Meindert Boysen
Programme Director, Centre for Health Technology Evaluation
National Institute for Health and Care Excellence
10 Spring Gardens
London
SW1A 2BU

5th August 2014

Dear Meindert,

RE: Appraisal 695: Ulcerative colitis (moderate, severe) - infliximab (review TA140), adalimumab (review TA262) & golimumab (2nd line)

MSD welcomes the opportunity to comment on the Assessment Group (AG) report which sets out the systematic review, evidence synthesis, and economic modelling exercise undertaken by the AG.

MSD's key comments are focused on the following areas:

1. The AG has not included the approved Patient Access Scheme (PAS) for golimumab in the report. We understand that updated analyses will be presented as an addendum to the Committee, however, we remain concerned that the initial analyses include an outdated price

2. The AG criticise the MSD treatment pathway which compares surgery to TNF-α inhibitors yet describe a similar pathway (considering surgery and TNF-α inhibitors after conventional therapy failure and including a model state after TNF-α inhibitor failure similar to 'relapse management')

3. The AG has considered a subgroup of patients for whom surgery is not an "*acceptable or preferable option*". For these patients we believe that TNF-α inhibitors represent the only treatment choice once conventional therapy has failed

4. The AG has used indirect data at 32-52 weeks to inform modelling over the entire time horizon, which will overstate the benefits of adalimumab. It would be more appropriate to use data from the whole maintenance period (8-52 weeks), in-line with the design of the RCTs

5. The AG has incorporated a lower rate of adalimumab dose escalation (from every other week to weekly) than evidence suggests happens in clinical practice in the UK

We would like to reiterate that patients with moderate to severe UC are in need of effective treatments once conventional therapy has failed, particularly patients for whom colectomy is inappropriate or not preferable. We have demonstrated that for the appropriate patients at the right stage in the treatment pathway, TNF-α inhibitors can be considered a cost-effective option.

Kind regards,

1. Omission of PAS within AG modelling

The PAS for golimumab has not been considered in the acquisition cost and has been omitted from all cost-effectiveness modelling in the AG report. While we appreciate that the PAS was being reviewed by PASLU at the time the AG compiled their report, it would have been prudent to include the proposed PAS in a sensitivity analysis (MSD presented cost-effectiveness modelling both with and without the PAS for golimumab).

The PAS has now been approved by the Department of Health (decision communicated on 16th July 2014) and it is therefore necessary to conduct an analysis using the AG model in which the PAS is included. We have not been able to conduct this analysis using the executable form of the AG model due to redaction of data within the model. However, we anticipate that including the PAS for golimumab will have a considerable impact on its cost-effectiveness.

NICE has agreed that accurate information must be presented to the Committee to support decision-making and we understand that analyses including the PAS will be performed and presented as an addendum to the AG report. However, we remain concerned that the first analyses to be presented to the Committee will include an outdated price for golimumab and will not be an accurate assessment of cost-effectiveness.

2. Treatment pathway considered by the AG

The treatment pathway modelled by MSD has been criticised for not including conventional therapy as a comparator. The rationale for this exclusion is that TNF-α inhibitors are indicated for use only after conventional therapy has failed, or in patients for whom conventional therapy is inappropriate. This position is also reflected in the population specified in the appraisal scope: "*adults aged 18 years and over with moderately to severely active UC who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications against such therapies"*. Further, ECCO guidelines on the management of immunomodulator-refractory UC state that: "*continued medical therapy that does not achieve a clear clinical benefit is not recommended*" (Dignass *et al*. 2012). Through discussion with UK gastroenterologists, we understand that the key decision point regarding the use of TNF-α inhibitors is at the point where conventional therapy has failed, and where the remaining treatment options are TNF-α inhibitors and surgery (a choice which is reflected in the MSD modelling approach).

We acknowledge that patients may continue to receive background conventional therapy, however, this scenario is reflected in the RCTs in which patients receiving conventional therapy at study entry continued to receive conventional therapy throughout i.e. any continuing benefits from conventional therapy are captured through the clinical data inputted into the model.

We note that in the treatment pathway described by the AG in Figure 1 (p.15 of the AG report), the options for patients who have not responded to 5-ASAs, prednisolone, azathioprine, and 6-mercaptopurine are stated to be "*tacrolimus, anti-TNF-α, or iv steroids*". Tacrolimus is not licensed for UC and is rarely used. The remaining options (TNF-α inhibitors and IV steroids) are considered in the

biologic arm of our model (initial treatment with the TNF-α inhibitor is followed by relapse management with IV steroids). Following this, the AG treatment pathway states that surgery can be considered. We believe that both treatment pathways are largely in agreement, the only difference being that surgery is considered at the same line of therapy as TNF-α inhibitors in the pathway modelled by MSD. This is necessary because no TNF-α inhibitors are currently recommended by NICE guidance in the sub-acute setting, meaning that surgery is presently the only choice once other treatment options have failed. We also note that the description of when patients might require surgery closely aligns to the patients who would be considered for treatment with a TNF-α inhibitor (p.16 of the AG report).

Finally, the AG model (described on p.217 of the AG report) contains four key phases: 'on biological treatment', 'non-biologic treatment', 'post-surgery', and 'dead'. If patients receiving biologics lose their response they are assumed to transition to the active UC state and subsequently receive conventional therapy (in 'non-biologic treatment'). We note that the 'non-biologic treatment' phase is akin to the relapse management phase considered in the MSD model (in which patients not responding to TNF-α inhibitor treatment receive IV steroids prior to receiving surgery and moving to the post-surgery state).

3. <u>Subgroup of patients for whom surgery is not acceptable or preferable</u>

The AG report has described the cost-effectiveness of TNF-α inhibitors in a population for whom surgery is "*not an acceptable or preferable option*". We agree that surgery will not be suitable for some patients and that for those patients, access to TNF-α inhibitors represents an opportunity to manage their condition where conventional therapy has failed. This group may be relatively small and is subject to clinical judgment, though through discussions with UK clinicians and reference to published literature, we understand patient groups for whom surgery is not appropriate or preferable to include:

Female patients who wish to conceive children (now or in the future)

Some types of surgery for UC can severely reduce female fertility. In patients who have undergone ileal pouch-anal anastomosis (IPAA), there is a significant reduction in the ratio of patient to reference population fecundity to 0.20 ($p<0.0001$) (Ørding Olsen *et al*. 2002). In a meta-analysis of published data, IPAA was shown to increase the risk of infertility in women with UC by approximately threefold (from 15% to 48%) (Waljee *et al*. 2006). Further, in a study comparing surgical and non-surgical management of UC, the infertility rate was significantly higher in females who had pelvic pouch surgery compared with females managed non-operatively (38.1% compared with 13.3%; $p<0.001$) (Johnson *et al*. 2004).

Patients at higher risk of post-operative complications

Post-operative complications (including pouchitis, pouch leakage, pelvic abscesses, pouch fistulae, small bowel obstruction, anastomotic stricture, post-operative bleeding, faecal incontinence, sexual dysfunction, infections, delayed wound healing, and nerve damage) frequently occur after surgery for UC, predominantly in elderly patients with multiple comorbidities. Patients aged over 64 years are more likely to experience complications than patients aged between 18 and 34 years (odds ratio [OR], 1.95; 95% confidence interval [CI], 1.07-3.54) and those with two or more comorbidities are also more likely to experience complications compared with patients that have none (OR, 1.89; 95% CI, 1.06-3.37) (De

Silva *et al*. 2011). Surgery is unlikely to be suitable for patients with contraindications to it e.g. in patients with respiratory disease, obesity, coagulopathy, cardiovascular disease etc.

Patients with disease limited to the rectum (ulcerative proctitis)

Patients with ulcerative proctitis may receive reduced or delayed benefits from surgery compared with those with more extensively-located disease. We understand that patients with proctitis are typically excluded from trials of TNF-α inhibitors, however, given the tendency to refractory disease in these patients an additional treatment option would be of benefit. One study of patients with refractory proctitis who were treated with infliximab showed that "*infliximab therapy seems to be effective in inducing and maintaining a clinical response in refractory ulcerative proctitis*", with complete response observed in 69% of patients in induction and in 82% of patients in maintenance (Bouguen *et al*. 2010).

Patients treated in hospitals with low volumes of surgery

Surgeries in a low-volume hospital are associated with poorer outcomes than surgeries in a high-volume hospital (Kaplan *et al*. 2008). As NHS surgeons perform a median of four procedures per year and 30% of NHS hospitals perform fewer than two procedures per year (Burns *et al*. 2011), surgeons may not obtain the required proficiency in the technique (estimated to require 40 surgeries per year [Tekkis *et al*. 2005]). The draft NICE IBD Quality Standard recommends that surgery is performed in high volume centres: "*People with inflammatory bowel disease who need surgery have it undertaken by a colorectal surgeon in a unit where the operations are performed regularly"* but acknowledges that this may be difficult for some patients: "*some people have to travel a significant distance from their home depending on local service configuration, which may make access to treatment difficult and also reduce the ability of family and friends to support people in some socioeconomic groups"*.

Patients who do not want to undergo surgery

Some patients may simply have a strong desire not to undergo a major surgical procedure. A study on patient preference measuring their willingness and the amount of life they would trade or gamble to avoid certain outcomes has shown that UC patients will choose to receive escalating medical therapy rather than undergo surgery (Byrne *et al*. 2013). For certain patients the psychological impact of surgical intervention would be too great (i.e. patients with "*psychological problems, emotional instability, poor motivation or* [who] *are non-compliant"* [Frizelle 2001]) and for others the consequences of surgery (i.e. a colostomy bag that needs frequent re-emptying) will negatively impact their working practices or other activities of daily living.

Not recommending TNF-α inhibitors for patients failing conventional therapy dictates that a patient with UC is condemned to surgery if disease symptoms persist. Discussions with clinical experts at the scoping workshop for this appraisal strongly indicated that the majority of patients would prefer the choice of trying a medical intervention prior to undergoing surgery, and this choice does not currently exist for the moderate to severe UC population.

In summary, we assert that patients who are identified as being inappropriate for surgery or that would prefer not to undergo surgery should be offered TNF-α inhibitors as an alternative effective treatment after conventional therapy has failed.

4. Inappropriate incorporation of maintenance data into economic model

The AG network meta-analysis (NMA) reports that in the analysis for 'maintenance, 8-32 weeks, patients starting in remission': "*all treatments except adalimumab were associated with beneficial treatment effects relative to placebo*" (Figure 20). However, the analysis for 'maintenance, 32-52 weeks, patients starting in remission' shows: "*…the greatest effect being associated with adalimumab. However, only the treatment effects of adalimumab were statistically significant at a conventional 5% level*" (Figure 26).

This is of critical importance because the AG model uses a cycle length of half a year using data extracted from trials at a mid-time point, unlike our analysis which is based on response over the whole year. Infliximab and golimumab demonstrated very good induction and maintenance data over the first half of the year, but had a reduced effect over the second half. For golimumab, this may be due to the more stringent requirements for patients to meet the primary efficacy endpoint of continuous clinical response. In contrast, adalimumab has poor induction and maintenance data for the first half of the year and improved maintenance data over the second half. The transitions from the second half of the year are repeated and used throughout the remainder of the AG model "*It should be noted that beyond 1-year, the model repeatedly uses the transition probabilities derived within the maintenance phase 2 NMA*" (p.221 of the AG report). This will be overstating the benefit for adalimumab over the long-term. We suggest that the AG conducts a sensitivity analysis in which maintenance response over the whole year is considered, in line with the RCT trial designs.

5. Consideration of adalimumab dose escalation

The AG model considers that 27.4% of patients receiving adalimumab are up-titrated to once weekly dosing in the maintenance setting (informed by the ULTRA study). In our model, although UK gastroenterologist clinical experts indicated that up to 80% of patients receiving adalimumab are up-titrated, we used a conservative assumption of 50% (between the two estimates). A recent independent study indicates that adalimumab dose escalation was necessary in 76% of patients (Ferrante *et al*. 2011). This suggests that the value used by the AG may be too low and they may need to consider input from UK clinical experts to correctly appraise the cost of treatment with adalimumab and therefore the resulting cost-effectiveness of the TNF-α inhibitors being compared.

Further to these key points, we have also presented a brief response to the AG's critiques of the MSD modelling approach and a list of factual inaccuracies identified in the AG report in Appendices 1 and 2, respectively.

Appendix 1: AG critiques of the MSD modelling approach

1. **Deviations from the NICE Reference Case and final NICE scope, particularly with respect to omission of conventional non-surgical management as a comparator**

1a. Omission of non-biologic non-surgical treatments

The rationale for excluding conventional therapy as a comparator in the MSD model is described in point 2 ('Treatment pathway considered by AG') on p.2 of this response.

1b. Time horizon

Results of a fully incremental analysis with a 10 year, 20 year, and lifetime horizon are provided below for the infliximab and golimumab models (Table 1 and Table 2, respectively). Findings are similar with longer time horizons as with a 10 year time horizon. Cost-effectiveness acceptability curves over the lifetime horizon are provided in Figure 1 and Figure 2 (for infliximab and golimumab, respectively).

Table 1: 10 year, 20 year, and lifetime results, infliximab model with PAS

Treatment	QALYs	Costs (£)	Incremental QALYs	Incremental cost (£)	ICER (£)
10 year follow-up					
Probabilistic model results					
infliximab	5.71	44,189	0.17	12,842	75,988
golimumab	5.54	31,348	0.57	15,523	27,163
adalimumab	5.48	32,123	-	-	Dominated
colectomy	4.97	15,825	-	-	-
Results based on point estimates of parameters					
infliximab	5.66	42,920	0.16	12,166	77,599
golimumab	5.51	30,753	0.56	14,964	26,569
adalimumab	5.45	31,237	-	-	Dominated
colectomy	4.94	15,790	-	-	-
20 year follow-up					
Probabilistic model results					
infliximab	9.14	50,394	0.19	13,727	72,281
golimumab	8.95	36,666	0.61	16,266	26,658
adalimumab	8.90	37,535	-	-	Dominated
colectomy	8.34	20,400	-	-	-
Results based on point estimates of parameters					
infliximab	9.11	48,740	0.17	12,524	74,921
golimumab	8.94	36,216	0.59	15,781	26,689
adalimumab	8.89	36,615	-	-	Dominated
colectomy	8.35	20,435	-	-	-
Lifetime follow-up					
Probabilistic model results					
infliximab	13.33	54,513	0.19	13,701	71,509
golimumab	13.14	42,640	0.61	16,488	26,847

adalimumab	13.08	43,318	-	-	Dominated
colectomy	12.52	26,152	-	-	-
Results based on point estimates of parameters					
infliximab	13.30	54,513	0.17	12,541	74,763
golimumab	13.13	41,972	0.59	15,827	26,695
adalimumab	13.07	42,366	-	-	Dominated
colectomy	12.54	26,145	-	-	-

Table 2: 10 year, 20 year, and lifetime results, golimumab model with PAS

Treatment	QALYs	Costs (£)	Incremental QALYs	Incremental cost (£)	ICER (£)
10 year follow-up					
Probabilistic model results					
infliximab	5.70	44,382	0.16	13,004	80,318
golimumab	5.54	31,379	0.56	15,611	27,994
adalimumab	5.49	32,096	-	-	Dominated
colectomy	4.98	15,768	-	-	-
Results based on point estimates of parameters					
infliximab	5.65	43,092	0.15	12,197	80,866
golimumab	5.50	30,895	0.55	15,101	27,322
adalimumab	5.45	31,370	-	-	Dominated
colectomy	4.95	15,794	-	-	-
20 year follow-up					
Probabilistic model results					
infliximab	9.11	50,688	0.19	13,896	73,020
golimumab	8.92	36,792	0.62	16,329	26,514
adalimumab	8.87	37,918	-	-	Dominated
colectomy	8.30	20,463	-	-	-
Results based on point estimates of parameters					
infliximab	9.11	48,941	0.16	12,558	77,955
golimumab	8.95	36,383	0.58	15,927	27,421
adalimumab	8.89	36,773	-	-	Dominated
colectomy	8.37	20,457	-	-	-
Lifetime follow-up					
Probabilistic model results					
infliximab	13.53	56,848	0.18	13,923	76,711
golimumab	13.35	42,926	0.59	16,610	27,940
adalimumab	13.30	43,878	-	-	Dominated
colectomy	12.75	26,361	-	-	-
Results based on point estimates of parameters					
infliximab	13.44	54,912	0.16	12,575	77,779
golimumab	13.28	42,337	0.58	15,974	27,426
adalimumab	13.22	42,722	-	-	Dominated
colectomy	12.70	26,363	-	-	-

Figure 1: Cost-effectiveness acceptability curves - lifetime time horizon, infliximab model with PAS

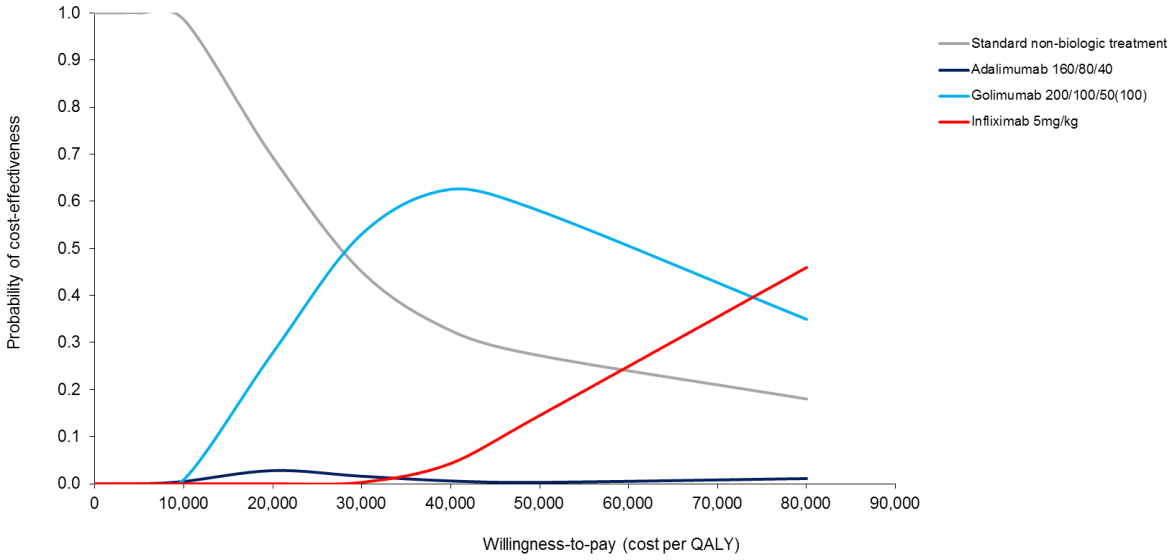
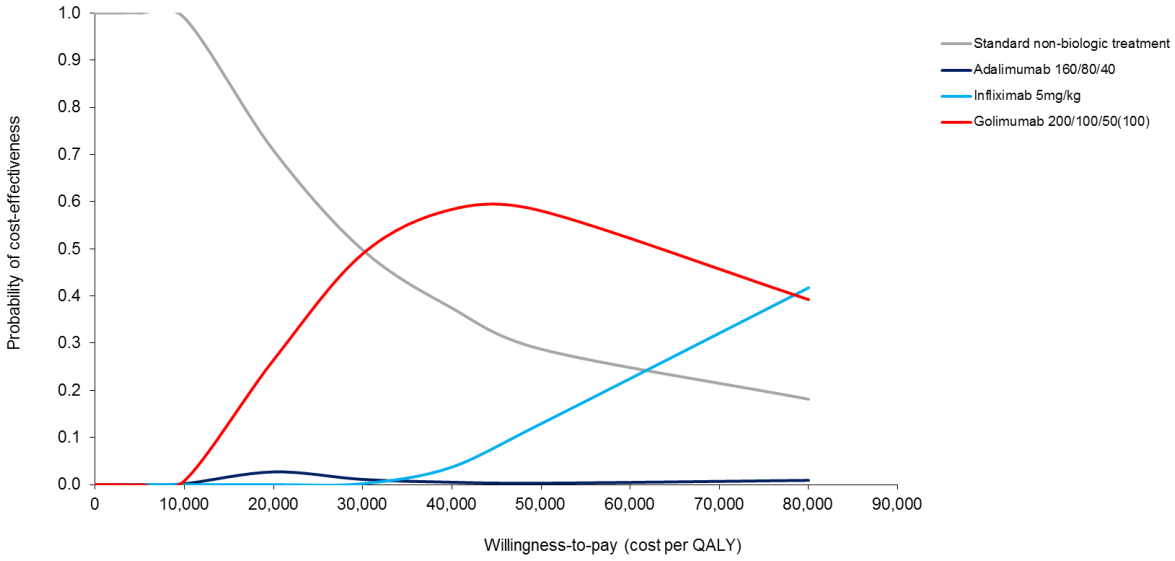


Figure 2: Cost-effectiveness acceptability curves - lifetime time horizon, golimumab model with PAS



2. Assumption that treatment failure is equivalent to severe exacerbation

2a. Place of surgery within treatment pathway in MSD model

The rationale for excluding conventional therapy as a comparator in the MSD model is described in point 2 ('Treatment pathway considered by AG') on p.2 of this response.

The AG report states that: "*The pathway represented by the model after failure of biologic therapy (or in its absence) appears to assume that failure to achieve an induction response, or that the loss of response during maintenance therapy, is synonymous with an acute UC exacerbation"*. We disagree that the treatment pathway modelled in our submission only relates to acute UC (excluded from the scope of this analysis), as surgery is also considered for patients with moderate to severe UC when conventional therapies have failed (as outlined in NICE CG166…"*These recommendations apply to anyone with ulcerative colitis considering elective surgery*"), and in the absence of any NICE recommendations for TNF-α inhibitors.

We also note that the modelled pathway/structure of the MSD model is very similar to the AG model:

- Both models consist of an induction and maintenance phase
- In both models, patients receiving TNF-α inhibitor treatment are assumed to continue receiving the same TNF-α inhibitor for as long as they continue to maintain response/remission
- Both models assume a non-biologic treatment phase until colectomy occurs. In the MSD model this is called the 'relapse management' phase, and in the AG model 'non-biologic treatment'
- In both models the duration of the phase between discontinuation of the TNF-α inhibitor and colectomy (i.e. 'relapse management' or 'non-biologic treatment' phase) is determined by the assumed colectomy rate and probability of maintaining response with non-biologic treatment

Further, the following statement by the AG is not an accurate reflection of the MSD model: "*After failing biologic treatment, the MSD model assumes that all patients who have failed biologic therapy will receive iv steroids and rapidly progress to colectomy (27% of all relapsing patients during each 2-month cycle). This fails to reflect the possibility that patients may continue to receive, and may still obtain clinical benefit from, non-biologic medical treatment options as defined in the model scope (5-ASAs, immunotherapies and/or steroids)*". After failure of the TNF-α inhibitor, patients start the relapse management phase in the relapse state where IV steroids are provided. Upon successful treatment, which has a 100%-27%=73% probability, patients transition to a response state. In this state IV steroids are no longer provided but patients do continue mesalazine, azathioprine, 6-mercaptopurine, and oral steroids. The probability of losing response and transitioning back to relapse (17.1% per 2-month period) is based on NMA results of maintenance treatment with placebo, similar to the AG model. The 73% probability of transitioning from relapse to response in the relapse management phase, combined with the 17.1% probability of transitioning from response back to relapse every 2 months means that the colectomy rate in the MSD model is effectively smaller than the 27% stated by the AG.

Overall, the primary difference between the MSD and AG model in terms of flow of patients through the model are the colectomy rates used.

2b. Colectomy rates

The cumulative incidence of colectomy estimated with the MSD model is a direct result of the discontinuation rates of TNF-α inhibitors and the efficacy of subsequent non-biologic drug treatment in the relapse management phase (i.e. failure of conventional treatment and unsuccessful IV steroid use). Although the modeled cumulative incidence of colectomy (Figure 3) may be an over-estimate, a comparison of our modeled estimates with data in ACT1/2 and recent observational studies of infliximab-treated cohorts (

Table 3) suggest that our modeled estimates are closer to observed data than the colectomy rate used by the AG.

The AG report also states: "*These rates are very high and fail to reflect both the possibility of benefit from further medical therapies and the element of patient choice in deciding whether or not to undergo colectomy. If surgery really was the only remaining treatment option for these patients, it would not have been possible (or ethical) to undertake any of the trials included in this assessment*". We disagree with this statement as patients who entered the trials on conventional therapy continued to receive background conventional therapy. However, that conventional therapy had not been successful in managing their condition, necessitating the consideration of another treatment option. With the present lack of recommendations for moderate to severe UC, currently the only option for these patients is surgery. TNF-α inhibitors are an alternative to surgery.

Figure 3: Cumulative incidence colectomy (light blue) over time with infliximab in MSD model

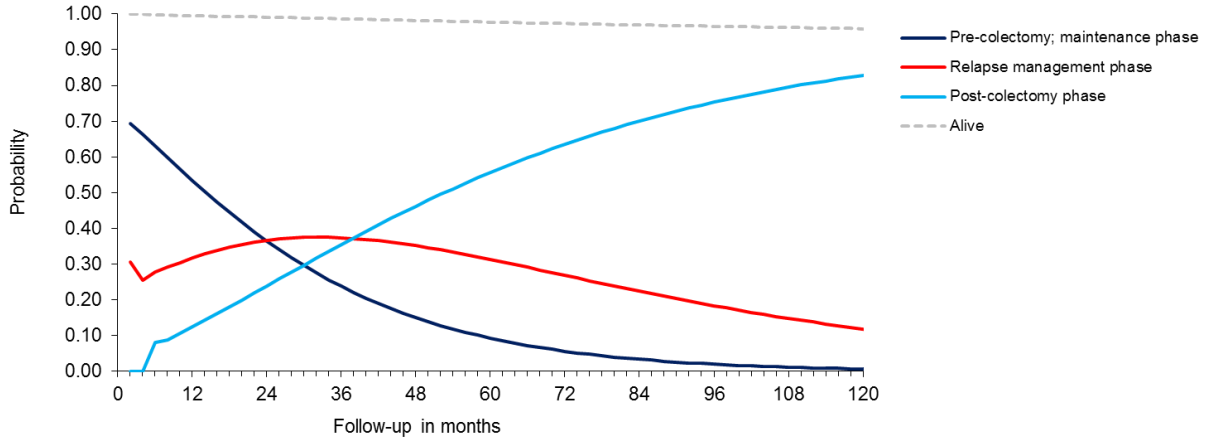


Table 3: Estimated cumulative incidence of colectomy in infliximab studies

	Follow-up (mths)	Proportion free of colectomy at end of follow-up	Proportion colectomy at end of follow-up	Proportion colectomy estimated assuming constant underlying hazard rate			
				12 mths	24 mths	60 mths	120 mths
RCTs							
ACT1/ACT2 combined	12	0.90	0.10	0.10	0.18	0.39	0.63
Observational infliximab studies							
Armuzzi 2013; all	12	0.90	0.10	0.10	0.18	0.39	0.63
Armuzzi 2013; steroid	12	0.81	0.19	0.19	0.34	0.65	0.87
Oussalah 2010	60	0.70	0.30	0.07	0.13	0.30	0.51
Rostholder 2012	24	0.73	0.27	0.15	0.27	0.54	0.79

2c. Use of Turner et al. (2007) study

The review by Turner includes studies describing the short-term outcome of moderate to severe UC patients treated with IV corticosteroids, with or without ciclosporin, for exacerbations. Given that the UC patient at the time/decision point considered in the MSD model is characterised by a history of failed conventional treatment (according to the decision problem outlined in the scope) and has failed biologic treatment and relapsed (i.e. has moderate to severe active UC), the data provided by Turner to inform colectomy rates among these relapsed patients was considered appropriate. Please note that we did not apply the colectomy rate informed by the Turner study to patients in the response health state.

3. Questionable use of evidence concerning improving and worsening of inflammation

For all NMA analyses the response and remission outcomes were analysed separately, where response includes patients in remission.

In the transformation of the odds ratios obtained with the NMA for induction treatment to probabilities of response and remission for the model, response in the cost-effectiveness model EXCLUDED remission. More specifically, 1 minus probability of response from the NMA model was the probability of failed induction treatment in the model. The probability of response in the cost-effectiveness model was calculated as 1 minus probability of failed induction minus probability of remission as obtained with the NMA model. The concept used for induction treatment is comparable to the use of the multinomial model.

We agree that a multinomial model for the NMA of maintenance of remission/response is a more elegant approach and reflects the data generating process. However, as indicated by the AG, the lack of reported information for ULTRA2 did not allow such an approach, and we could only perform an NMA that included adalimumab assuming remission and response to be binomial variables. Although patients in the response health state can improve and transition to remission, or flare up and transition to relapse, the duration of biologic treatment in the MSD model is a reflection of time to relapse. As such,

the dichotomisation of losing/leaving response may not yet have biased the economic model outcomes. To ensure that the modeled duration of treatment with a TNF-α inhibitor matched the observed discontinuation rates of biologic treatment in the trial for golimumab as closely as possible, we had to assume that patients losing remission did not transition straight to relapse.

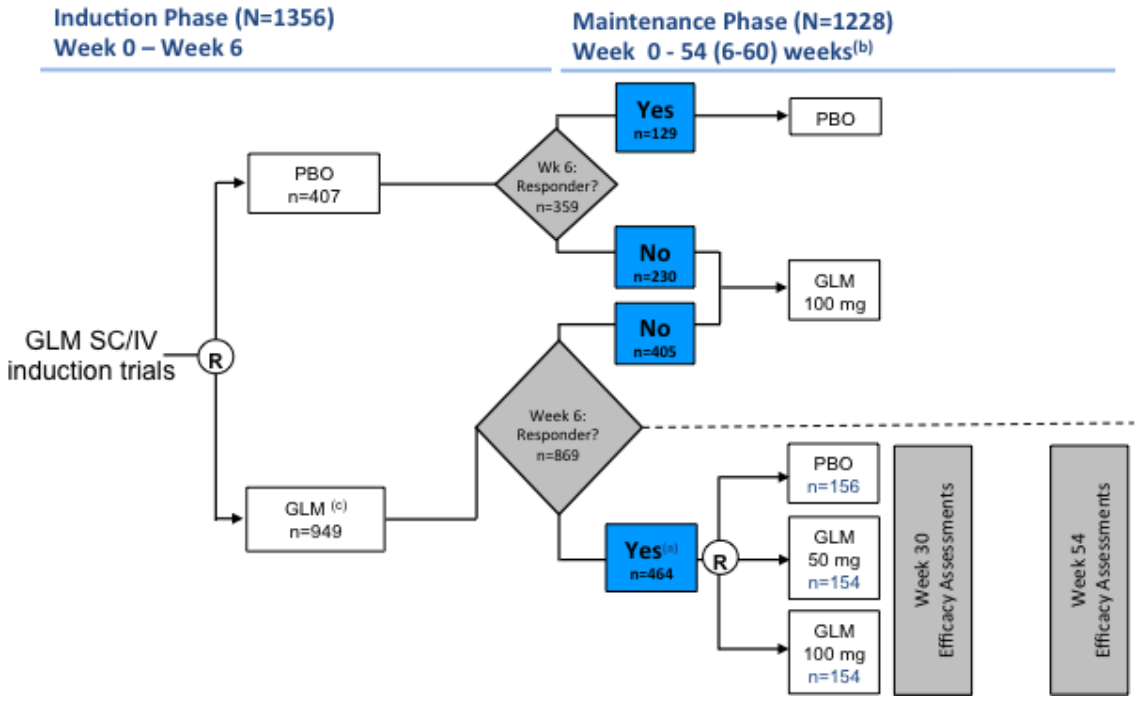
4. **Questionable validity of use of novel methods for including non-randomised data from PURSUIT**

In general the NMAs were performed in line with established methods published by the NICE DSU. Given the distinction between induction and maintenance treatment, different NMAs were performed for each setting. In Table 4 an overview is provided indicating which specific assumptions were made to allow comparisons of PURSUIT, ACT, and ULTRA data and whether the findings were used in the economic model. The maintenance analysis that required 'novel' assumptions/approaches did not form the basis for the economic model.

Table 4: Overview of NMAs performed for response and remission using logistic regression models with a binomial likelihood for r/n

Outcome	Cases (r)	Population (n)	Used in CE model	Comments
Induction response	Responders at end of induction	Patients at beginning of induction treatment	Yes	N/A
Induction remission	Patients in remission at end of induction	Patients at beginning of induction treatment	Yes	N/A
Response at end of maintenance treatment	Responders at end of maintenance	Patients at beginning of induction treatment	No	The PURSUIT trial investigating the efficacy of golimumab was split in two stages: the induction stage and the maintenance stage. At the induction stage, patients were randomised to placebo or various induction doses of golimumab. After 6 weeks, patients were then re-randomised or re-allocated depending on their response. Given the reallocation and re-randomisation the calculation of proportion of response and remission at the end of maintenance out of all patients starting induction with the same treatment cannot be done without making certain assumptions, as described in the submission. However, given these necessary assumptions this NMA was not used for the model. In fact, it would not have matched the model structure where induction and maintenance phase are stratified.
Remission at end of maintenance treatment	Patients in remission at end of maintenance	Patients at beginning of induction treatment	No	
Sustained response	Responders at end of induction AND maintenance	Patients at beginning of induction treatment	No	For PURSUIT the number of cases was calculated by multiplying the sample size at the beginning of the induction phase with probability of induction response/remission at the end of the induction phase and with the probability of maintaining response/remission at the end of the maintenance phase. For the placebo arm no randomised maintenance data was available (see Figure 4 below). However, this is not an issue for the NMA because the data are still conditional upon the randomised induction phase and trials are indirectly compared as such.
Sustained remission	Patients in remission at end of induction AND maintenance	Patients at beginning of induction treatment	No	
Sustained response among patients starting maintenance	Responders at end of maintenance	Patients in response at end of induction and starting with the same maintenance treatment	Yes	For PURSUIT, the control arm consisted of patients that responded to induction treatment with placebo and continued maintenance with placebo. This is NOT the placebo group of the PURSUIT-Maintenance trial, because this later population received golimumab as induction as illustrated in Figure 4 below. However, the selection mechanism at the end of the placebo and golimumab induction phases can be considered the same and as such there is no reason to believe that comparison of sustained response/remission with golimumab relative to placebo among patients starting maintenance is biased.
Sustained remission among patients starting maintenance	Patients in remission at end of maintenance	Patients in remission at end of induction and starting with the same maintenance treatment	Yes	

Figure 4: PURSUIT design



5. Lack of clarity regarding the NMA model

5a. Lack of clarity

Given the distinction between induction and maintenance treatment, the following NMAs can be performed to synthesise the available evidence:

1. NMA induction
   a. Response/remission cases (numerator) out of all patients at beginning of induction (denominator)

2. NMA maintenance
   a. Response/remission cases at a certain time point during maintenance treatment (numerator) out of all patients at beginning of induction with the same treatment (denominator);
   b. Response/remission at a time point during maintenance treatment AND response/ remission at end of induction/start maintenance (numerator) out of all patients at beginning of induction with the same treatment (denominator);
   c. Response/remission cases at a time point during maintenance treatment (numerator) out of all patients that showed response/ remission at end of induction/ start maintenance with the same treatment (denominator);
   d. Response/remission cases at a time point t during maintenance treatment (numerator) out of all patients that showed response/ remission at time point t-1 during maintenance with the same treatment (denominator)

For the MSD cost-effectiveness model, analysis 1 was used to obtain efficacy estimates associated with induction treatment. In order to obtain the required estimates of relative treatment effects in terms of sustained response for the economic model, we used option 2c; the observed number of patients with a sustained response at the end of the maintenance phase (numerator) out of the number of patients showing and induction response (denominator) were obtained from ACT 1 and ULTRA 2 trials. The same approach was used for sustained remission. In PURSUIT, responders to golimumab 200mg/100mg or 400mg/200mg in the 6 week induction phase were re-randomised to placebo, golimumab 50mg, and golimumab 100mg groups for the maintenance phase. Hence, sustained response or remission was the proportion of patients showing a response or being in remission at 54 weeks out of patients randomised to maintenance treatment. Since only responders to induction treatment with golimumab were included in the randomised maintenance trial, the sustained responders with placebo maintenance treatment given placebo induction response were obtained from the non-randomised maintenance patients. Please note that the primary treatment effects obtained with this NMA reflect the conditional probability of maintaining response or remission with a biologic given induction with the same biologic relative to the conditional probability of response or remission with placebo following induction treatment with placebo.

In the AG report it seems that option 2c and 2d have been used. What is unclear is how maintenance efficacy data of PURSUIT were used in the analyses. Did the analysts use the placebo arm of the randomised maintenance phase, which only includes patients that received induction treatment with golimumab, or was a similar approach used for the placebo arm of maintenance treatment as for the MSD model by using maintenance placebo data among placebo induction responders? Unfortunately the source data have been redacted in the appendix, which compromises the transparency of the NMA performed by the AG.

5b. Baseline model

The odds ratios of induction response and remission with each treatment relative to placebo were obtained with a random effects NMA model for dichotomous endpoints using the DSU bugs code for dichotomous outcomes (Dias *et al*. 2013). The average of the estimated log odds of a response with placebo of each trial as obtained with the NMA model (the study effects; the 'mu') was transformed into probability of response with placebo (0.355; 95% credible interval [CrI] 0.322, 0.390). This estimate was assumed to reflect the outcome with non-biologic drug induction treatment. In the same way the probability of remission with placebo was obtained (0.089; 95% CrI: 0.071, 0.110).

The efficacy of drug maintenance treatment with each of the TNF-α inhibitors and conventional treatment was obtained with fixed effects NMA models (Dias *et al*. 2013). A fixed effects model was used because there is only one trial for each direct comparison in this placebo-controlled 'star-network' meaning there are no data to estimate between-study heterogeneity. Again, the average of the estimated log odds of a sustained response with placebo of each trial as obtained with the NMA model (the study effects; the 'mu') was transformed into a 2-month probability (0.829; 95% CrI: 0.797, 0.858) assuming a time-independent constant underlying loss of response rate. This estimate was assumed to

reflect the 2-month sustained response with non-biologic drug treatment. In a similar fashion, the 2-month probability of sustained remission with placebo was obtained (0.861; 95% CrI: 0.806, 0.908).

6. Inconsistencies between results of the MSD golimumab and infliximab models

The golimumab model used utility data and distribution of drugs/resource use as reported in the golimumab trials, and the infliximab model used the corresponding infliximab data. Given that two MSD products were included in this appraisal, two dossiers and associated models were submitted. Accordingly we opted to use trial-specific data in each of the models, as we would have if a submission for only one MSD product was made.

The models for infliximab and golimumab are exactly the same; one model was a copy of the other and the only modifications made were treatment labels, and use of trial-specific utility estimates and drug cost/resource use. It is therefore unclear why the AG has noted that: "*when the PURSUIT utility vector and resource use assumptions were inserted into the infliximab model, the results still did not coincide*".

7. Lack of clarity regarding the identification, selection and use of certain model parameters

Dosing schedules were obtained from SPCs (summary of product characteristics) and the frequency of background therapies was informed through usage in the ACT and PURSUIT trials. Unit costs were obtained from the BNF (drug acquisition and administration costs). Consultation with nine consultant gastroenterologists informed the resource use estimates (with associated costs obtained from the NHS Reference Costs and PSSRU). Costs relating to adverse events were obtained from NHS Reference Costs.

8. Complex implementation of the model

The AG report states that: "*the model employs 30 worksheets, many of which were locked as read-only*". MSD provided Bijal Joshi with the password for these sheets in an email dated 14th March 2014.

We agree that the model structure is relatively straightforward. However, we do not agree that the implementation in Excel is unnecessarily complex. All the input source data along with distributions are contained in only one worksheet (grey tab 'PSA Input'). Input values are sampled from these distributions and feed into the four Markov sheets, one for each of the competing interventions compared in the model (grey tabs 'MODEL arm TREATMENT A', 'MODEL arm TREATMENT D' etc.). The final three core model calculation sheets (grey tabs '1 way sensitivity analysis', 'PSA Output', and 'Prob CE') concern storage of the results of the one way and probabilistic sensitivity analysis. In short, the eight grey tabs are the core model sheets for input, calculation, and output. All the other sheets with blue tabs are user interfaces: sheets with dark blue tabs concern model input and feed into the single grey tab 'PSA Input' sheet; sheets with light blue tabs are output sheets and summarise the model output.

We attempted to create an Excel model that was as transparent as possible; however, for future purposes we will consider the AG feedback.

9. **Failure to undertake an incremental analysis**

Given that for all possible pairwise comparisons, incremental costs, QALYS, and ICERs were obtained with the model, the information was available for a full incremental interpretation. We did not present ICERs in a fully incremental way purely due to our focus on the comparison of the drug of interest from the manufacturer relative to competing interventions. The cost-effectiveness acceptability curves represent the joint decision-problem of four competing interventions, and are in line with the table of ICERs for a full incremental analysis.

10. **Inclusion of a PAS for golimumab which has not yet been agreed by the Department of Health**

We have explained our rationale for including the PAS within our analysis in point 1 ('Omission of PAS within AG modelling') on p.2 of this response. The text on p.196 of the AG report is misleading as it describes a scenario without the PAS which does not reflect the price of golimumab in reality.

Appendix 2: Issues and factual inaccuracies in the AG report (additional to those noted in our key points and response to critiques of MSD model)

Page	Issue/factual inaccuracy
4	The AG has suggested a potential bias arising from the PURSUIT design as those patients who received golimumab in PURSUIT-Maintenance were re-randomised golimumab-induction-responders (distinct from the treat-through designs of ACT and ULTRA where both responders and non-responders went on to receive active drug in the maintenance phase). However, one must also be conscious of another major difference in study design in terms of how both continuity of response and loss of response were defined. The primary endpoint of PURSUIT-Maintenance was continuous clinical response throughout 54 weeks, i.e. for a patient to achieve the primary endpoint, they were required to be in a state of clinical response at each 4-weekly assessment (not only at 2 or 3 time points), and a loss of response at any individual time point (detected via partial Mayo scoring and confirmed by colonoscopy) constituted treatment failure. This is in contrast to the designs of ACT and ULTRA. In ACT, patients were assessed for response at weeks 8, 30, and 54 (ACT 1 only), and patients in clinical response at each of these time points met the criteria for sustained clinical response. In ULTRA2, patients were assessed for response at weeks 8, 32, and 52, and a patient in clinical response at both weeks 8 and 52 met the criteria for sustained clinical response. In ULTRA2 inadequate response was defined as: (1) partial Mayo score equal to or above baseline score on 2 consecutive visits at least 14 days apart (for patients with a partial Mayo score of 4-7 at baseline); (2) partial Mayo score ≥7 on 2 consecutive visits at least 14 days apart (for patients with a partial Mayo score of 8 or 9 at baseline) - one could make the argument that it was therefore more difficult for a patient to achieve the primary endpoint in PURSUIT.
5	Data for infliximab biosimilars are included in the AG report, however, their inclusion is inappropriate because at the time of the technology appraisal guidance being issued biosimilars will not yet be available for use in UK clinical practice (prior to the Remicade patent expiry). In addition, there are no clinical data relating to the UC indication to support considerations of

	clinical effectiveness, and no list price to support considerations of cost-effectiveness.
10	The AG report states that: "*The risk of relapse and disease flares is increased by poor adherence to medication regimens*". We note that the 4-weekly administration of golimumab and the 8-weekly administration of infliximab in the maintenance setting could be expected to facilitate better adherence, compared with medicines that need to be administered more frequently.
14	The economic burden of UC is reported in Euros; it would be useful to convert this value to GBP.
25	Consistency in reviewer inclusion/exclusion decisions was assessed during first passing (title and abstract screening) but not for second passing (full-text screening). It is therefore unclear whether all inclusion decisions were consistent between reviewers.
32	The AG acknowledges that: "*None of the included RCTs applied Truelove and Witts' disease severity criteria*". This was expected and was discussed at the scoping workshop. Using the Mayo score to classify disease severity may be more appropriate given the evidence base.
33	It is inappropriate to include UC-SUCCESS in the AG systematic review as the population deviates from that specified in the final scope: the patients in UC SUCCESS were required to be either azathioprine-naïve or free from azathioprine treatment for at least 3 months before enrollment (only ~10% of patients had prior immunomodulator exposure at baseline so this represents a very different population to that in the TNF-α inhibitors licences and the scope of this appraisal).
95	In Figure 14, it is not acknowledged that infliximab is statistically significantly superior to adalimumab in clinical response/remission in induction (-0.52; 95% CrI: -10.03, 0.00) (in Figure 41 a credible interval meeting 0.00 is stated to be a significant result).
105	The AG acknowledges the limited data in the NMA and used a random effects model with an informative prior distribution for between-study heterogeneity (whereas we use a fixed effects model for the maintenance NMA). We note that where there is only one trial per connection in the network, the choice of the prior distribution for the between-study heterogeneity is directly incorporated in the posterior distributions of the treatment effects and is not updated based on the data. As such, the results of the NMA (i.e. 95% CrI and associated rankings) are arguably more influenced by the choice of prior distribution than by actual trial data. In this context we would like to emphasise the very wide intervals for remission in Table 16 and 17.
156	The AG report states that: there is "*a potential inconclusive benefit for intervention-treated patients compared with placebo*". It is not clear why this conclusion has been reached given the data from Sandborn *et al*. 2009 which indicates a 41% reduction in the risk of colectomy with infliximab relative to placebo (p. 74 of the AG report).
177	The MSD NMA includes data for unlicensed doses from the pivotal trials, however, only odds ratios relating to licensed doses were included in the MSD cost-effectiveness model.
180	The possible transitions in the model were limited by the available data for all competing TNF-α inhibitors. The duration of drug maintenance treatment is reflected with the time spent in the health states 'Response (pre-colectomy; maintenance)' and 'Remission (pre-colectomy; maintenance)' and is based on efficacy data. Upon losing remission patients transition to the health state 'Response (pre-colectomy; maintenance)'. Upon losing response, patients transition to the health state 'Relapse (relapse management)'. To ensure that the modeled duration of treatment with a TNF-α inhibitor matched the observed discontinuation rates of biologic treatment in the trial for golimumab, we had to assume that patients losing remission do not transition straight to 'Relapse (relapse management)'; the corresponding transition probability was set to 0. However, we would like to note that the model structure/programming would allow such transitions, which was the reason for the corresponding arrow in the figure.
184	The description of the golimumab induction dosing regimen in Table 46 is incorrect. Patients do not receive 2x200mg then 2x100mg; rather patients receive 2x100mg at week 0 and 1x100mg at

	week 2. Given that the cost per cycle is calculated for 8 weeks (patients will receive 2x100mg at week 0, 1x100mg at week 2, and 1x50mg or 100mg at week 8), the cost of £3051.88 is correct when the PAS for golimumab is applied (which it has not been elsewhere in the AG report).
189	The AG report states that: "*The DSAs indicate that the cost-effectiveness results for golimumab versus colectomy are sensitive to the utility values for pre-colectomy remission and pre-colectomy relapse*". This is not the conclusion presented in the tornado diagrams in the MSD submission.
214	The population considered in the AG model comprises those who have failed at least one prior therapy "*as reflected in the RCTs*": this does not reflect the RCT populations. Further, the licences for golimumab, infliximab, and adalimumab refer to 'moderately to severely active ulcerative colitis in adult patients who have had an inadequate response to conventional therapy including corticosteroids and 6-mercaptopurine (6-MP) or azathioprine (AZA), or who are intolerant to or have medical contraindications for such therapies' - meaning that the relevant population should include failure of two prior therapies.
215	In the AG model, patients in the conventional treatment group and those who have previously achieved but subsequently lost response to biologic therapy are assumed to continue receiving conventional therapy irrespective of whether response or remission is achieved. The rationale for continuing treatments which have been ineffective is unclear.
219	In Table 65 the evidence source for pre- and post-surgical utility states is given as Woehl *et al*. (reference 109). The same reference was (rightly) criticised in the MSD submission (p. 176, p.181 of the AG report); it was referenced by Tsai *et al*. but does not contain the relevant information. Incorrect referencing of the Woehl study also occurs on p. 225 and p.227 of the AG report.
220	The transition probabilities in Table 66 indicate that after patients fail on biologic treatment, they go onto conventional therapy. At this point the AG uses the placebo transitions and permits patients to move to remission/response as a consequence of being placed on conventional therapy despite the patient having already failed both biologic and conventional therapy.
222	The AG identified 6 potentially relevant studies for colectomy rates (Table 67), selecting Solberg *et al*. (2009) primarily due to its size. The cumulative colectomy rate after 10 years was 9.8% (95% CI: 7.4%, 12.4%). However, the population comprises newly-diagnosed UC patients who may be different to a moderate to severe population initiating treatment with a biologic. In fact, the average disease duration at baseline in the trials forming the basis for the economic model is ~6 years (Table 8, page 53). Furthermore, the Solberg study enrolled patients between 1990 and 1994, which suggests that the patients will not have received biologics during follow-up. As such, we have concerns around the use of this study to determine colectomy rates. Although limited data were reported from included trials for the outcome of time to surgical intervention, Sandborn *et al*. (2009) combined data from ACT1/2 and reported that the cumulative incidence of colectomy through 54 weeks was higher for the placebo group (17%) than for the infliximab group (10%). These infliximab estimates correspond to a cumulative incidence of colectomy of 19% at 2 years, 41% at 5 years, and 65% at 10 years, and a 6-month incidence rate of 0.053. This 0.053 rate is ~10 times larger than the 6-month rate of 0.0051 used by the AG. Overall, it appears that the colectomy rate has been underestimated by the AG. Our model assumed that within 1 year, 15%-20% patients are in the colectomy/post-colectomy health states. This was criticised for being higher than the proportion of patients undergoing surgery in RCTs (colectomy rates in the RCTs included in the systematic review were 0.7% to 5.8% in individual trial arms at ~1 year). However, reference to the RCT rates further underlines that the estimate used by the AG is too low.

228	In Table 72 the induction dosing regimens are not described correctly. When the PAS is applied to the golimumab cost, the cost for the induction phase is £2288.91 and the cost for the maintenance phase is £4976.34 irrespective of whether 50mg or 100mg golimumab is received.
229	The AG use only one UK clinician's estimates (Alan Lobo) to inform the non-biologic drug resource use estimates; we surveyed nine consultant gastroenterologists to determine the same and feel this provides a more comprehensive view of UK clinical practice. In addition, the costs (for induction) assumed in the AG model are lower than those assumed in the MSD model. As these costs were only included in the TNF-α inhibitor arms of the MSD model, our approach is more conservative in terms of the resulting ICERs.
230	The "*other UC health state resource use*" considered by the AG were informed by Tsai *et al*. (these estimates were used in previous infliximab submissions). These resource use estimates have been shown to be out of date; when we used a similar questionnaire with the nine consultant gastroenterologists there were changes from the original values.
235	Updated analyses are required which incorporate the approved PAS for golimumab.
247	The AG report notes that training for subcutaneous injection "*is associated with additional resource use and costs*". However, this cost would be expected to be very minimal, and patients receiving golimumab have the opportunity to be supported through 'Simply for Me'.

References

Bouguen G, Roblin X, Bourreille A, Feier L, Filippi J *et al*. Infliximab for refractory ulcerative proctitis. Aliment Pharmacol Ther. 2010; 31(11): 1178-1185. doi: 10.1111/j.1365-2036.2010.04293.x. Epub 2010 Mar 8.

Burns EM, Bottle A, Aylin P, Clark SK, Tekkis PP *et al*. Volume analysis of outcome following restorative proctocolectomy. Br J Surg. 2011; 98(3): 408-417. doi: 10.1002/bjs.7312. Epub 2010 Nov 11.

Byrne CM, Tan KK, Young JM, Selby W, Solomon MJ. Patient and clinician preferences for surgical and medical treatment options in ulcerative colitis. Colorectal Dis. 2014; 16(4): 285-292. doi: 10.1111/codi.12538.

de Silva S, Ma C, Proulx MC, Crespin M, Kaplan BS *et al*. Postoperative complications and mortality following colectomy for ulcerative colitis. Clin Gastroenterol Hepatol. 2011; 9(11): 972-980. doi: 10.1016/j.cgh.2011.07.016. Epub 2011 Jul 30.

Dias S, Welton N, Sutton A, Ades A. NICE DSU Technical Support Document 2: A generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials. Available at http://www.nicedsu.org.uk/Evidence-Synthesis-TSD-series(2391675).htm. Accessed on 24/02/2014.

Dignass A, Lindsay JO, Sturm A, Windsor A, Colombel JF *et al*. Second European evidence-based consensus on the diagnosis and management of ulcerative colitis part 2: current management. J Crohns Colitis. 2012; 6(10): 991-1030. doi: 10.1016/j.crohns.2012.09.002. Epub 2012 Oct 3.

Ferrante M, Karmiris K, Compernolle G, Ballet V, Vermeire S *et al*. Efficacy of adalimumab in patients with ulcerative colitis: restoration of serum levels after dose escalation results in a better long-term outcome. Gut 2011; 60 (Suppl 3) A72.

Frizelle FA. Surgical management of ulcerative colitis. Surgical Treatment: Evidence-Based and Problem-Oriented. Holzheimer RG, Mannick JA, editors. Munich: Zuckschwerdt; 2001.

Johnson P, Richard C, Ravid A, Spencer L, Pinto E *et al*. Female infertility after ileal pouch-anal anastomosis for ulcerative colitis. Dis Colon Rectum. 2004; 47(7): 1119-1126. Epub 2004 May 28.

Kaplan GG, McCarthy EP, Ayanian JZ, Korzenik J, Hodin R, Sands BE. Impact of hospital volume on postoperative morbidity and mortality following a colectomy for ulcerative colitis. Gastroenterology. 2008; 134(3): 680-687. doi: 10.1053/j.gastro.2008.01.004. Epub 2008 Jan 10.

National Institute for Health and Care Excellence. CG166: Ulcerative colitis: Management in adults, children and young people. Available at http://guidance.nice.org.uk/CG166, Accessed on 03/03/2014.

National Institute for Health and Care Excellence. Inflammatory bowel disease Quality Standard [GID-QSD70]. Available at http://www.nice.org.uk/Guidance/InDevelopment/GID-QSD70/Documents, Accessed on 04/08/2014.

Ørding Olsen K, Juul S, Berndtsson I, Oresland T, Laurberg S. Ulcerative colitis: female fecundity before diagnosis, during disease, and after surgery compared with a population sample. Gastroenterology. 2002; 122(1): 15-19.

Sandborn WJ, Rutgeerts P, Feagan BG, Reinisch W, Olson A *et al*. Colectomy rate comparison after treatment of ulcerative colitis with placebo or infliximab. Gastroenterology. 2009; 137(4): 1250-60; quiz 1520. doi: 10.1053/j.gastro.2009.06.061. Epub 2009 Jul 28.

Solberg IC, Lygren I, Jahnsen J, Aadland E, Høie O *et al*. Clinical course during the first 10 years of ulcerative colitis: results from a population-based inception cohort (IBSEN study). Scandinavian Journal of Gastroenterology – Supplement. 2009; 44(4): 431-440.

Tekkis PP, Fazio VW, Lavery IC, Remzi FH, Senagore AJ *et al*. Evaluation of the learning curve in ileal pouch-anal anastomosis surgery. Ann Surg. 2005; 241(2): 262-268.

Tsai HH, Punekar YS, Morris J, Fortun P. A model of the long-term cost effectiveness of scheduled maintenance treatment with infliximab for moderate-to-severe ulcerative colitis. Aliment Pharmacol Ther. 2008; 28(10): 1230-1239.

Turner D, Walsh CM, Steinhart AH, Griffiths AM. Response to corticosteroids in severe ulcerative colitis: a systematic review of the literature and a meta-regression. Clin Gastroenterol Hepatol. 2007; 5: 103-110.

Waljee A, Waljee J, Morris AM, Higgins PD. Threefold increased risk of infertility: a meta-analysis of infertility after ileal pouch anal anastomosis in ulcerative colitis. Gut. 2006; 55(11): 1575-1580. Epub 2006 Jun 13.

Woehl A, Hawthorne B, Morgan C, Punekar Y, McEwan P. The epidemiology and healthcare resource use in patients with Crohn's disease: a population based UK study. Value Health. 2007; 10(6): A355.