a single value of 0.56 for adalimumab (based on 6 week CHARM trial TNF arm events 96/172) and 0.56 for infliximab (based on 6 week ACCENT 1 TNF arm events 63/113). These are the proportions of patients in remission at 6 weeks in the respective trials. These probabilities are applied at the first and every subsequent 4 week cycle in the model. The control group data from the trials does not feature in the Leeds model.

Table 2 shows the baseline cost effectiveness estimates from the Leeds model. The episodic treatment strategy dominates standard care for both infliximab and adalimumab. The ICERs for maintenance therapy, compared to episodic therapy, are in the region of £5m. It is worth that the numbers of QALYs generated for each of the three strategies are almost identical in the infliximab and adalimumab models, since the parameter values are almost all identical. However, the costs of adalimumab and infliximab are different and therefore the costs of the episodic and maintenance strategies also differ between the adalimumab and infliximab models. These differences mean that maintenance infliximab, if compared to standard care, generates an ICER in excess of £60k, whereas for adalimumab this ICER is less than £1k.

Table 2: Baseline Cost-effectiveness results of the Leeds models

Strategy	***Cost***	***QALY***	
Infliximab severe			
Standard Care	£13,418	0.8121	
Episodic IXB	£12,026	0.8948	
Maintenance IXB	£19,138	0.8962	
Comparison	***Incremental Cost***	***Incremental QALY***	***ICER***
Episodic vs Standard Care	£-1,392	0.0827	Dominant
Maintenance vs Standard Care	£5,720	0.0841	£68,014
Maintenance vs Episodic	£7,112	0.0014	£5,024,522
Strategy	***Cost***	***QALY***	
Adalimumab severe			
Standard Care	£13,418	0.8121	
Episodic ALB	£7,037	0.8949	
Maintenance ALB	£14,042	0.8963	
Comparison	***Incremental Cost***	***Incremental QALY***	***ICER***
Episodic vs Standard Care	£-6,381	0.0828	Dominant
Maintenance vs Standard Care	£624	0.0842	£7,411
Maintenance vs Episodic	£7,005	0.0014	£4,949,900

2.2 THE SCHERING PLOUGH MODEL

The SP model is not substantially different from the Leeds model in terms of structure. It too is a cohort, state transition model which includes similar states to the Leeds model. In particular, there is no health state for partial response. However, because of differences in parameter values used and implementation, there are substantial differences between the results. The SP model is based on analysis of the key clinical trials, and uses these data both for the standard care and treatment arms of the model. The first transition is based on the TARGAN study (induction) and then the ACCENT 1 trial is used. The model is therefore based on the trials for the first year which are extrapolated to 5 years.

Table 3 shows that episodic infliximab dominates standard care whilst the ICER for maintenance is in excess of £400,000. Whilst this ICER is high, it is substantially lower than the equivalent ICER in the Leeds model. Maintenance compared to

standard care generates an ICER of £26k, again substantially lower than the Leeds estimate.

Table 3 - Baseline Cost-effectiveness results for the Schering Plough Infliximab model

Strategy	*Cost*	*Incr Cost*	*QALYs*	*Incr QALYs*	*C/E*	*Incr C/E (ICER)*	
						Comparator	*Value*
Infliximab severe							
Episodic IXB	£25,501		2.1330		£11,956		
Standard Care	£26,209	£708	1.9586	-0.1744	£13,381	Vs Episodic	(Dominated)
						Vs Standard Care	£25,903
Maintenance IXB	£31,040	£5,539	2.1451	0.0121	£14,470	Vs Episodic	£457,769

There is a concern relating to how the SP model allocates treatment costs to patients. Drug costs are calculated for each cycle on the basis of a mean cost per year. In the case of episodic treatment, this is calculated as the cost of a dose (based on the mean weight of 60kg), plus administration, 2.2 times over a 52 week period. The cost per cycle is the annual cost divided by 6.5 (8 week cycles). This approach does not explicitly link drug costs with the disease course in the model. We would expect to see episodic costs incurred each time patients relapse, incurring a cost of between £4k and £6k depending on assumptions about vial wastage and the distribution of weight, including for every patient in the episodic arm at the start of the model. Instead, the SP model allocates a cost of £795 for every cycle in "active" disease whilst on treatment, which may generate very different estimates of cost.

2.3 THE ABBOTT MODEL

The Abbott model is a decision analytic model that estimates the cost-effectiveness of adalimumab maintenance therapy compared to standard care. There is no comparison with episodic treatment. Unlike the Schering Plough and Leeds models that use transition state model structures and therefore estimate transition probabilities, the Abbott model takes the distribution of patients at different time points over the 56 weeks direct from the CHARM clinical trial. Transitions between states are not required in this model structure.

Over the 1 year timeframe, the proportion of time spent in each of the health states is estimated and costs and utilities allocated on that basis. It is also worth noting that the

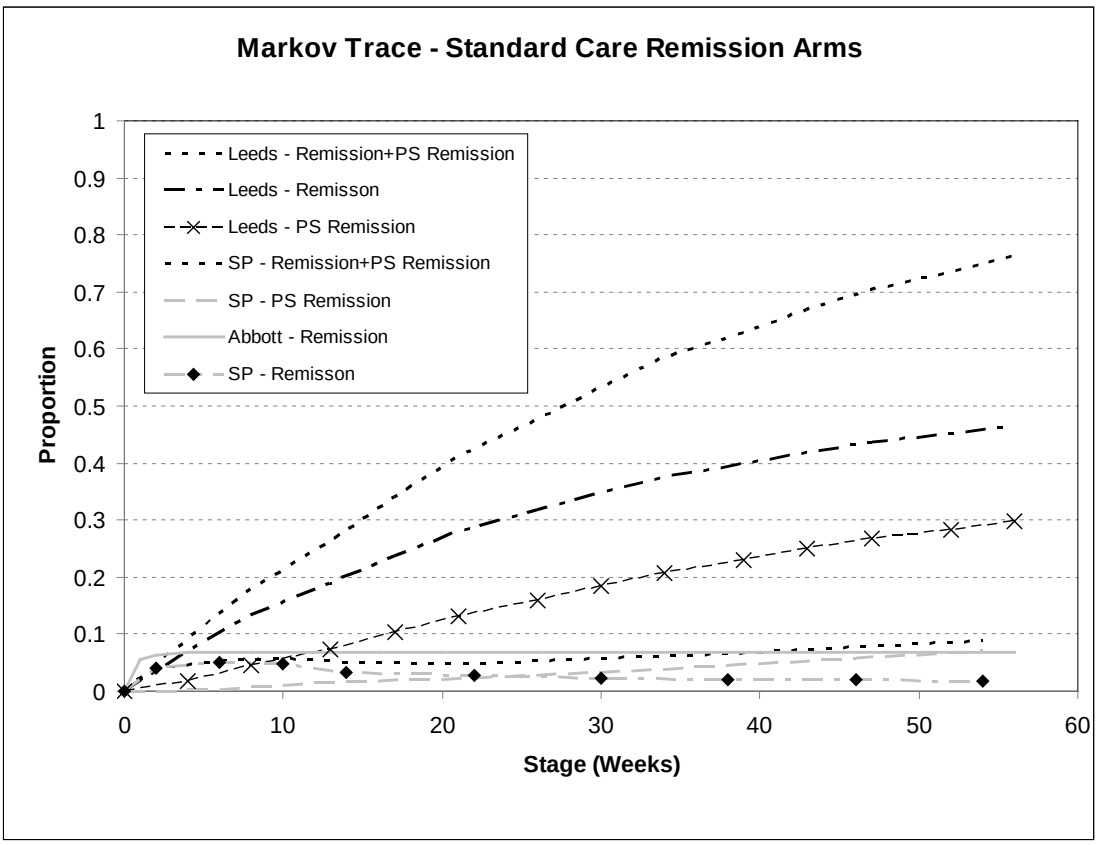
model distinguishes remission and other intermediate health states and thereby allows treatment benefits other than full remission to be reflected. Abbott extrapolated the CHARM results to provide a lifetime extended analysis. The basecase results, shown in Table 4, estimate that maintenance therapy adalimumab has an ICER of £10,959 when compared to standard care (note that these results are for severe patients only using the last value carried forward method to input missing data).

Table 4 - Baseline Cost-effectiveness results for the Abbott Adalimumab model

Strategy	***Cost***	***Incr Cost***	***QALYs***	***Incr QALYs***	***C/E***	***Incr C/E (ICER)***
Adalimumab severe						
Standard Care	£9,892		0.7339		£13,479	
Maintenance ALB	£11,182	£1,290	0.8516	0.1177	£13,131	£10,959

Overall, there is some consistency in the results between the three models to the extent that the ICERs for maintenance therapy for adalimumab and episodic therapy for infliximab are relatively low when standard care is the comparator. However, the models are substantially different in terms of their inputs, structures and other outputs (ICERs, mean costs, effects and Markov traces). For example, the standard care remission rates over 1 year, displayed in Figure 1, show how the Leeds model predicts that this proportion approaches 0.8 at 1 year and is monotonically increasing. This proportion is made up of patients in the “remission” and “post surgical remission” states. The two manufacturer models, based on the clinical trials, are consistent in that they both suggest less than 0.1 of the standard care population are in remission at 1 year. These differences are explored further in Section 4 below.

Figure 1: Comparison of Standard Care remission rates – Leeds, SP and Abbott model outputs

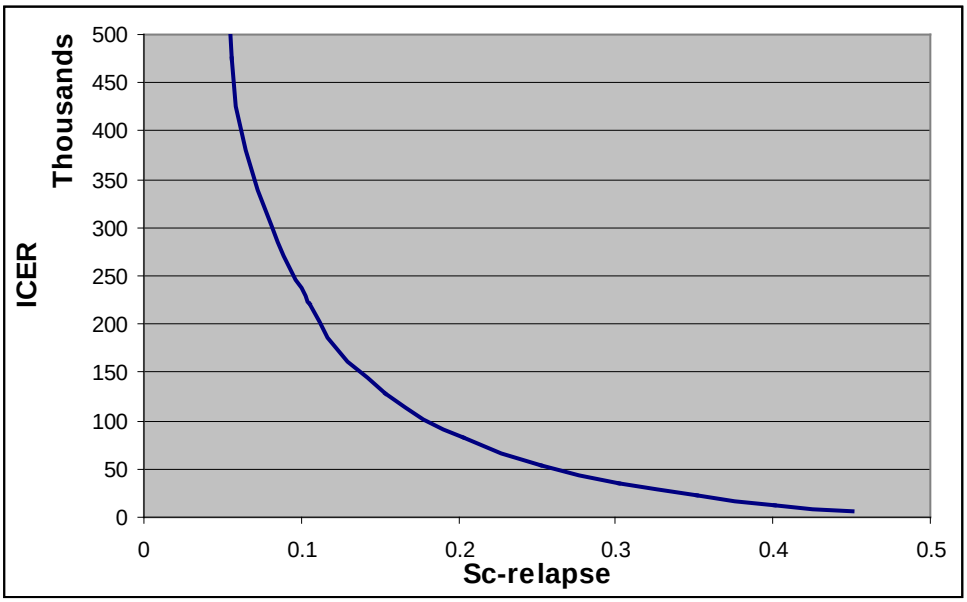


3. LEEDS MODEL – THE IMPORTANCE OF THE RELAPSE RATE

As highlighted above, patient transitions in the Leeds model are governed by probabilities derived from the Silverstein study. One such probability that is of critical importance to the model results and has been highlighted in consultation, is that which governs patients movement from remission to relapse. This parameter is of particular importance in comparisons which include episodic as well as standard care in comparisons with maintenance therapy. In the Leeds model this parameter is labelled sc_relapse and takes a fixed value of 0.0059 for each 4 week cycle, every cycle. The value is relevant both to the standard care and episodic treatment arms. The relatively low probability means that patients that are in remission are extremely unlikely to relapse and therefore episodic treatment is given only infrequently. This is also a factor in explaining the stark differences in model outputs outlined in Figure 1 above. Episodic care is therefore a substantially lower cost strategy than maintenance therapy, yet generates only slightly lower benefit.

Figure 2 below shows how the ICER is extremely sensitive to this parameter. Where the probability is 0.33 or greater, the ICER for maintenance therapy falls below £30,000 per QALY. Given the importance of this parameter and suggested analyses from consultees that this probability could be as high as 0.42 (see Abbott 7th October 2008 and previous responses), we conducted a systematic review of published evidence relating to this probability and a review of the analyses presented by Abbott.

Figure 2: Sensitivity analysis of sc-relapse versus ICER for maintenance (adalimumab)



3.1 USE OF SILVERSTEIN BASED TRANSITION PROBABILITIES AND ALTERNATIVES

Consultation raised numerous concerns about the use of the Silverstein cohort as the basis for modelling the cost effectiveness of infliximab and adalimumab. These concerns were made both in general terms and with specific reference to the standard care relapse rate. Overwhelmingly, these concerns outlined the differences between the patients included in the Silverstein study and those indicated for biologic therapy.

Silverstein et al. describe the lifetime clinical course and costs of Crohn's disease in a 24-year population-based retrospective cohort of patients with Crohn's disease in

Olmsted County, Minnesota. The study reports a general population with Crohn's disease 'who had Crohn's disease diagnosed between January 1, 1970, and December 31, 1993'. The focus of the study is not moderate to severe, refractory patients indicated for biologic therapy.

Abbott suggests two sources of evidence that conflict with the estimates derived from Silverstein et al. in relation to the relapse rate.

First, it is suggested that the CHARM placebo arm patients are a more reasonable proxy for episodic patients. Using these data, Abbott estimated the four-week probability of moving from remission to relapse as 0.4213 (42%).

Second, Abbott presented evidence from their own systematic review of evidence relating to remission rates. The Abbott review focused on the percentage of time a patient would spend in remission over a 26 week period, finding it to be 14.57% (when adjusted for sample size and duration), consistent with the results from the CHARM trial and from previous published reviews[3,4]. The predictions from the Leeds model, based on the Silverstein et al relapse rate of 0.0059 are that approaching 80% of patients are in remission at week 52 (see Figure 1).

3.2 SEARCH STRATEGY FOR POST REMISSION STANDARD CARE RELAPSE RATES

A comprehensive search was undertaken to identify literature which specified the standard care relapse rate for patients with Crohn's disease who had already achieved remission. Standard care is defined as 'concurrent therapies for Crohn's disease, including azathioprine, 6-mercaptoprurine, methotrexate, 5-aminosalicylates, sulfasalazine, oral mesalamine, CD-related antibiotics and predisolone or budesonide'. [5]

Searches were not restricted by language, publication date, or publication type. Databases searched were Medline, EMBASE, The Cochrane Library including the Cochrane Database of Systematic Reviews and Cochrane Controlled Trials Register

(CENTRAL) databases and DARE. The initial search strategies are included in Appendix 1.

3.3 INCLUSION CRITERIA

Studies were included if they were randomised control trials (or systematic reviews of randomised controlled trials) which reported relapse rate data following remission in standard care patients with moderate to severe Crohn's disease (CD). Moderate to severe Crohn's Disease was defined as a baseline Crohn's Disease Activity Index (CDAI) score of 220-450 points but papers alternatively reporting that the patients being studied had moderate to severe CD were also included. Studies were included if they gave a clear definition of the relapse rate usually based on the CDAI score for relapse, for example, Feagan et al (2000) who define relapse as a ≥100 point increase in CDAI above baseline. If this was not available, studies were included if they clearly stated how they defined relapse rate. The definitions of the relapse rate as used by the studies are included in the review. Remission is usually defined as a CDAI score of <150[6]. Studies were included if they clearly reported other definitions of remission.

Studies were included if they discussed one or more of the following three sub-groups if the other inclusion criteria were met:

- Severe active Crohn's disease
- Fistulating Crohn's disease
- Post surgical Crohn's disease patients.

These sub-groups were chosen to reflect as far as possible the moderate to severe active Crohn's Disease patient groups for which anti-TNF treatment is recommended in the UK.

3.4 DATA EXTRACTION

The main variable of interest from the studies is the standard care relapse rate following remission among those patients with moderate to severe Crohn's disease. A secondary variable of interest is the standard care median time to loss of response. Full data extraction tables are included in Table 6.

3.5 RESULTS

The search strategy identified 249 potentially relevant articles. Of these 193 articles were excluded following a title sift providing 56 for further selection. Of these 56, 21 further articles were excluded following an abstract sift. Of the remaining 35 articles, 31 were excluded after full text sifting. Four studies were identified as included articles (Table 5).

One study compared re-treatments with 10mg/kg infliximab with the effects of placebo following prior response to 5, 10 or 20/mg/kg of initial infliximab treatment. One study assessed patients who had entered remission after initial treatment with 25mg methotrexate via intramuscular injection (IM) once weekly, and then randomly assigned them to receive methotrexate 15mg IM once a week or placebo.[7] One study assessed participants for response following a 5mg/kg intravenous infusion (IV) of infliximab and then randomized responders to either placebo or 5mg/kg infliximab throughout or 5mg/kg infliximab at weeks 2 and 6 and thereafter 10mg/kg infliximab until week 46[8]. The final study by Sands et al (2004)[9] administered participants with infliximab 5mg/kg, then assessed for response, then randomly allocated responders to either placebo or 5mg/kg infliximab.

Appendix 2 and 3 briefly outline reasons for exclusion of potential studies at the full text sifting stage in Abbott documentation[10] and from WMHTAC via email communication respectively.

Table 5: 4 identified RCTs with relevant relapse rate data following remission

Study	Year	Intervention
Randomised controlled trials		
Rutgeerts et al	1999	Repeated treatments 10mg/kg infliximab vs placebo in pts showing a clinical response @week 8 to an initial infliximab treatment of 5, 10 or 20mg/kg (4 infusions@ wks 12,20,28,36.) Re-treatment began at 12 weeks
Feagan et al	2000	25mg methotrexate once weekly by IM injection for 16-24 weeks, then randomly allocated to: -Placebo (n=36) -Methotrexate (n=40)
Hanauer et al	2002	Week 0=5mg/kg IV infusion Infliximab then: Assessed for response to treatment then: Gp 1: placebo Gp 2: 5mg/kg Infliximab @ 2, 6 and every 8 weeks thereafter until week 46 Gp 3: 5mg/kg Infliximab @ 2, 6 and 10mg/kg Infliximab thereafter until week 46
Sands et al	2004	IV infusion of Infliximab 5mg/kg @ wks 0,2,6. If a responder then: Gp 1: placebo Gp 2: 5mg/kg Infliximab @ wks 14,22,30,38,46 and followed until wk 56.

Table 6 below presents further details of each study including, amongst other data: description of study populations; follow up period, disease activity definitions; and use of anti-TNFs.

3.5.1 Relapse rate

Moderate to severe active Crohn's disease

In Feagan et al a total of thirty six patients relapsed. Of these 22 of 36 were in the placebo arm (61.1%), and 14 of 36 were in the methotrexate active group (35%) with a 40 week follow up. These translate to 4 week probabilities of 0.09 and 0.048 respectively. Feagan et al used methotrexate to induce remission and methotrexate is defined as standard care. Feagan et al examined those with chronically active CD refractory to prednisolone.

Rutgeerts et al reported a median time to loss of response of 37 weeks from the placebo arm with a follow up of 48 weeks as compared with over 48 weeks in the infliximab re-treatment group. A 4 week probability of 0.072 was calculated.

Hanauer et al reported a median time to loss of response of 19 weeks from the placebo arm as compared with over 46 weeks in the infliximab re-treatment group with a follow up to 54 weeks. A 4–week probability of 0.136 was calculated.

Fistulising active Crohn's disease

Sands et al (2004) examined fistulizing CD patients and reported that 61 patients (62%) in the placebo arm had a loss of response (relapse) assessed at week 54 compared with 40 patients (42%) in the infliximab maintenance group. The median time to loss of response from the placebo arm was 14 weeks, which leads to an estimated 4 week transition probability of 0.180. In this study, remission was induced by infliximab therapy.

Post surgical

No RCTS were found with direct relapse rate data were identified after application of the inclusion criteria.

Table 6: Full data extraction tables for studies meeting inclusion criteria

Study	Type of study	Study sample*	Interventions	Responders, Remission and relapse definitions	Relapse rate post remission (% of patients) with time points and/or median time to loss of response data.	Use of anti-TNFs
Hanauer et al 2002	Multicentre randomized double blind trial North America, 55 sites.	CD of at least 3 months duration. From 573 pts assessed for response, 335 responders randomly assigned to 3 groups. Gp 1 n110 Gp 2=n113 Gp 3=n112 Median age 35 yrs (18-76yrs) CDAI entry score 220 to 400 prior to treatment to obtain response	Week 0=5mg/kg IV infusion Infliximab then: Assessed for response to treatment then: Gp 1: placebo Gp 2: 5mg/kg Infliximab @ 2, 6 and every 8 weeks thereafter until week 46 Gp 3: 5mg/kg Infliximab @ 2, 6 and 10mg/kg Infliximab thereafter until week 46	Responders: Decrease in CDAI score of 70 points or more from baseline and at least 25% reduction in score (some patients would be in remission but % not clearly reported)	Median time to loss of response was 46 weeks (IQR 17 to >54) in groups 2 and 3. Gp 1 placebo= 19 weeks (10-45) No clear relapse rate reported.	No. Excluded from study for previous infliximab treatment or other TNF targeting agents.

IV=Intravenous
CD=Crohn's Disease
FU=follow up.

Study	Type of study	Study sample*	Interventions	Responders, Remission and relapse definitions	Relapse rate post remission (% of patients) with time points and/or median time to loss of response data.	Use of anti-TNFs
Sands et al 2004	Multicentre randomized double blind placebo-controlled trial. North America, Europe, Israel; 45 sites.	Fistulizing CD From 306 pts, 195 randomized as responders at week 14. Placebo=n99 Infliximab maintenance n96. 48% complete response before randomization 147of 306) at week 14. Median age of responder then placebo: 36 (IQR 29-46) Median age of responder then active drug: 37 (IQR 28-47)	All=IV infusion of Infliximab 5mg/kg @ wks 0,2,6. If a responder then: Gp 1: placebo Gp 2: 5mg/kg Infliximab @ wks 14, 22, 30, 38, 46 and followed until wk 56.	Responders: Reduction of 50% from baseline in number of draining fistulas at weeks 10 and 14. Relapse termed loss of response: recrudescence of draining fistulas, need for change in medication for CD, need for additional therapy for persistent or worsening disease activity, need for surgery for CD or discontinuing medication for perceived lack of efficacy. Remission termed as complete response=absence of draining fistulas.	Median time to loss of response was weeks (IQR 17 to >54) in groups 2 and 3. Gp 1 placebo= 14weeks (10-45) Infliximab maintenance=>40 weeks. Relapse (loss of response): Infliximab maintenance: 42% Placebo group: 62% Time points not clearly reported .	Yes, see interventions cell.

IV=Intravenous
CD=Crohn's Disease
FU=follow up.

Study	Type of study	Study sample*	Interventions	Responders, Remission and relapse definitions	Relapse rate post remission (% of patients) with time points and/or median time to loss of response data.	Use of anti-TNFs
Feagan et al 2000	Multicentre randomized placebo controlled trial North America: 7 sites	Chronically active disease with at least three months of symptoms despite daily doses of at least 12.5 mg of prednisone with at least one attempt to discontinue treatment. Total of 76 patients Median age Methotrexate gp=32 (2) Placebo=34(2)	76 in remission after 25mg methotrexate once weekly by IM injection for 16-24 weeks, then randomly allocated to: -Placebo (n=36) -Methotrexate (n=40) FU of 40wks.	Remission: absence of need for Prednisolone therapy. + CDAi score ≤150. Relapse=≥100 point increase in CDAI above baseline or initiation of either/both prednisolone and antimetabolite.	36 pts relapsed (22 placebo, 14 methotrexate gp) Inferring: 22 of 36 placebo (61.1%) 14 of 40 methotrexate gp (35%)	Not mentioned, presumably excluded from study.

Study	Type of study	Study sample*	Interventions	Responders, Remission and relapse definitions	Relapse rate post remission (% of patients) with time points and/or median time to loss of response data.	Use of anti-TNFs
Rutgeerts et al 1999	Randomized double blind placebo controlled parallel group trial North America, Europe: sites not clearly reported	Active CD Refractory to standard therapy Median age=35 (20-65) Sample=73 responders: N=37 infliximab arm N=36 placebo arm Reduced to n=69 receiving initial treatment.	Repeated treatments 10mg/kg infliximab vs placebo in pts showing a clinical response @week 8 to an initial infliximab treatment of 5, 10 or 20mg/kg (4 infusions@ wks 12, 20, 28, 36). Re-treatment began at 12 weeks FU over 36 weeks (wks 12-48)	Baseline CDAi used to determine clinical response and remission at week 0. Clinical response=≥70 point decrease in CDAI from baseline Remission= CDAI <150 points 38% in remission at week 12 when re-treatment began.	Median time to loss of response: Infliximab re-retreatment arm=>48 weeks Placebo arm=37 weeks No clear relapse rate reported.	Yes, patients had to show a response to an initial infliximab treatment.

3.6 SUMMARY OF IDENTIFIED TRIAL DATA

The available data directly reporting the relapse rate are minimal but suggest 4 week probabilities ranging from 7 to 14% for moderate to severe CD may be typical. This contrasts with the 0.59% estimate used in the Leeds base case model derived from the Silverstein cohort. The Leeds model suggests that the median time to relapse is approximately 9 years, whereas these sources suggest median estimates of 0.5 years.

These estimates are substantially lower than those proposed by Abbott based on their analysis of the CHARM data. In part this may be due to the patient level analysis conducted by Abbott.

It is important to note that these differences in estimates of the relapse rate do not necessarily result in different conclusions regarding cost effectiveness from the Leeds model when other transition probabilities are also considered. The review data discussed above suggest that there are substantial differences between the predicted proportion of patients in remission from the Leeds model compared to the literature. It is not solely the sc_relapse transition probability that determines the proportion of patients in remission in the Leeds model. Therefore, whilst using a very high relapse rate (0.42) as in the CHARM data results in relatively low ICERs for biologic maintenance therapy, similar results could also be obtained using lower relapse rates (such as those identified in our review) but with alterations to other parameters in the model. This is further explored below.

3.7 COST EFFECTIVENESS USING CHARM RELAPSE RATE

Table 7: Cost effectiveness using Leeds model and CHARM relapse rate - adalimumab

	Cost (£)	***QALY***	
Standard Care	16,836	0.7738	
Episodic	15,060	0.8264	
Maintenance	15,566	0.8786	
Comparison	***Incremental Cost***	***Incremental QALY***	***ICER***
Episodic vs Standard Care	-1,775	0.053	Dominant
Maintenance vs Standard Care	-1,270	0.105	Dominant
Maintenance vs Episodic	506	0.052	£9,687

Table 7 shows the result of using the relapse rate of 0.42 in the Leeds model for adalimumab. Whilst standard care is dominated by both episodic and maintenance therapy, the ICER for maintenance compared to episodic is £9687.

Table 8: Cost effectiveness using Leeds model and CHARM relapse rate - infliximab

	Cost (£)	***QALY***	
Standard Care	16,832	0.7738	
Episodic	28,281	0.8264	
Maintenance	21,451	0.8786	
Comparison	***Incremental Cost***	***Incremental QALY***	***ICER***
Episodic vs Standard Care	11,450	0.053	217,672
Maintenance vs Standard Care	4,620	0.105	44,079
Maintenance vs Episodic	-6,830	0.052	Dominant

Table 8 shows the results of using the Leeds model for infliximab with the 0.42 relapse rate. The ICER for episodic versus standard care exceeds £200k, whilst the ICER for maintenance versus standard care is £44k.

Figure 3: Standard care Markov trace – amended Leeds model using 0.42 sc-relapse

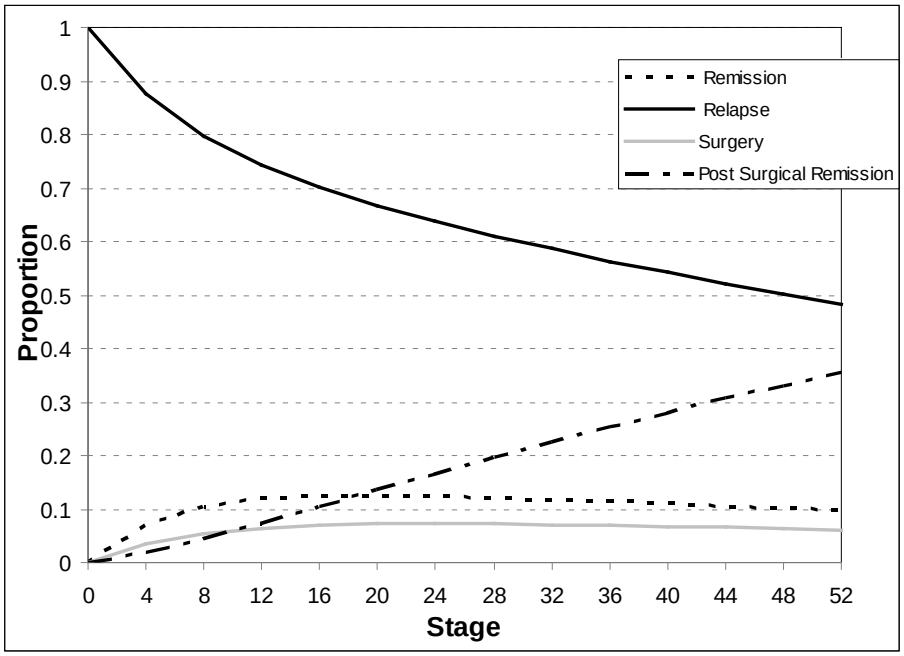


Figure 3 shows the model output for standard care using the amended Leeds model and the 42% relapse rate. The model output is now more consistent with the proportions in the Abbott and SP models (see Figure 1) being around 10% at year 1. However, it should be noted that other elements of the model remain inconsistent, for example, the post surgical remission rate remains high.

4. RECONCILIATION BETWEEN MANUFACTURER AND LEEDS MODELS

We identified differences between the manufacturer and Leeds models and made changes to the Leeds model in a stepwise fashion. The purpose of this is to attempt to reconcile the different modelling approaches, rather than an attempt to present a set of ICER estimates that are considered preferable or plausible. It should be noted that small inconsistencies between results presented in this section and those reported elsewhere (including the Assessment Group reports) are down to rounding errors. It should also be noted that we only consider the deterministic version of the model in this report. Probabilistic analysis in the Leeds model is not implemented in a manner that is useful for decision making being both partial and inappropriate.

4.1 SP AND LEEDS MODELS

4.1.1 Base case

Table 9 and Table 10 show the base case analyses for the two models. Incremental changes to the Leeds model are described in the following sections.

Table 9 - Leeds Model Basecase

	Cost	***QALY***	
Standard Care	£13,418	0.8121	
Episodic	£12,026	0.8948	
Maintenance	£19,138	0.8962	
Comparison	***Incremental Cost***	***Incremental QALY***	***ICER***
Episodic vs Standard Care	£-1,392	0.083	**Dominant**
Maintenance vs Standard Care	£5,720	0.084	**£68,014**
Maintenance vs Episodic	£7,112	0.001	**£5,080,000**

Table 10 - Schering Plough Basecase

	Cost	***QALY***	
Standard Care	£26,209	1.958601	
Episodic	£25,501	2.132989	
Maintenance	£31,040	2.145098	
Comparison	***Incremental Cost***	***Incremental QALY***	***ICER***
Episodic vs Standard Care	£-708	0.174	**Dominant**
Maintenance vs Standard Care	£4,831	0.186	**£25,903**
Maintenance vs Episodic	£5,539	0.012	**£457,386**

We aimed to explain differences in model results due to a) cost, utility and discount parameter differences, b) structural issues such as time horizon and the patient pathway, c) transition probabilities that could be quantified, d) residual transition probabilities that could not be quantified. These changes are discussed below and the results of each step reported in Table 11.

Since the modelled health states are not equivalent, it is not entirely feasible to substitute transition probabilities from one model to another without restructuring the

health states being modelled. For example, in the SP model, once patients have “failed” treatment and entered the “non-responding active” state, no further distinction is made between health state, other than those relating to surgery and death. The Leeds model replicates all health states for patients on and off treatment. At the extreme, reconciliation would result in rebuilding the manufacturer model in a different piece of software.

Table 11: Results from changes made to Leeds severe, infliximab model

Change made		Treatment	Cost (£)	QALY	Cost Effectiveness			
					Comparison	Incr Cost (£)	Incr QALY	ICER (£/QALY)
1	Time Horizon	Standard Care	24,925	4.2345	Episodic vs SC	-3,492	0.117	Dominant
		Episodic	21,433	4.3512	Maintenance vs SC	40,449	0.126	320,261
		Maintenance	65,374	4.3608	Maintenance vs Episodic	43,941	0.010	4,577,188
2	Utility Values – relapse and remission	Standard Care	24,925	3.6455	Episodic vs SC	-3,492	0.140	Dominant
		Episodic	21,433	3.7858	Maintenance vs SC	40,449	0.152	266,462
		Maintenance	65,374	3.7973	Maintenance vs Episodic	43,941	0.011	3,820,957
2a	*Utility values – post surgical remission*	*Standard Care*	*24,925*	*3.3649*	*Episodic vs SC*	*-3,492*	*0.230*	*Dominant*
		Episodic	*21,433*	*3.5946*	*Maintenance vs SC*	*40,449*	*0.243*	*166,320*
		Maintenance	*65,374*	*3.6081*	*Maintenance vs Episodic*	*43,941*	*0.014*	*3,254,889*
2b	*Utility values – from Buxton et al.*[11]	*Standard Care*	*24,925*	*3.574*	*Episodic vs SC*	*-3,492*	*0.190*	*Dominant*
		Episodic	*21,433*	*3.7641*	*Maintenance vs SC*	*40,449*	*0.206*	*196,736*
		Maintenance	*65,374*	*3.7796*	*Maintenance vs Episodic*	*43,941*	*0.015*	*2,834,903*
3	Discount Rate	Standard Care	23,816	3.3607	Episodic vs SC	-3,281	0.136	Dominant
		Episodic	20,535	3.497	Maintenance vs SC	37,131	0.147	252,936
		Maintenance	60,947	3.5075	Maintenance vs Episodic	40,412	0.011	3,848,762
4	Cost of Infliximab	Standard Care	23,816	3.3607	Episodic vs SC	-5,412	0.136	Dominant
		Episodic	18,404	3.497	Maintenance vs SC	23,051	0.147	157,023
		Maintenance	46,867	3.5075	Maintenance vs Episodic	28,463	0.011	2,710,762
5	Transition probabilities – Relapse to Remission	Standard Care	21,944	3.3853	Episodic vs SC	1,614	0.038	42,251
		Episodic	23,558	3.4235	Maintenance vs SC	19,774	0.069	288,251
		Maintenance	41,718	3.4539	Maintenance vs Episodic	18,160	0.030	597,368
6	Transition probabilities – Remission to Relapse	Standard Care	38,531	3.1675	Episodic vs SC	6,190	0.009	719,767
		Episodic	44,721	3.1761	Maintenance vs SC	14,011	0.056	250,644
		Maintenance	52,542	3.2234	Maintenance vs Episodic	7,821	0.047	165,349
7	Episodic Therapy Remission to Relapse Rate	Standard Care	38,531	3.1675	Episodic vs SC	5,521	0.016	340,802
		Episodic	44,052	3.1837	Maintenance vs SC	14,011	0.056	250,644
		Maintenance	52,542	3.2234	Maintenance vs Episodic	8,490	0.040	213,854
8	Health State Costs	Standard Care	47,949	3.1675	Episodic vs SC	3,232	0.016	199,506
		Episodic	51,181	3.1837	Maintenance vs SC	11,840	0.056	211,807
		Maintenance	59,789	3.2234	Maintenance vs Episodic	8,608	0.040	216,826

9	Surgery Probabilities	Standard Care Episodic Maintenance	43,888 48,995 52,318	3.0124 3.0383 3.1228	Episodic vs SC Maintenance vs SC Maintenance vs Episodic	5,107 8,430 3,323	0.026 0.110 0.084	197,181 76,359 39,325
10	Routing of maintenance therapy after surgery	Standard Care Episodic Maintenance	43,888 48,995 50,839	3.0124 3.0383 3.1046	Episodic vs SC Maintenance vs SC Maintenance vs Episodic	5,107 6,951 1,844	0.026 0.092 0.066	197,181 75,390 27,813
11	Routing of episodic therapy after surgery	Standard Care Episodic Maintenance	43,888 48,790 50,839	3.0124 3.0372 3.1046	Episodic vs SC Maintenance vs SC Maintenance vs Episodic	4,902 6,951 2,049	0.025 0.092 0.067	197,661 75,390 30,401
12	Post-surgery remission rates	Standard Care Episodic Maintenance	52,773 57,319 56,940	2.8752 2.9075 3.0206	Episodic vs SC Maintenance vs SC Maintenance vs Episodic	4,546 4,167 -379	0.032 0.145 0.113	140,743 28,659 Dominant
12	*Also incorporating 2a*	*Standard Care* *Episodic* *Maintenance*	*52,773* *57,319* *56,940*	*2.7868* *2.8254* *2.9111*	*Episodic vs SC* *Maintenance vs SC* *Maintenance vs Episodic*	*4,546* *4,167* *-379*	*0.039* *0.124* *0.086*	*117,772* *33,524* *Dominant*

Note: Changes are cumulative, excluding those in italics

4.1.2 Change 1 – Time Horizon

A clear difference between the two models is that the Leeds model only has a 1 year time horizon, whereas the Schering Plough base case model is run for 5 years. Implementing the same time horizon in the Leeds model results in an ICER for maintenance care compared to standard care that deteriorates, whilst the ICER compared to episodic care improves slightly to £4.6m. It is interesting to note that whilst the costs of episodic and standard care are closer to those in the SP model, maintenance care is more than double the estimated SP cost. The average QALYs for each strategy are also vastly higher than for the SP model. The mean annual utility for a standard care patient is 0.846 in the Leeds model, compared to 0.39 in the SP model (although it should be noted that the SP values were discounted).

4.1.3 Change 2 – Utility Values for relapse and remission

The second change was to use the utility values used in the Schering Plough model for remission and "Non –responding active" in place of the Leeds "remission" and "relapse" utilities. These values make a greater distinction in utility values between remission and relapse - 0.28 compared to 0.22. The results of implementing this change are shown in Table 11. The incremental benefits of both episodic and maintenance therapy are improved and consequently the ICER for maintenance vs standard care also improves. Only marginal changes in maintenance vs episodic ICERs occur from this change.

Change 2a – Incorporating utility values for post surgical remission

An additional difference between the models relates to post surgical remission. The Leeds model assumes that the utility for this state is equivalent to the remission state utility. SP assign a value that is 0.16 lower for the post surgical state. This results in the incremental benefits of each of the treatment strategies increasing, thereby lowering the ICERs.

Change 2b – Incorporating utility values from Buxton et al..

As outlined in our previous report, a potential source of utility values, based on a simple regression function between CDAI score and EQ5D results in larger

differences between the remission and relapse states. The results from substituting these values into the Leeds model are reported in Table 11.

4.1.4 Change 3 – Discount Rate

The base case Leeds analysis does not include discounting of either costs or benefits since it operated over a 1 year time horizon. Having extended the time horizon it is also appropriate to include discounting at the recommended 3.5% rate. Results are shown in the row numbered 3 in Table 11. Costs and benefits for all strategies are reduced as expected. The ICER for maintenance versus standard care reduces slightly but compared to episodic care the ICER rises slightly.

4.1.5 Change 4 – Cost of Infliximab

The cost of infliximab was calculated by Leeds to be much higher than the SP estimate. These estimates differed due to assumptions about the mean weight of patients, the possibility for vial sharing, and administration costs. To implement the SP estimates in the Leeds model we reduced the cost of induction from £5809 to £4066, the cost of maintenance from £3872 to £2710 in the first cycle, and from £968 to £678 in the subsequent cycles. Reductions in the costs of episodic and maintenance strategies lower the ICERs for maintenance to £2.7m compared to episodic care.

4.1.6 Change 5 – Transition Probabilities for Relapse to Remission

As outlined above, the reliance on constant transition probabilities from the Silverstein cohort and single arms of trials is a source of considerable difference in approach to the manufacturer models. The Leeds model used a constant probability of 0.0713 for relapse to remission standard care and 0.56 for both episodic and maintenance therapy. This is a contributory factor for the difference in the numbers of standard care patients in remission (see Section 2, Figure 1) and is also a contributory factor to the differences between the models in terms of treated patients (episodic and maintenance) in remission. Figure 4 and Figure 5 below show further differences between the models. In the Leeds model, the use of a relatively high, constant response rate (0.56) means that over 75% of patients are in the remission state after 2 cycles. The numbers of patients in either “remission” or “post surgical remission” is around 80% over the entire 5 years, for both episodic and maintenance therapy arms. The Leeds model generates ICERs for maintenance therapy vs standard care that

increase over time because approaching 80% of standard care patients are predicted to be in remission after 5 years. In effect, the model suggests that the benefit of treatment is largely in moving patients to remission more quickly but that they will move to remission in the absence of anti-TNF therapy over time. The SP model predicts much lower remission rates for both maintenance and episodic therapy patients.

Figure 4: Remission rates for Leeds and SP models – Maintenance therapy

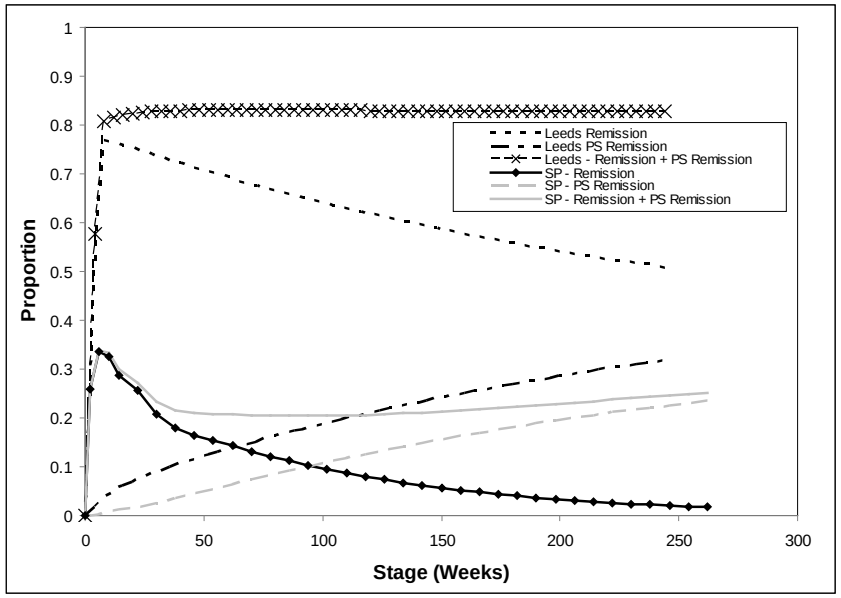
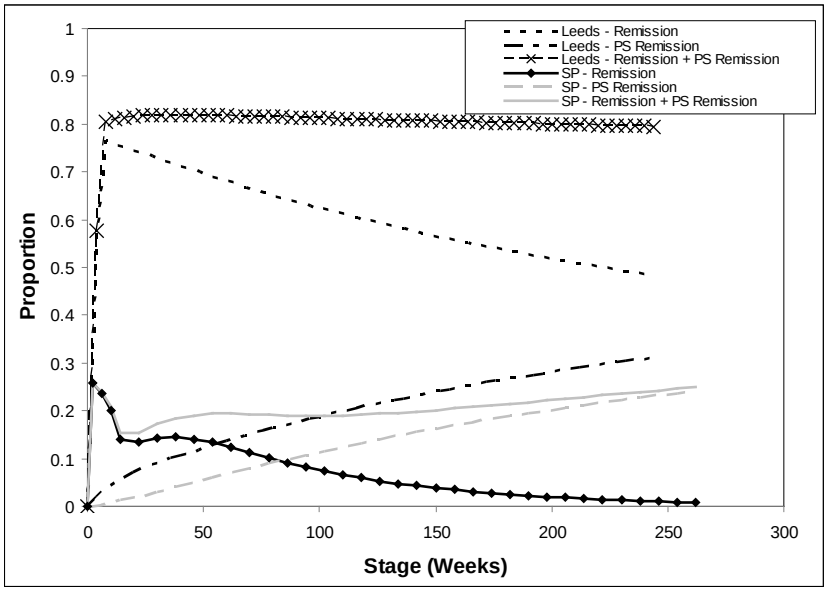


Figure 5: Remission rates for Leeds and SP models – Episodic therapy



We took the trial based transition probability matrix from the SP model and adjusted the Leeds model to use these probabilities. Since the appropriate transition depends on the time previous spent in a state, rather than time from the start of the model, we created tunnel states within TreeAge Pro to achieve this. Some caution is warranted in considering the results of this approach and all other changes to transition probabilities. Since the source of the data, the permitted transitions, the precise definitions of health states and cycle length all differ to some degree between the two models, these analyses suffer some limitations. They are conducted with the intention of helping to reconcile the differences between the two models.

The effect of introducing the SP data for relapse to remission rates sees both episodic and maintenance therapy worsen compared to standard care since the relatively large treatment effect (0.55) is no longer applied every cycle. The trial data reflect a much lower response rate over time for patients on treatment.

The ICER for maintenance versus episodic care improves to £0.6m.

4.1.7 Change 6 – Transition Probabilities for Remission to Relapse

As with Change 5, tunnel states were introduced so that the trial based transition probabilities could be used to inform the probabilities of patients moving from remission to relapse. These estimates replace the 0.0059 constant probability applied from the Silverstein study to the episodic and standard care arms, and the 0.00076 constant probability applied to the maintenance arm.

This change has a substantial effect on the ICERs. In particular, the ICER for episodic therapy rises to over £0.7m. The ICERs for maintenance care vs standard care and episodic care both fall. The results highlight how sensitive the Leeds models are to changes in their effectiveness parameters.

4.1.8 Change 7 – Episodic Therapy Remission to Relapse Rate

The changes to the transition probabilities implemented in change 6 for patients moving from remission to relapse differ between the episodic and standard care arms of the model. An alternative is to use the placebo data for both episodic and standard

care treatment arms. This slightly lowers the cost and increases the QALYs for the episodic care arm, thereby improving the ICER for episodic versus standard care and worsening the maintenance versus episodic care ICER.

Figure 6: Comparison of SP and Revised Leeds model – Episodic

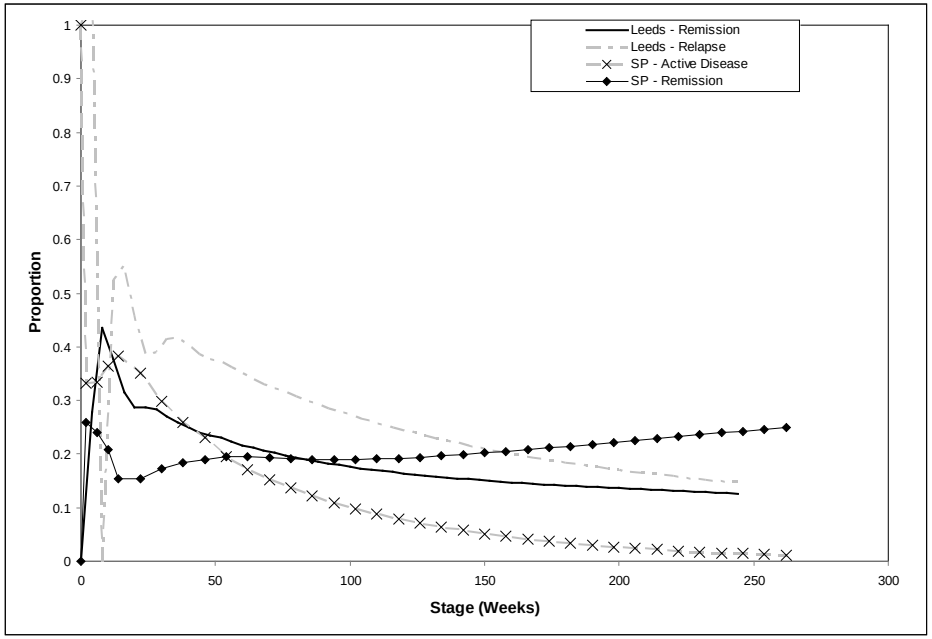


Figure 7: Comparison of SP and Leeds models – Maintenance

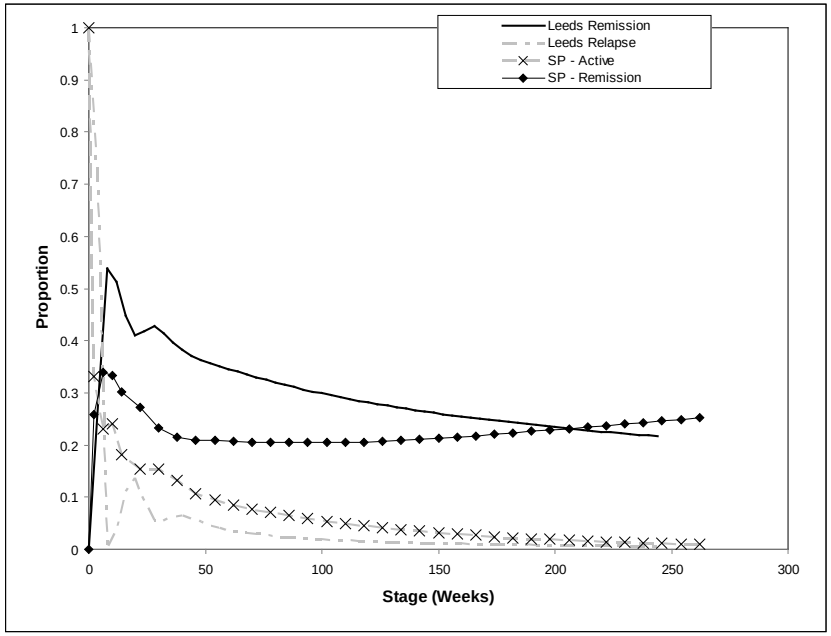


Figure 6 and Figure 7 demonstrate the impact that changing the remission and relapse transition probabilities have had in aligning the Leeds model to the SP model to some extent. It can be seen that the models have a substantial degree of similarity in these outputs, much more so than the base case models. Differences remain and are driven by transitions to and from other states than those governed by the probabilities we have been able to adjust.

4.1.9 Change 8 – Health State Costs

The health state costs included in both models contain drug administration costs, as well as any hospital attendances and resource use costs. The Leeds model estimated the Health State Costs at £52 (Remission), £1489 (Severe Relapse), £475 (Moderate Relapse), £4592 (Surgery) and £72 (Surgical Remission). The Leeds model was updated with the Schering Plough estimates for these health state costs. These revised estimates are £354 (Remission, Severe Relapse and Moderate Relapse), £5277 (Major Surgery), £420 (Surgical Remission) and £922 for a new parameter for the cost per cycle whilst on placebo therapy.

Table 11 shows that the costs of all strategies rise but proportionally less for episodic and maintenance therapy. ICERs are therefore marginally improved when infliximab is compared to standard care. The change in the ICER comparing maintenance versus episodic is negligible.

4.1.10 Change 9 – Surgery Probability

The Schering Plough model has lower estimates in general for the probability of moving to the surgery state, which in turn influences the proportion of patients in post surgical states. These transition probabilities were substituted into the Leeds model which results in lower costs and QALYs for all three treatment strategies. The result of this change sees a slight improvement in the ICER for episodic care vs standard care. There is a significant improvement in the ICERs for maintenance vs both standard care and episodic therapy, which is driven by greater incremental QALYs with the revised surgery rates.

4.1.11 Change 10 – Routing of maintenance therapy patients after surgery

An assumption made in the SP model is that post surgery, maintenance therapy ceases. In contrast, the Leeds model stops maintenance therapy for any cycle whilst patients are in the surgical state, with maintenance restated when moving to post surgical remission. To align the models, this has been changed so that patients after surgery on maintenance therapy move to standard care. This results in a small reduction in mean maintenance cost and QALYs because patients no longer pass through the "transitional" state, where additional maintenance costs were incurred and utility at a higher rate to surgery was generated.

The movement of patients after surgery to standard care sees a significant fall in the cost of maintenance therapy relative to the fall in QALYs, resulting in an improvement of the ICERs when compared to standard care (£75k) and episodic care (£28k).

4.1.12 Change 11 – Routing of patients after episodic therapy surgery

In line with Change 10, patients who are on episodic therapy are now routed to standard care after having surgery, rather than continuing in the "on treatment" component of the model.

This change has very little effect on the resulting ICERs partially since the episodic "transitional" state did not entail additional drug costs in this arm of the mode, unlike the maintenance arm. Episodic care is slightly less cost effective compared to standard care, but improves against maintenance care, due to a slight narrowing of the incremental QALYs between the two.

4.1.13 Change 12– Post-surgery Remission Rates

The SP model has a substantially higher probability of transiting from surgery to post surgical remission (85.6%) compared to the Leeds model (26.1%). In the Leeds model, most patients undergoing surgery remain in the same state (and therefore incur the high costs of surgery) in the following cycle (67.1%). We substituted the SP probability into the Leeds model.

This change has a substantial impact on the ICERs. For episodic care vs standard care the ICER increases from £108k to £141k. This is driven by a large increase in the incremental costs of episodic care. On the other hand, maintenance therapy is now a dominant strategy over episodic care, due to being less costly, and the ICER compared to standard care has now fallen from £75k to £29k.

Change 12a – including SP post surgical remission utility

Adding the previous changes to that made in 2a reduces the incremental QALY gain of maintenance compared to standard care, thereby increasing the ICER to £34k.

Figure 8 - Updated Leeds model vs SP: SC remission

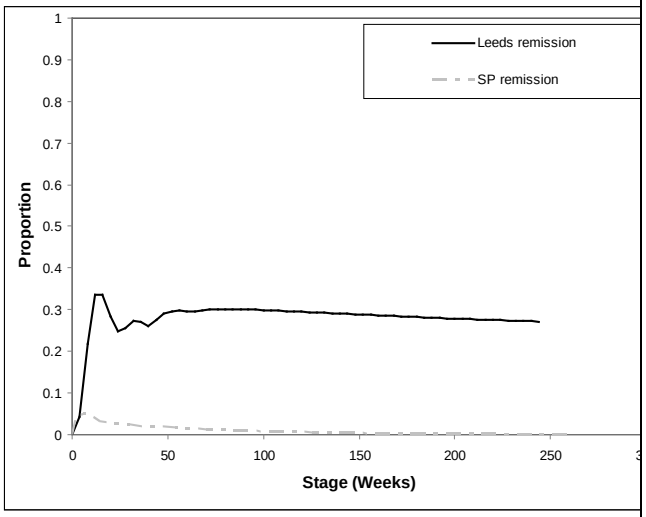


Figure 9 - Remission and PS remission combined, SC

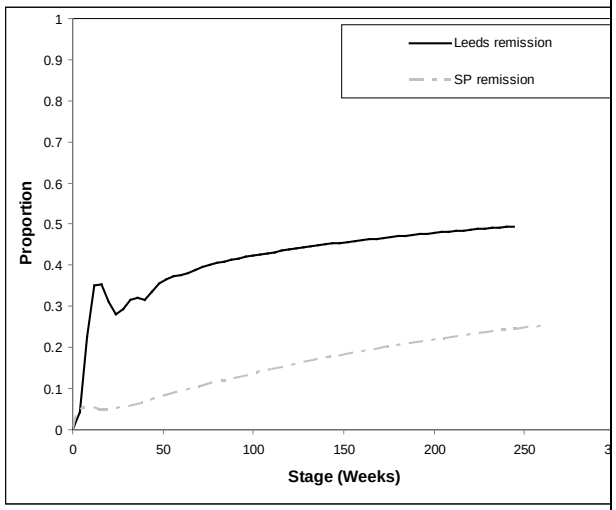
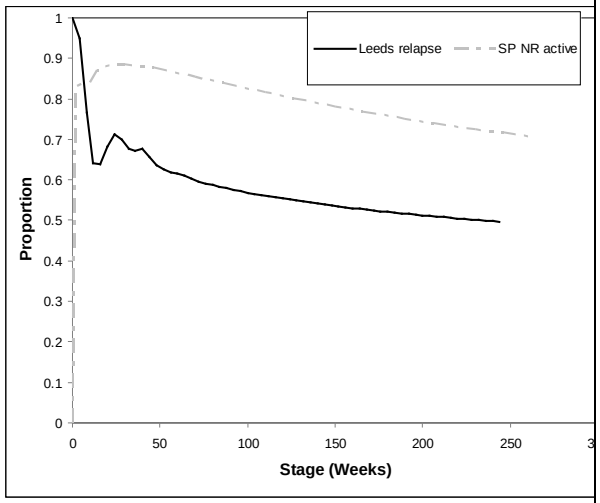


Figure 10 – Relapse, SC



For illustration, we have also calculated the ICERs for infliximab using all the cumulative changes but excluding those that use the SP approach for infliximab drug costs. Table 12 and Table 13 show the new estimated ICERs which rise both for episodic and maintenance. Maintenance versus standard care rises to £59k in version 12 and £69k in version 12a.

Table 12: Amended Leeds model (change 12) with Leeds infliximab drug costs

Strategy	***Cost***	***QALY***	
Standard Care	£52,773	2.8752	
Episodic IXB	£60,230	2.9075	
Maintenance IXB	£61,405	3.0206	
Comparison	***Incremental Cost***	***Incremental QALY***	***ICER***
Episodic vs Standard Care	£7,457	0.032	£230,867
Maintenance vs Standard Care	£8,632	0.145	£59,367
Maintenance vs Episodic	£1,175	0.113	£10,389

Table 13: Amended Leeds model (change 12a) with Leeds infliximab drug costs

Strategy	***Cost***	***QALY***	
Standard Care	£52,773	2.7868	
Episodic ALB	£60,230	2.8254	
Maintenance ALB	£61,405	2.9111	
Comparison	***Incremental Cost***	***Incremental QALY***	***ICER***
Episodic vs Standard Care	£7,457	0.039	£193,187
Maintenance vs Standard Care	£8,632	0.124	£69,445
Maintenance vs Episodic	£1,175	0.086	£13,711

4.1.14 Summary of reconciliation between SP and Leeds models

We have implemented changes to key cost and utility parameters, made structural changes to the model to incorporate aspects consistent with the SP model, and adjusted some of the key sets of transition probabilities.

These changes have aligned the models to a degree yet substantial differences in the outputs of the models remain. Examination of the Markov traces for the amended version of the Leeds model and the SP model highlight the influence of the remaining transitions based on the Silverstein cohort compared to those based on the SP trial data, and the permitted transitions between states, differences which in turn drive the

estimates of cost and effect. For example, whilst the relapse and remission rates for the episodic and maintenance arms are much closer (see Figure 6 and Figure 7), differences remain in the standard care arm. Figure 8, Figure 9 and Figure 10 illustrate that the standard care remission rate in the revised Leeds model remains higher than the SP model predicts. Without changing the health states and transitions that the Leeds model uses, effectively rebuilding the SP model in TreeAge Pro, it is not possible to reconcile every element of the Markov process.

However, it is clear that the SP model, which does faithfully replicate the relevant clinical trials as far as we can tell, produces results that differ from the Leeds model because the Leeds model is based on transition probabilities derived from a quite different source.

4.2 USING THE SP/LEEDS ANALYSES TO CONSIDER ADALIMUMAB

Whilst making changes to the Leeds model to incorporate the Schering Plough structure and parameter values, the revised model after each change was also run with the adalimumab drug costs, for completeness. The original Leeds submissions generated near identical estimates of effectiveness between infliximab and adalimumab. The results alongside each change are given in Table 14. The final result shows consistency with the Abbott result, in that maintenance therapy generates a relatively low ICER. However, there remain substantial differences between the Abbott and Leeds models, as was highlighted above in relation to the SP model.

Table 14: Results of updated Leeds model with adalimumab costs

	ICER		
Change	**Episodic vs Standard Care**	**Maintenance vs Standard Care**	**Maintenance vs Episodic**
Leeds model starting point	Dominant	£7,037	£5,003,571
1.	Dominant	£199,430	£3,599,167
2.	Dominant	£165,929	£3,004,522
3.	Dominant	£155,606	£3,039,333
4.	Dominant	£155,606	£3,039,333
5	n/a	n/a	n/a
6.	133,372	£189,589	£199,810
7.	33,580	£189,589	£253,249
8.	Dominant	£150,751	£256,222

9.	232	£39,293	£51,266
10.	232	£35,879	£49,804
11.	232	£35,879	£48,947
12.	Dominant	£2,689	£7,445
Abbott Basecase Results	-	£11,998	-

4.3 SUBSTITUTING ABBOTT AND LEEDS PARAMETER VALUES

Table 15: Leeds adalimumab model output with Abbott utilities and health state costs

Strategy	***Cost***	***Incr Cost***	***QALYs***	***Incr QALYs***	***ICER***	
Adalimumab severe						
Standard Care	£4,560	£-858	0.7551			
Episodic ALB	£3,702	£6,244	0.8178	0.0627	Vs Episodic	Dominant
Maintenance ALB	£10,804	£6244 £7,101	0.8188	0.0637 0.0010	Vs Standard Care Vs Episodic	£98,019 £7,101,390

Table 15 shows the results of substituting the parameter values from the Abbott model that are easily used given the differences in structure between the two approaches, namely the utility values and the health state costs.

The results, compared to the Leeds base case model shown in

Table 2, show that the incremental cost of maintenance compared to standard care is much greater than in the original Leeds model, leading to a rise in the ICER from £7k to £98k.

Table 16: Results from Abbott model with Leeds costs and utilities

Strategy	***Cost***	***Incr Cost***	***QALYs***	***Incr QALYs***	***ICER***
Adalimumab severe					
Standard Care	£16,725		0.8351		
Maintenance ALB	£14,854	£-1,871	0.9512	0.116	Dominant

Table 16 shows the results obtained from the Abbott model with the Leeds health state costs and utility values. In this situation, maintenance adalimumab dominates standard care.

These results provide further support for the finding that it is the source of the patient transitions i.e. the Silverstein based transitions compared to the clinical trial based analyses, that are the drivers of the differences in model results.

5. CONCLUSIONS

At the heart of the differences between the manufacturer and Leeds models is the source of the data used to estimate the distribution of patients between various health states. Both manufacturers present models that draw substantially on the clinical trials for infliximab and adalimumab. The Leeds model relies almost exclusively on data derived from a cohort of patients that may be substantially different from those indicated for anti-TNF therapy. The model then uses a single probability from the clinical trials to estimate one year transitions for both maintenance and episodic treatment.

One key driver of differences in the cost effectiveness estimates is the relapse rate. Whilst direct evidence of the relapse rate for standard care patients is scarce, we find trial evidence to suggest that the rate may far exceed that used in the Leeds base case, with estimates of the 4 week transition probability between 0.07 and 0.14 compared to 0.0059 used in the Leeds base case model.

The CHARM data discussed by Abbott suggests an appropriate rate may be 0.4213. Using the CHARM based estimates does reduce the proportions of standard care patients in remission to a degree that is compatible with substantial evidence from systematic reviews of clinical trials.

Our model reconciliation focuses substantially on the SP model, since this has a structure that permits some transition probabilities to be considered in the Leeds model. Whilst the results and model outputs cannot be fully resolved without substituting the entire set of transition probabilities (effectively rebuilding the SP model in different software), it is clear that there are differences in results which stem from different approaches to costing infliximab. Other than this difference, this reconciliation work, together with the substituting of parameter values to and from the Leeds, SP and Abbott models, clearly identifies that the differences between model outputs is driven by the difference between clinical trial based models and the Silverstein based Leeds analysis.

Adalimumab maintenance therapy generates an ICER of below £10k when the transition probability for remission to relapse for standard and episodic care is substituted into the Leeds model with no other alterations, such that a credible proportion of patients are in remission at 1 year.

Adapting the Leeds model to more closely reflect the manufacturer analyses, by incorporating values from the SP model, suggests that episodic adalimumab dominates standard care and maintenance adalimumab is cost effective compared to episodic adalimumab (ICER = £7445).

Using Abbott's own model, which estimated the ICER for maintenance adalimumab versus standard care at £11k, but substituting cost and utility values used by Leeds, results in maintenance dominating standard care.

For infliximab, the manufacturer estimated that episodic treatment dominated standard care in the base case analysis and that maintenance generated an ICER exceeding £450k when compared to episodic care. The manufacturer estimates of

drug costs, which underpin these ICERs, are substantially lower than the Leeds estimates.

Nevertheless, the Leeds base case model was in broad agreement with these findings. When a relapse rate was used that produced more credible model outputs in terms of the proportion of patients in remission, the ICER for episodic infliximab rose to in excess of £200k per QALY. Maintenance versus standard care generates an ICER of £44k per QALY.

Amending the Leeds model to more closely reflect the manufacturer models resulted in high ICERs for episodic infliximab in all scenarios. For maintenance versus standard care, ICERs range from £29k to £69k depending on assumptions about drug costs.

Appendices

Appendix 1 - SEARCH STRATEGIES

MEDLINE

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1950 to Present>
Search Strategy:
---
1 exp Inflammatory Bowel Diseases/ or exp Crohn Disease/ (46369)
2 chron$ disease.ti,ab. (12218)
3 1 or 2 (58457)
4 tumour necrosis factor alpha inhibitor$.mp. [mp=title, original title, abstract, name of substance word, subject heading word] (49)
5 cytokine inhibitor$.mp. [mp=title, original title, abstract, name of substance word, subject heading word] (337)
6 infliximab.mp. (5050)
7 adalimumab.mp. (1207)
8 certolizumab pegol.mp. (65)
9 natalizumab.mp. (415)
10 anti-TNF$.mp. (4512)
11 8 or 6 or 4 or 7 or 10 or 9 or 5 (9504)
12 11 and 3 (1644)
13 exp Remission, Spontaneous/ or exp Remission Induction/ (40040)
14 remission.ti,ab. (64256)
15 13 or 14 (91071)
16 12 and 15 (405)
17 exp Recurrence/ (128115)
18 relapse$.ti,ab. (78846)
19 recurrence$.ti,ab. (145528)
20 recrudescence$.ti,ab. (1826)
21 loss of response.mp. (443)
22 probability of relapse.mp. [mp=title, original title, abstract, name of substance word, subject heading word] (439)
23 (predict$ adj1 relapse).mp. [mp=title, original title, abstract, name of substance word, subject heading word] (771)
24 22 or 21 or 18 or 23 or 19 or 17 or 20 (303498)
25 24 and 16 (92)
26 limit 25 to "all adult (19 plus years)" (40)
27 from 26 keep 1-40 (40)

EMBASE

Database: EMBASE <1980 to 2008 Week 47>
Search Strategy:
---
1 exp Crohn Disease/ (23956)
2 exp Enteritis/ (79354)
3 chron$ disease.ti,ab. (8757)
4 1 or 2 or 3 (87875)
5 tumour necrosis factor alpha inhibitor$.mp. (41)
6 cytokine inhibitor$.mp. (369)
7 infliximab.mp. (10849)
8 exp Infliximab/ (10774)
9 exp Tumor Necrosis Factor Alpha Inhibitor/ (1535)
10 adalimumab.mp. or exp Adalimumab/ (3426)
11 certolizumab pegol.mp. or exp Certolizumab Pegol/ (516)

12 natalizumab.mp. or exp Natalizumab/ (1328)
13 anti-TNF$.mp. (3993)
14 6 or 11 or 7 or 9 or 12 or 8 or 10 or 13 or 5 (15884)
15 4 and 14 (4311)
16 exp Remission/ (26052)
17 remission.ti,ab. (51747)
18 16 or 17 (62927)
19 18 and 15 (971)
20 exp Recurrent Disease/ (52792)
21 relapse$.ti,ab. (66112)
22 recurrence$.ti,ab. (116532)
23 recrudescence$.ti,ab. (1300)
24 loss of response.mp. (387)
25 probability of relapse.mp. (383)
26 (predict$ adj1 relapse).mp. (1586)
27 25 or 22 or 21 or 24 or 26 or 23 or 20 (207921)
28 27 and 19 (172)
29 limit 28 to (adult <18 to 64 years> or aged <65+ years>) (40)
30 from 29 keep 1-40 (40)

COCHRANE LIBRARY - CDSR

#1 exp Inflammatory Bowel Diseases/ or exp Crohn Disease/
#2 chron* disease
#3 (#1 OR #2)
#4 tumour necrosis factor alpha inhibitor*
#5 cytokine inhibitor*
#6 infliximab
#7 adalimumab
#8 certolizumab pegol
#9 natalizumab
#10 anti-TNF*
#11 (#4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10)
#12 (#3 AND #11)
#13 MeSH descriptor Remission, Spontaneous explode all trees
#14 MeSH descriptor Remission Induction explode all trees
#15 (#13 OR #14)
#16 (#12 AND #15)
#17 exp Recurrence/
#18 relapse*
#19 recurrence*
#20 recrudescence*
#21 loss of response
#22 probability of relapse
#23 predict* NEAR/1 relapse
#24 (#17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23)
#25 (#16 AND #24)

COCHRANE LIBRARY - CENTRAL

#1 exp Inflammatory Bowel Diseases/ or exp Crohn Disease/
#2 chron* disease
#3 (#1 OR #2)
#4 tumour necrosis factor alpha inhibitor*
#5 cytokine inhibitor*
#6 infliximab
#7 adalimumab
#8 certolizumab pegol

#9 natalizumab
#10 anti-TNF*
#11 (#4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10)
#12 (#3 AND #11)
#13 MeSH descriptor Remission, Spontaneous explode all trees
#14 MeSH descriptor Remission Induction explode all trees
#15 (#13 OR #14)
#16 (#12 AND #15)
#17 exp Recurrence/
#18 relapse*
#19 recurrence*
#20 recrudescence*
#21 loss of response
#22 probability of relapse
#23 predict* NEAR/1 relapse
#24 (#17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23)
#25 (#16 AND #24)

Appendix 2

Reasons for exclusion of potential studies cited by Abbott from the review of standard care relapse rate (those entering full text sift)

Author, date, ref id	Reason for exclusion	Comments
Bitton 2008 (ref id 317)	Cohort study and reporting unclear if moderate to severe patients, mentions that patients are inactive.	Abbott used reference to support that patients under stress may relapse but relapse rate data not reported.
Regueiro 2008(ref id 315)	Abstract	Full report obtained Regueiro 2009 (ref id 339)
Loftus 2008 (ref id 324)	Abstract	Full report obtained. Abbott 2008 (ref id 333)
Sandborn 2007 (ref if 314)	No relapse rate data, relapse rate not defined	RCT of Certolizumab Pegol, excluded from WMHTAC TAR as not licensed in UK
Bernklev 2005 (ref id 318)	Population based inception cohort study, severity not defined, CDAI not reported, not clear that in remission prior to relapse	34% of sample (n=169 of 497) had CD

Appendix 3

Reasons for exclusion of potential studies provided to DSU by WMHTAC (those entering full text sift)

Author, date, ref id	Reason for exclusion	Comments
Brignola 1994 (ref id 91)	Cohort study	Patients not receiving treatment whilst in remission, so not clear if moderate to severe in severity. No information regarding illness severity prior to remission.
Bitton 2008 (ref id 317)	Cohort study	No information regarding illness severity prior to remission. Reporting unclear if moderate to severe patients, mentions that patients are inactive.
Schreiber 1999 (ref id 338)	Prospective case control study	Median time to relapse data for high and low anti TNF secretion groups on standard care to achieve remission but excluded as not an RCT
Feagan 2008 (ref id 337)	Low risk patients in quiescent sub-population	EPIC 1 and II data
Consigny 2006 (ref id 90)	No information regarding illness severity	RCT level data
Zankel 2005 (ref id 92)	Population based cohort study	None
Simillis	Meta analysis, no relapse data	Defined recurrence as need for re-operation, relapse rate not defined. Checked for potential references
Keh 2005 (ref id 99)	Retrospective cohort analysis, relapse not defined, no relapse data mentioned	Post operative recurrence data.
Yamamoto 2005 (ref id 96)	No relapse data, literature review re factors affecting recurrence after surgery	Checked for potential references
Scarpa 2007 (ref id 101)	Retrospective cohort	Surgical predictors of recurrence
Scarpa 2007(ref id 93)	Cross sectional study	HR-QoL data after ileocolonic resection

Appendix 4

Systematic reviews examining placebo rates of remission and response

First Author (date)	Inclusion criteria	Number of included studies	Dates of review	Data base(s) searched	Limits	Outcomes	Time points	Follow up duration	Placebo size (n)	Mean CDAI score	Pooled estimates of the placebo rate of remission	Pooled estimates of the placebo rate of response
Su (2004) (ref id 313)	Placebo-controlled randomised clinical trials, patients had active disease at entry, clinical response and remission defined in methodology, data on response and remission in patients receiving placebo, CDAI scores used	21 studies used CDAI ≤150 to define remission 8 studies used CDAI to define response as ≥100 point reduction in CDAI	1966-2001	MEDLINE	English language studies, Human studies	Response versus no response Remission versus no remission	Studies with multiple time points: primary end point was recorded. Studies with primary end point not defined: final end point was recorded	Remission-2 to 52 weeks (range). Response-2 to 52 weeks (range).	Remission-Range: n=4 to n=77 (total n=707) Response-Range: n=6 to n=80 (total n=340)	264.65	18% (95% CI, 14% to 24% range, median 19%) Heterogeneity between studies (range 0% to 50% p=0.0003)	19% (95% CI, 13% - 28%, median 17%) Heterogeneity between studies (range 0% to 46% p=0.03)

First Author (date)	Inclusion criteria	Number of included studies	Dates of review	Data base(s) searched	Limits	Outcomes	Time points	Follow up duration	Placebo size (n)	Mean CDAI score	Pooled estimates of the placebo rate of remission	Pooled estimates of the placebo rate of response
(Abbott) (2008)(ref id 333)	Standard care treated placebo arm data in randomised controlled trials (PC-RCTs) evaluating new biologics, reported remission rate or CDAI score at baseline and 1+ time points reported	20 with 21 treatment arms with remission data	1990-August 2007	PC-RCTs from Su et al published after 1990 supplemented by MEDLINE, Digestive Diseases Weekly (DDW), American College of Gastroenterology (ACG)	Study arms with baseline <CDAI 150 excluded	Time spent in remission, CDAI score, measures of variance, % of patients with previous anti-TNF therapies	Pooled estimates over 26 weeks	Trial duration average 15 weeks (range 2-28 weeks)	Average n=60 (range 8 to 326) Total n=1257	Mean baseline= 296	Simple average time in remission across all 21 arms was 8.48% Average weighted by sample size was 11.61% Average weighted by sample size and duration of trial was 14.57% Anti-TNF naïve patients 72.48%	Not reported

First Author (date)	Inclusion criteria	Number of included studies	Dates of review	Data base(s) searched	Limits	Outcomes	Time points	Follow up duration	Placebo size (n)	Mean CDAI score	Pooled estimates of the placebo rate of remission	Pooled estimates of the placebo rate of response
Tine (2008) (ref id 319)	Control arm data from RCTs of biological therapies, active luminal CD in acute phase at baseline CDAI >150	28 trials	1997-2007	MEDLINE,EMBASE, Cochrane Library, Cochrane Controlled Trials Register	English language papers, adults Exclusions: no active or inflammatory disease	Clinical response and remission as per CDAI score Response=reduction of 100 or 70 points in CDAI.	At 6 weeks in 64% of RCTs Time at which evaluated for pooling when remission and response were primary outcomes	Not clear	Total n=1913 (range 3-325) 68% moderate to severe patients (19 of 28 RCTs)	294 (n=26)	17% (95% CI, 13% to 21% range,) Heterogeneity between studies (0 to 34% p=0.0001)	33% (95% CI, 28% to 38% range,) Heterogeneity between studies p=0.0001)

Appendix 5 Summary of CHARM trial

Study	Type of study	Study sample*	Interventions	Responders, Remission and relapse definitions	Primary end points with time points	Use of anti-TNFs
Colombel et al 2007 (ref id 323)	Phase 3 multicentre randomized double blind placebo controlled 56 week efficacy and safety study Crohn's Trial of the fully human antibody Adulimumab for remission maintainance (CHARM) 92 sites in Europe, United states, Canada, Australia, South Africa	Moderate to severely active CD (defined as a baseline CDAI score of 220-450 points, aged 18-75yrs of age with known CD of at least 4 months duration	Sub-cutaneous Adalimumab Open label induction therapy with Adalimumab 80mg (week 0), then 40mg (wk 2) (n=854) At wk 4 stratified by response then randomized to double blind treatment in one of three blinded groups(n=788); -placebo (n=261) -Adalimumab 40mg every other week (eow) (n=260) -Adalimumab 40mg weekly (n=257) N=499 (58% of 854) responded to adalimumab induction and were randomized ('randomized responde' 'population) 60 weeks follow up	**4 week responders:** decrease in CDAI ≥70 points at week 4 vs baseline who achieved clinical remission (CDAI score <150) at weeks 26 and 56. **Clinical response:** Decrease in CDAI score from baseline by ≥ 70 points and ≥ 100 points **Loss of clinical response:** decrease in CDAI < 70 points lower than baseline on 2 consecutive visits **Disease flare:** increase in CDAI of ≥70 points compared with week 4 and a CDAI score >220	% of week 4 randomized responders in remission (CDAI score <150) at weeks 26 and 56 was significantly greater in both adalimumab treatment groups vs placebo: **Week 26** Placebo:17% - Adalimumab 40mg every other week (eow) =40% Adalimumab 40mg weekly=47% **Week 56** Placebo:12% - Adalimumab 40mg every other week (eow) =36% Adalimumab 40mg weekly=41% *(p<0.001) **NB TNF antagonist naïve remission rates at week 26:** -placebo rate 18% Adalimumab 40mg every other week (eow) =47% Adalimumab 40mg weekly=50% **NB TNF antagonist naïve remission rates at week 56:** -placebo rate 14%	Patients who had received infliximab or any anti TNF other than adalimumab more than 12 wks before screening could be enrolled provided rhey did not exhibit initial non-response to the agent Patients excluded if had been treated with Adalimumab or participated in a Adalimumab clinical study. NB. At week 4 subjects stratified for previous exposure to anti-TNFs *Primary efficacy analysis adjusted for previous anti –TNF use (50% had received anti-TNF before baseline).

			(n-505 completed the 56 week study) Mean age of randomized responders: 36.7 (11.6) (n-499) Baseline CDAI score of randomized responders: 316.6 (62.5)	**Significant improvement in symptoms:** decrease in CDAI of ≥ 70 points vs baseline	Adalimumab 40mg every other week (eow) =42% Adalimumab 40mg weekly=48%	

References

1 DSU Report, Use of tumour necrosis factor alpha (tnf a) inhibitors (adalimumab and infliximab) for Crohn's disease. 20th November 2008

2 Silverstein, M.D., Loftus, E.V., Sandborn, W.J., Tremaine, W.J., Feagan, B.G., Nietert, P.J., Scott Harmsen, W.S., and Zinsmeister, A.R. (1999) Clinical course and costs of care for Crohn's disease: Markov model analysis of a population –based cohort, Gastroenterology, Vol.117:49-57.

3 Su, C., Lichtenstein, G. R., Krok, K., and et al A meta-analysis of the placebo rates of remission and response in clinical trials of active Crohn's disease. Gastroenterology 2004; 126 1257-1269.

4 Tine, F., Rossi, F., Sferrazza, A., Orlando, A., Mocciaro, F., Scimeca, D., Olivo, M., and Cottone, M. Meta-analysis: remission and response from control arms of randomized trials of biological therpaies for active luminal Crohn's disease. Alimentary Pharmacology & Therapeutics 2008; 27 1210-1223.

5 Colombel, J-F., Sandborn, W. J., Rutgeerts, P., Enns, R., Hanauer, S. B., Panaccione, R., Schreiber, S., Byczkowski, D., Li, J., Kent, JD., and Pollack, PF. Adalimumab for maintenance of clinical response and remission in patients with Crohn's disease: the CHARM trial. *Gastroenterology* 2007; **132** 52-63.

6 Rutgeerts, P., D'Haens, G., Targan, S., Vasiliauskas, E., Hanauer, S., Present, D. H., Mayer, L., Van Hogezand, RA., Braakman, T., DeWoody, KL., Schaible, TF., and van Deventer, S. J. Efficacy and safety of retreatment with anti-tumor necrosis factor antibody (Infliximab) to maintain remission in Crohn's disease. *Gastroenterology* 1999; **117** 761-769.

7 Feagan, B. G., Fedorak, R. N., Irvine, E. J., Wild, G., Sutherland, L., Steinhart, A. H., Greenberg, G. R., Koval, J., Wong, C. J., Hopkins, M., Hanauer, S. B., and McDonald, J. W. A comparison of methotrexate with placebo for the maintenance of remission in Crohn's disease. North American Crohn's Study Group Investigators. *New England Journal of Medicine* 1-6-2000; **342** 1627-1632.

8 Hanauer, S. B., Feagan, B. G., Lichtenstein, G. R., Mayer, L. F., Schreiber, S., Colombel, J. F., Rachmilewitz, D., Wolf, D. C., Olson, A., Bao, W., Rutgeerts, P., and ACCENT, I. Study Group Maintenance infliximab for Crohn's disease: the ACCENT I randomised trial. *Lancet* 4-5-2002; **359** 1541-1549.

9 Sands, B. E., Anderson, F. H., Bernstein, C. N., Chey, W. Y., Feagan, B. G., Fedorak, R. N., Kamm, M. A., Korzenik, J. R., Lashner, B. A., Onken, J. E., Rachmilewitz, D., Rutgeerts, P., Wild, G., Wolf, D. C., Marsters, P. A., Travers, S. B., Blank, M. A., and van Deventer, S. J. Infliximab maintenance therapy for fistulizing Crohn's disease. *New England Journal of Medicine* 26-2-2004; **350** 876-885.

10 Abbott Abbott's response to the Appraisal Consultation document of adalimumab and infliximab for the treatment of Crohn's Disease. 2008.

11 Buxton, M.J., Lacey, L.A., Feagan, B.G., Niecko, T., Miller, D.W., and Townsend, R.J. (2007) Mapping from disease specific measures to utility: An analysis of the relationship between the inflammatory Bowel Disease Questionnaire and Crohn's Disease Activity Index in Crohn's Disease and measures of utility, Value in Health, Vol.10(3):214-220.