NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Premeeting briefing

Vedolizumab for treating moderately to severely active ulcerative colitis

This premeeting briefing presents:

- the key evidence and views submitted by the manufacturer, the consultees and their nominated clinical specialists and patient experts and
- the Evidence Review Group (ERG) report.

It highlights key issues for discussion at the first Appraisal Committee meeting and should be read with the full supporting documents for this appraisal.

Please note that this document includes information from the ERG before the manufacturer has checked the ERG report for factual inaccuracies.

Key issues for consideration

Clinical effectiveness

- Who would be offered vedolizumab in clinical practice? People whose treatment with TNF alpha inhibitors has failed and/or people who are TNF alpha naïve?
- How generalisable is GEMINI I to clinical practice in England in terms of the population (given the eligibility criteria were different between USA and non-USA sites, and there were only 2 UK sites), the treatments that people received, and how clinical response to treatment was measured (using the Mayo score)?
- In GEMINI I people responding to vedolizumab were followed up for 1 year. How long would a person receive treatment with vedolizumab in clinical practice? Would a person in remission continue to take vedolizumab over the longer term? It was assumed in the company's model that vedolizumab and TNF alpha

inhibitors would be used for up to one year, after which treatment would be discontinued irrespective of response. Does this reflect clinical practice?

- In clinical practice, if a patient did not respond to treatment with vedolizumab at week 6, would treatment continue? At what point would a decision be made to stop treatment? 6 or 10 weeks (as defined in the marketing authorisation)?
- The company presented clinical data from GEMINI I for a subgroup of people who had not had prior treatment with a TNF alpha inhibitor, and a subgroup of people in whom treatment with an TNF alpha inhibitor had failed. The company noted that the results from these subgroups were consistent with the results from the whole population and the ERG noted that GEMINI I was not powered to detect statistically significant differences between subgroups. Is there any evidence that the characteristics defining these subgroups influence the effectiveness and/or cost effectiveness of vedolizumab?
- The relative effectiveness of vedolizumab compared with the TNF alpha inhibitors (adalimumab, infliximab and golimumab) was estimated using a mixed treatment comparison. The company stated that the clinical trials of infliximab and golimumab that were included in the mixed treatment comparison, only included a population who had not had a prior TNF alpha inhibitor. In its cost effectiveness analysis the company only presented a comparison of the cost effectiveness of vedolizumab with the TNF alpha inhibitors in this population. Would the relative effectiveness of vedolizumab compared with the TNF alpha inhibitors be expected to differ in people who had and had not had prior treatment with a TNF alpha inhibitor? Can a robust estimate of clinical and cost effectiveness of vedolizumab compared with the TNF alpha inhibitors be made for the whole population given the available data?
- The company noted that people who were likely to need surgery were not eligible to participate in GEMINI I and that there were no data for the effectiveness of vedolizumab relative to surgery. It further stated that surgery is generally reserved for people with severe disease not amenable to medical management, those intolerant to medical therapies, or those with certain disease-related complications. The ERG noted that surgery would not be an acceptable treatment option for all patients. Should surgery be considered a comparator?

Cost effectiveness

- The ERG noted that in the company's base case people who were in remission after surgery had a worse quality of life than people with moderate to severe ulcerative colitis. The ERG considered this to be implausible. The company and the ERG presented scenarios using different utility values, which all had a effect on the ICER (Table 1). In particular, in all of the company's sensitivity analyses and the ERG scenario in which the Woehl utlility values were used, surgery became more effective (was associated with more QALYs) than vedolizumab. To what extent does having surgery affect quality of life? Which utility values, of those presented by the company and ERG, are the most appropriate?

Table 1. Utility values used in the scenarios presented by the company and the ERG and the corresponding ICERs

Scenario	Health state						ICER*	
	Remission	Mild	Moderate to Severe	Surgery	Post surgery remission	Post surgery comp.	vs. CT	vs. surgery
Company base case	0.86	0.80	0.68	0.42	0.60	0.42	£33,297	Surgery dominated ‡
Company's sensitivity analyses								
Punekar and Hawkins	0.88	assumed 0.76	0.42	0.42	0.60	0.42	£17,857	£117,134 (SW)
Arseneau et al.	0.79	assumed 0.76	0.32	0.32	0.63	0.49	£18,008	£26,438 (SW)
Tsai et al.	0.88	0.76	0.42	0.61	0.61	0.55	£18,627	£46,733 (SW)
ERG's sensitivity analyses (ICERs also include model corrections)								
Woehl et al	0.87	0.76	0.41	0.41	0.71	0.54	£17,140	£21,881 (SW)
Swinburn et al	0.91	0.80	0.55	0.55	0.59	0.42	£15,267	Surgery dominated‡

*for ITT population
Comp: complications CT: Conventional therapy; ‡ Surgery is dominated by vedolizumab meaning it is more costly and less effective; SW: South west quadrant, (in these scenarios surgery is more effective and more costly than vedolizumab- the ICER shows the money saved with vedolizumab for the QALYs lost)

- The ERG presented a scenario in which the costs of caring for a stoma were included for people who had this surgical procedure. This reduced the ICER for vedolizumab compared with conventional therapy in the whole population to £19,630 per QALY gained from the company base case of £33,297 per QALY gained. Have the costs associated with surgery and post surgery care been adequately captured in the model? Should stoma care be included? What is the cost of stoma care?

- The ERG presented a scenario in which the surgery health state transition probabilities were amended, because it considered the company's assumptions resulted in an overestimation of the costs and underestimation of the benefits of surgery. This scenario did not favour vedolizumab, and increased the ICER substantially across all populations for vedolizumab compared with the next most effective comparator. Which of the health state transitions can be considered more appropriate, those presented by the ERG or that presented by the company?
- The company stated that vedolizumab is innovative because it has a different mechanism of action compared with conventional therapies and TNF alpha inhibitors, and because it is targeted to the gut so avoiding the effects on other parts of the body. Patient groups stated that nonspecific anti-inflammatory agents, such as corticosteroids, are associated with adverse effects for many people. In GEMINI people reduced their dose of corticosteroids if they had a response to vedolizumab. Have all of the benefits of vedolizumab been captured in the QALY calculation?

1 Remit and decision problems

1.1 The remit from the Department of Health for this appraisal was: To appraise the clinical and cost effectiveness of vedolizumab within its licensed indication for treating moderately to severely active ulcerative colitis in adults who are intolerant of, or whose disease has had an inadequate response or loss of response to conventional therapy or a tumour necrosis factor-alpha antagonist.

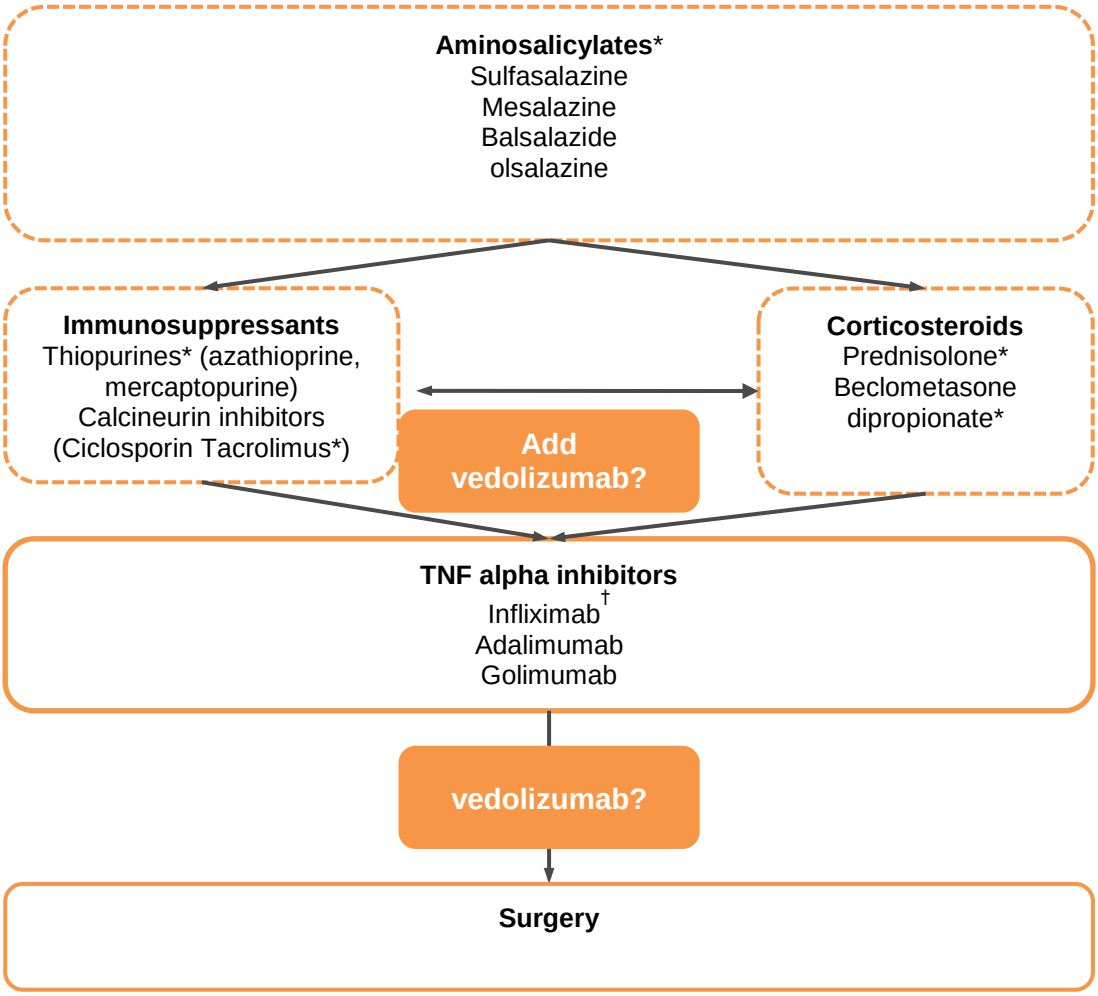
Table 2 Decision problem

	Final scope issued by NICE	Decision problem addressed in the submission	Comments from the manufacturer	Comments from the ERG
Pop.	Adults with moderately to severely active ulcerative colitis (excluding those with acute severe ulcerative colitis that is a medical emergency and requires inpatient treatment) who are intolerant of, or whose disease has had an inadequate response or loss of response to conventional therapy (immunosuppressants and/or corticosteroids) or a TNF-alpha inhibitor.		This population is in line with the VEDO label and the pivotal study, GEMINI I.	GEMINI I included only 2 UK sites. However, the population reflected the population defined in the licensed indication and the final NICE scope.
Int.	Vedolizumab		No comments	No comments
Com.	Established clinical management without vedolizumab, which may include a combination of aminosalicylates (sulfasalazine, mesalazine, balsalazide or olsalazine), corticosteroids (beclometasone, budesonide, hydrocortisone or prednisolone), thiopurines (mercaptopurine or azathioprine), calcineurin inhibitors (tacrolimus or ciclosporin), TNF-alpha inhibitors (infliximab, adalimumab or golimumab) and surgical intervention.	• Conventional therapy, as defined in the GEMINI I study (aminosalicylates, corticosteroids, immunomodulators, antibiotics, probiotics and antidiarrheal) and those used in UK clinical practice based on the UK IBD audit • TNF-alpha antagonists licensed in the UK for ulcerative colitis (infliximab, adalimumab and golimumab).	Surgical intervention is not included as a comparator. Reasons included: • Different populations eligible for vedolizumab compared with surgery. Surgery generally reserved for people with severe disease not amenable to medical management, those intolerant to medical therapies, or those with certain disease-related complications. • In the GEMINI study patients were excluded if there was an anticipated requirement for major surgery. • Lack of comparable evidence to vedolizumab.	• Surgical intervention was not included as a comparator in the company's review of clinical effectiveness; it was however included in the health economic analysis. Surgery is not acceptable to some patients. • Calcineurin inhibitors were not included in the company's clinical effectiveness review, network meta-analysis or health economic analysis. There is limited data for the use of calcineurin inhibitors compared with biologics.

Out.	The outcome measures to be considered include: • mortality • measures of disease activity • rates of and duration of response, relapse and remission • rates of hospitalisation • rates of surgical intervention • time to surgical intervention • adverse effects of treatment (including leakage and infections following surgery) • health-related quality of life.			Data on relapse rates were not reported in the company's submission.
SubG.	People who have been previously treated with one or more TNF-alpha inhibitors and people who have not received prior TNF-alpha inhibitor therapy	• Anti-TNF naïve population • Anti-TNF failure population • Mixed population		

2 The technology and the treatment pathway

Figure 1 Treatment pathway for people with moderately and severely active ulcerative colitis



Pathway based on NICE Clinical Guideline 166, British Society of Gastroenterology Guidelines for the management of inflammatory bowel disease, clinical expert statements, and the company's submission. Throughout the pathway treatments may be added to, or replace, the previous treatment.
*Recommended by NICE Clinical Guideline 166
†Not recommended for subacute manifestations of ulcerative colitis (NICE technology appraisal guidance 140)
Dashed lines represent 'conventional therapy'

2.1 Vedolizumab (Entyvio, Takeda) is a humanised monoclonal antibody. It targets the $\alpha_4\beta_7$ integrin which is expressed in certain white blood cells that are found in the gut. $\alpha_4\beta_7$ integrin is responsible for recruiting these

cells to inflamed bowel tissue. Vedolizumab therefore specifically targets the gut. Vedolizumab is administered by intravenous infusion.

2.2 There is currently no cure for ulcerative colitis. Treatment aims to relieve symptoms during a flare-up and then maintain remission (see Ulcerative colitis: management in adults, children and young people [NICE clinical guideline [CG]166]) and can include a combination of anti-inflammatory drugs, immunosuppressants, corticosteroids or TNF alpha inhibitors (see **Error! Reference source not found.**).

2.3 NICE Technology Appraisals have given recommendations for, or are currently considering, the following technologies for ulcerative colitis:

- A multiple technology appraisal of infliximab, adalimumab and golimumab for moderately to severely active ulcerative colitis after the failure of conventional therapy (including review of TA140 and TA262) is ongoing, with guidance anticipated to be published in January 2015.
- NICE was unable to make recommendations for adalimumab for treating moderately to severely active ulcerative colitis because the company did not make an evidence submission (Adalimumab for the treatment of moderate to severe ulcerative colitis (NICE technology appraisal guidance 262).
- NICE technology appraisal guidance 140 Infliximab for subacute manifestations of ulcerative colitis does not recommend infliximab for treating 'subacute' manifestations of moderately to severely active ulcerative colitis
- NICE technology appraisal guidance 163 Infliximab for acute exacerbations of ulcerative colitis recommends infliximab for people with acute, severely active ulcerative colitis only if ciclosporin is considered inappropriate for them or they are taking part in a research study

Table 3 Technology and biological comparators

	Vedolizumab	Infliximab	Adalimumab	Golimumab
Marketing authorisation	For the treatment of adult patients with moderately to severely active ulcerative colitis who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy or a tumour necrosis factor-alpha antagonist. Summary of product characteristics states that concomitant use of biologic immunosuppresents is not recommended	Infliximab, adalimumab and golimumab have the same marketing authorisation in the UK for the 'treatment of moderately to severely active ulcerative colitis in adult patients who have had an inadequate response to conventional therapy including corticosteroids and mercaptopurine or azathioprine, or who are intolerant to or have medical contraindications for such therapies'.		
NICE recommendation	-	TA140: Not recommended for subacute manifestations of UC	TA262: Unable to appraise due to non-submission by manufacturer	-Subject to on-going appraisal
Administration	Intravenous infusion	Intravenous infusion	Subcutaneous injection	Subcutaneous injection

	Vedolizumab	**Infliximab**	**Adalimumab**	**Golimumab**
Frequency of administration	300 mg at 0, 2 and 6 weeks, then every 8 weeks thereafter. Increased frequency to every 4 weeks allowed for people who show decrease in response.	Initial dose of 5 mg/kg at 0, 2 and 6 weeks, then every 8 weeks thereafter.	Initial dose of 160 mg at week 0 and 80 mg at week 2, then 40 mg every other week.	Initial dose of 200 mg, followed by 100 mg at week 2, then 50 mg or 100 mg (depending on body weight) every 4 weeks.
Setting	Hospital – usually day patient	Hospital – usually day patient	Can be administered at home by patient or family member	Hospital – usually day patient
Acquisition cost (excluding VAT)	Basic NHS list price: £2050 per vial [redacted]	£419.62 for a 100 mg vial	£352.14 for a 40 mg pre-filled pen or pre-filled syringe, or a 40 mg/0.8-mL vial	£762.97 for a 50 mg pre-filled pen/syringe or pre-filled syringe and £1525.94 for a 100 mg.. Merck Sharp & Dohme has agreed a patient access scheme to make the 100 mg dose available to the NHS at the same cost as the 50 mg dose.

	Vedolizumab	**Infliximab**	**Adalimumab**	**Golimumab**
Number of cycles	Until relapse, intolerance or discontinuation due to side effects. Reconsider continuation if response is not observed by week 10.	The MA states that continued therapy should be reconsidered in people who have not had a response in 12-14 weeks with infliximab or golimumab or within 2-8 weeks with adalimumab.		
Total cost (based on list price)	[redacted]	The cost of infliximab induction therapy is £5035; the monthly cost of infliximab maintenance therapy is £840, assuming the average person weights 77 kg	The cost of adalimumab induction therapy is £2113; the monthly cost of adalimumab maintenance therapy is £704	The cost of golimumab induction therapy is £2289; the monthly cost of golimumab maintenance therapy is £763.Assuming the average person weighs 80 kg

3 Comments from consultees

The Comments from consultees included 2 submissions from the British Society of Gastroenterology (BSG), 1 submission from a Commissioning Support Unit, 1 submission from UK Clinical Pharmacy Association (UKCPA), and 1 submission from a patient organisation (Crohn's and Colitis UK).

3.1 Professional groups described the population in the decision problem. They stated that people with moderate to severely active ulcerative colitis have generally not responded adequately to therapy with aminosalicylates (such as mesalazine) or immunosuppressants (such as the thiopurines azathioprine or mercaptopurine), or are intolerant to these drugs (see Figure 1). They are likely to require repeated courses of corticosteroids, or may be steroid-dependent or steroid-refractory. Further treatment options for this population currently include ciclosporin, or TNF antagonists (infliximab or adalimumab). Ciclosporin is not widely used because of lack of experience of this drug in clinical practice and concerns about its toxicity. Long term funding for the biologic TNF alpha inhibitors infliximab and adalimumab can be difficult to obtain in the NHS because infliximab was not recommended by NICE and adalimumab was not appraised. Further options such as methotrexate and thioguanin are available but their use is not well documented. Surgery may be offered as an alternative to long-term steroid therapy for this population, but can be associated with severe adverse events.

3.2 Patient groups described their experience with currently available treatments. They stated that conventional therapies for ulcerative colitis are suboptimal. TNF alpha inhibitors produce remission in approximately one-third of patients, with many losing response over time, and a significant proportion of patients continue to experience flares or chronic symptoms. Nonspecific anti-inflammatory agents such as corticosteroids are associated with adverse effects for many people. Adverse reactions and long-term complications of all currently available treatments are of great concern to patients. Patient groups noted that vedolizumab has a

novel mechanism specific to the gut meaning there are reduced levels of systemic immunosuppression, which may result in fewer subsequent adverse effects. However, adverse events would need to be monitored in clinical practice. Patient groups stated that vedolizumab is a clear alternative where people are reluctant to use ciclosporin or have surgery, offering the potential for long-term therapy with the possibility of lower toxicity.

3.3 Professional groups noted that response to treatment was defined in the clinical trial using the Mayo score (of note this is also used in the cost effectiveness model submitted by the company for this appraisal). However, the Mayo score is not widely used in the UK because the NHS does not have the capacity for the required colonoscopies and sigmoidoscopies, and routine monitoring using colonoscopy may not be acceptable to patients.

3.4 Professional groups stated that vedolizumab treatment will be used exclusively in secondary care, and noted that it is given by intravenous infusion and requires day-case hospital attendance and the use of infusion facilities and nursing support. The monitoring for side-effects of therapy would be no different to that needed for other currently used immunosuppressive therapies.

3.5 Professional groups commented that, based on the use of biological drugs (such as TNF alpha inhibitors) in managing Crohn's disease, increased age is associated with an increased risk of infection with biological drugs. They therefore suggested considering people aged over 50 years as a subgroup and stated that the course of the disease for people aged over 50 years is different.

4 Clinical-effectiveness evidence

Overview of the clinical trials

4.1 Evidence for the clinical efficacy of vedolizumab in the company's submission was from GEMINI I, and evidence for the safety of vedolizumab was from GEMINI I as well as 3 additional studies, all identified through a systematic literature review.

GEMINI I

4.2 GEMINI I was a study in adults with moderately to severely active ulcerative colitis who were intolerant of, or whose disease had an inadequate response or loss of response to immunosuppressants, corticosteroids or TNF alpha inhibitors. It was conducted in 34 countries at 211 centres, 2 of which were based in the UK. The study consisted of separate induction and maintenance trials:

- **Induction trial (double blind cohort):** The induction trial included 374 people randomised (3:2) to have double-blind vedolizumab (300 mg) or placebo intravenously at weeks 0 and 2, concomitantly to conventional therapy (described in section 4.3 below). People were assessed for clinical response (the primary outcome) at six weeks. Clinical response was determined using sigmoidoscopy and defined as a reduction in the Mayo score of at least 3 points and a decrease of at least 30% from baseline, with an accompanying decrease in the rectal bleeding subscore of at least 1 point or an absolute rectal bleeding subscore of up to 1 point. Secondary outcomes included clinical remission (Mayo score of up to 2 points and no individual subscore greater than 1 point) and mucosal healing (defined as an endoscopic subscore of 1 point or less)
- **Induction (open label cohort):** an additional 521 people received open label vedolizumab (300mg) at weeks 0 and 2. People were assessed for clinical response (as defined above) at six weeks.

- **Maintenance trial:** People who had received vedolizumab and had a clinical response at week 6, from either induction cohort, could progress into the maintenance trial. 373 people were enrolled onto the maintenance trial and were randomised (1:1:1) to vedolizumab every 8 weeks (n=122), every 4 weeks (n=125) or placebo every 4 weeks (n=126), for up to 52 weeks. The primary outcome for the maintenance trial was clinical remission at week 52 (remission defined as above). Secondary outcome measures included durable clinical response (response at weeks 6 and 52), durable clinical remission (remission at weeks 6 and 52), mucosal healing at week 52 and glucocorticoid free remission at week 52 in patients receiving glucocorticoids at baseline.

Additionally, data continued to be collected for people who did not did not have a clinical response at 6 weeks in the induction study and from the induction open label cohort. These people continued on their assigned study drug (vedolizumab or placebo) and were followed until week 52.

4.3 GEMINI I included people with a baseline Mayo score of 6 to 12. The Mayo score includes assessment of stool frequency, rectal bleeding, an endoscopic assessment, and a global assessment by a clinician. The Mayo score ranges from 0 to 12, with scores of 6 or greater representing moderate to severe disease. People in GEMINI I had also demonstrated over the previous 5 years an inadequate response to, or intolerance of, at least one of: an immunosuppressant (oral azathioprine or mercaptopurine), a TNF alpha inhibitor (infliximab), or a corticosteroid (prednisone). During the trial people in both treatment arms could take mesalazine, up to 30 mg prednisone (or equivalent daily) and immunosuppressants. People taking corticosteroids had a reduced dose after week 6. In total 37.1% of people received concomitant corticosteroids only, 17.8% of people received concomitant immunosuppressant drugs only and 16.6% of people received concomitant corticosteroids and immunosuppressants. In the induction and maintenance trials, people were randomised according to 3 stratification factors: 1) concomitant glucocorticoids (corticosteroids) use,

2) concomitant immunosuppressant use, and 3) prior TNF alpha inhibitors use. In the maintenance trial people were also stratified by whether they were part of the double blind cohort 1 in the induction phase or the open label cohort 2. Across all study groups, mean age was 40.3 years, mean disease duration was 6.9 years, mean baseline Mayo score was 8.6, mean use of TNF alpha inhibitors before study enrolment was 48.2%, and mean failure of TNF alpha inhibitor therapy was 41.0% (by inadequate response [48.0%], loss of response [38.4%], or intolerance [13.6%]). Full details of the baseline characteristics of people in GEMINI I can be found in section 6.3.4 of the company's submission and table 14 page 84.

4.4 The number of people who discontinued treatment before the end of the induction phase was: 7 (3%) of those who received vedolizumab in cohort 1, 36 (7%) of those who received vedolizumab in cohort 2 and 14 (9%) of those who received placebo in cohort 1. The main reason for discontinuation was lack of efficacy. During the maintenance phase 45 (37%) people receiving vedolizumab every 8 weeks, 41 (33%) people receiving vedolizumab every 4 weeks and 78 (62%) people receiving placebo discontinued prematurely, mostly due to lack of efficacy or disease related adverse events.

ERG comments

4.5 The ERG commented on the baseline characteristics of GEMINI I, and stated that there were no relevant differences between the treatment arms during the induction or maintenance periods. However, there were differences between US and non-US sites in entry criteria (failure of an immunomodulator or TNF alpha inhibitor was required in the US, whereas elsewhere corticosteroid failure was sufficient for entry), and the protocol for concomitant immunosuppressant use during the study (immunosuppressant use was discontinued at week 6 in the US but continued elsewhere). The ERG commented that it was unclear how these differences may impact the results of the trial.

4.6 The ERG considered the proportion of people who discontinued from the trial. It noted that discontinuation during the induction phase was 6% and during the maintenance phase was 44%. The ERG noted that the company had presented an intention-to-treat (ITT) analysis and assumed that all people who discontinued treatment had not met the primary endpoint. The ERG stated that in general the validity of the study may be threatened if the proportion of people discontinuing is over 20%, and considered the disproportionate discontinuation rates observed in the maintenance phases was a serious threat to the validity of GEMINI I.

Clinical trial results – ITT population

Induction trial efficacy results

4.7 Analyses from the company's submission used the ITT population (unless stated otherwise). In the induction trial in GEMINI I (double blind cohort 1), vedolizumab was associated with statistically significantly greater rates of clinical response, clinical remission and mucosal healing compared with placebo (Table 4). People receiving vedolizumab in the open label cohort 2 had similar rates of response and remission to people receiving vedolizumab in the double-blind cohort 1; of the 521 people in the open-label vedolizumab cohort 2, 231 (44.3%) had a clinical response, 100 (19.2%) had clinical remission and 191 (36.7%) had mucosal healing at week 6.

Table 3 Efficacy endpoints at week 6 in induction trial cohort 1 (double-blinded) of GEMINI I, adapted from table 18, p100 of the company's submission

Study endpoint	Vedolizumab (n=225)	Placebo (n=149)	Percentage difference (95% CI)†	p-value
Response* – n (%)	106 (47.1%)	38 (25.5%)	21.7% (11.6 to 31.7)	<0.001
Remission – n (%)	38 (16.9%)	8 (5.4%)	11.5% (4.7 to 18.3)	0.001
Mucosal healing – n (%)	92 (40.9%)	37 (24.8%)	16.1% (6.4 to 25.9)	0.001

CI, confidence interval
*Primary endpoint
†Percentage differences were adjusted for 2 stratification factors: concomitant use or non-use of glucocorticoids, and concomitant use or non-use of immunosuppressive agents or prior use or non-use of anti-TNFs.

Maintenance trial efficacy results

4.8 Table 5 presents the trial efficacy outcomes for each arm of the maintenance trial. People who had vedolizumab maintenance, either every 4 weeks or every 8 weeks, were statistically significantly more likely to achieve clinical remission at week 52 compared to those who had placebo. In addition, vedolizumab maintenance treatment was associated with statistically significantly higher rates of durable clinical response (response at both week 6 and 52), durable clinical remission (remission at both week 6 and 52), mucosal healing (at week 52), and glucocorticoid-free remission (remission at week 52 with no concomitant corticosteroid use) compared with placebo. The company noted that there were no clear differences in clinical effectiveness between vedolizumab if taken every 8 compared with every 4 weeks.

Table 5 Efficacy endpoints for the maintenance trial of GEMINI I, adapted from table 19, p102 of the company's submission

Study endpoint	8 week vedo (n=122)	4 week vedo (n=125)	Placebo (n=126)	Between group percentage difference	
				8 week vedo vs. placebo (95% CI)	4 week vedo vs. placebo (95% CI)
Primary outcome					
Remission at week 52 – n (%)	51 (41.8%)	56 (44.8%)	20 (15.9%)	26.1% (14.9–37.2) **p<0.001**	29.1% (17.9–40.4) **p<0.001**
Secondary outcomes					
Durable clinical Response – n (%)	69 (56.6%)	65 (52.0%)	30 (23.8%)	32.8% (20.8–44.7) **p<0.001**	28.5% (16.7–40.3) **p<0.001**
Durable clinical Remission – n (%)	25 (20.5%)	30 (24.0%)	11 (8.7%)	11.8% (3.1–20.5) **p=0.008**	15.3% (6.2–24.4) **p=0.001**
Mucosal healing at week 52 – n (%)	63 (51.6%)	70 (56.0%)	25 (19.8%)	32.0% (20.3–43.8) **p<0.001**	36.3% (24.4–48.3) **p<0.001**
Glucocorticoid-free remission at week 52 – n (%) †	22 (31.4%)	33 (45.2%)	10 (13.9%)	17.6% (3.9–31.3) **p=0.01**	31.4% (16.6–46.2) **p<0.001**

CI, confidence interval; Vedo, vedolizumab
† Glucocorticoid-free remission was analysed for those who recieved oral glucocorticoids at baseline: vedolizumab every 8 weeks, n=70; vedolizumab every 4 weeks, n=73; placebo, n=72.

ERG comments

4.9 The ERG commented that the long-term efficacy and safety of vedolizumab, and the optimum duration of therapy remain unclear because in the GEMINI I trial people only received vedolizumab for up to 52 weeks, and the extension study to GEMINI I is ongoing. The ERG commented that there are no data on strategies for withdrawal of vedolizumab in people receiving it to maintain response or remission.

4.10 The ERG commented that establishing mucosal healing needs histologic data, which were not reported by the company. How mucosal healing had been confirmed was not reported in the company submission and therefore the clinical relevance of these data was uncertain.

4.11 The ERG noted that the definition of corticosteroid free clinical remission in the company's submission did not pre-specify the minimum duration of time over which a patient had to be corticosteroid-free. This time period is needed to show the clinical relevance of the endpoint, and therefore the clinical relevance of this endpoint was not clear.

Clinical trial results – subgroups

4.12 The company carried out a range of pre-specified sub-group analyses and stated that there was consistency in the effect of vedolizumab across different subgroups according to demographic factors, disease characteristics, and failure of previous treatments. The company stated that in the maintenance trial the efficacy of vedolizumab was not substantively affected by concomitant use of corticosteroids or immunosuppressants. See response to clarification A25 for these data

4.13 The company presented further data from pre-specified subgroup analyses of the following:

- **A subgroup without prior TNF alpha inhibitor exposure**
- **A subgroup in whom prior TNF alpha inhibitor treatment had failed** (that is, people who were intolerant of, or whose disease had an inadequate response or loss of response to a TNF-alpha inhibitor).

In both subgroups there was a higher proportion of people had a clinical response or clinical remission with vedolizumab than with placebo which was consistent with the results for the whole population.

Table 6 GEMINI1 efficacy endpoints at week 6 in induction study from ERG report table 7 page 32 (also includes data from the company's submission page 100, the clinical study report and the NEJM publication for GEMINI I)

Study endpoint	**Subgroup without prior TNF alpha inhibitor exposure**		
	Vedolizumab (n= 130)	Placebo (n=76)	Percentage difference from placebo (95% CI)
Clinical response, No. (%)	69 (53.1)	20 (26.3)	26.8 (13.7 to 39.9)
Clinical remission, No. (%)	30 (23.1)	5 (6.6)	16.5 (2.4 to 30.2)
Mucosal healing, No. (%)	64 (49.2)	19 (25.0)	24.2 (11.2 to 37.2)
	Subgroup in whom prior TNF alpha inhibitors have failed		
	Vedolizumab (n=82)	Placebo (n=63)	Percentage difference from placebo (95% CI)
Clinical response, No. (%)	32 (39.0)	13 (20.6)	18.4 (3.9 to 32.9)
Clinical remission, No. (%)	8 (9.8)	2 (3.2)	6.6 (-9.8 to 22.8)
Mucosal healing, No. (%)	25 (30.5)	13 (20.6)	9.9 (-4.3 to 24.0)

Table 7 Results at week 52 (maintenance phase) for subgroups with prior anti-TNF failure and without anti-TNF exposure, adapted from table 20, p103 of the company's submission

Study endpoint	**Subgroup without prior TNF alpha inhibitor exposure**				
	8 wk vedo (n=43)	4 wk vedo (n=40)	Placebo (n=38)	Difference vs. placebo (95% CI)[a]	
				8 wk vedo	4 wk vedo
Clinical remission (%)	45.8	47.9	19.0	26.8 (12.4–41.2)	29.0 (14.6–43.3)
Durable clinical response (%)	65.3	56.2	26.6	38.7 (24.0–53.4)	29.6 (14.6–44.6)
	Subgroup in whom prior TNF alpha inhibitors have failed				
	8 wk vedo (n=72)	4 wk vedo (n=73)	Placebo (n=79)	Difference vs. placebo (95% CI)	
				8 wk vedo	4 wk vedo
Clinical remission (%)	37.2	35.0	5.3	31.9 (10.3–51.4)	29.7 (7.4–49.4)
Durable clinical response (%)	46.5	42.5	15.8	30.7 (11.8–49.6)	26.7 (7.5–45.9)

CI, confidence interval; TNF, tumour necrosis factor; wks, weeks
[a]Confidence interval for difference from placebo. Although these endpoints were pre-specified, p-values are not provided because multiple testing adjustments were not made.

ERG comments

4.14 The ERG commented that the results of all subgroups analyses should be interpreted with caution because the numbers of people in each subgroup were small and the study was not powered for these assessments. This included comparing the 4-weekly and 8-weekly doses of vedolizumab, and the subgroup analyses relating to prior use of TNF alpha inhibitors.

Patient reported outcomes in GEMINI I

Health related quality of life

4.15 At week 6 of induction treatment, vedolizumab was associated with statistically significantly greater improvements from baseline in health-related quality of life compared with placebo, as assessed by the Inflammatory Bowel Disease Questionnaire (IBDQ) total score, the EQ-5D and the EQ-5D visual analogue scale (VAS) scores. These were clinically meaningful improvements, based on pre-specified criteria for minimally important differences. Vedolizumab induction treatment was also associated with statistically significantly greater improvements in the SF-36 physical and mental component summary scores at week 6 compared with placebo. In the maintenance trial vedolizumab every 4 weeks and every 8 weeks resulted in significantly greater improvements in health-related quality of life compared with placebo, as measured by differences in the IBDQ score from baseline to week 30 and week 52. These differences were considered to be clinically meaningful, based on the minimally important differences. The SF-36 physical and mental component scores at weeks 30 and 52 were higher for vedolizumab every 4 weeks and every 8 weeks compared with the placebo group, although statistical significance was only reached at week 52. Both vedolizumab groups were associated with improvements in health-related quality of life at weeks 30 and 52 as assessed by the EQ-5D and EQ-5D VAS, although

statistical significance was only reached with the EQ-5D VAS score at week 52 (table 8).

Table 8 Changes in health-related quality of life from baseline at week 6 of induction therapy (GEMINI I), table adapted from tables 25 and 26, p107 and p109-111 of the company's submission. For EQ-5D VAS and SF-36 scores see tables in company's submission

Induction			
Endpoint	**Adjusted mean change from baseline (95% CI)[b]**		**Difference in adjusted change from baseline vs. Placebo, mean (95% CI)**
	Placebo (n=144)	**Vedolizumab (n=219)**	
IBDQ Total Score[a]	10.9 (5.5 to 16.3)	28.9 (24.5 to 33.2)	18.0* (11.0 to 24.9)
EQ-5D Score[a]	–0.0 (–0.3 to 0.2)	–0.5 (–0.7 to –0.4)	–0.5* (–0.7 to –0.2)

Maintenance					
Endpoint	**Placebo**	**Vedo every 4 Weeks**		**Vedo every 8 Weeks**	
		Adjusted mean change from baseline (95% CI)[b]	Difference in adjusted change from baseline vs. placebo, mean (95% CI)	Adjusted mean change from baseline (95% CI)[b]	Difference in adjusted change from baseline vs. placebo, mean (95% CI)
Week 30 IBDQ	n=98	n=110		n=99	
	35.8 (28.4–43.3)	46.9 (39.9–53.9)	11.0 (0.8–21.2)	52.9 (45.6–60.3)	17.1 (6.6–27.6)
Week 52 IBDQ	n=58	n=86		n=77	
	32.8 (24.6–41.0)	58.5 (51.8–65.3)	25.7* (15.1–36.3)	58.9 (51.8–66.0)	26.1* (15.2–36.9)
Week 30 EQ-5D score	n=99	n=109		n=98	
	–0.9 (–1.2 to –0.6)	–0.9 (–1.2 to –0.7)	–0.0 (–0.4 to 0.4)	–1.2 (–1.5 to –0.9)	–0.3 (–0.7 to 0.1)
Week 52 EQ-5D score	n=58	n=86		n=76	
	–0.6 (–1.0 to –0.3)	–1.2 (–1.5 to –0.9)	–0.6 (–1.1 to –0.1)	–1.2 (–1.5 to –0.9)	–0.6 (–1.1 to –0.1)

Abbreviations: CI, confidence interval; IBDQ, Inflammatory Bowel Disease Questionnaire;

[*] Statistically significant result, p value not reported.

[a] Higher IBDQ indicates improvements in HRQOL; lower EQ-5D scores indicate improvements in HRQOL.

[b] Mean changes were adjusted for individual baseline measurements.

Other exploratory analyses presented by the company

4.16 The company presented exploratory analyses to assess for delayed response among people whose disease had not responded to treatment at week 6 and who remained in the study receiving vedolizumab or placebo every 4 weeks. Clinical response was assessed by the partial Mayo score (that is, the Mayo score without the sigmoidoscopy subscore). Response was defined as a reduction of at least 2 points and a decrease of at least 25% from baseline, with an accompanying decrease in the rectal bleeding subscore of at least 1 point or an absolute rectal bleeding subscore of up to 1 point). Response was achieved at week 10 and week 14 by greater proportions of people who had vedolizumab (32% [102/322] and 39% [126/322], respectively) than placebo (15% [12/82] and 21% [126/322], respectively). These analyses formed the basis of the recommendation in the label for vedolizumab which states that continued therapy for people with ulcerative colitis should be carefully reconsidered if no evidence of therapeutic benefit is observed by week 10.

ERG comments

4.17 The ERG commented that the additional *post hoc* delayed response analysis should be interpreted with caution because dosing frequency was increased if a clinical response was not achieved by week 6, and this was not appropriately randomised in GEMINI I.

Indirect comparison/MTC

4.18 The company presented a mixed treatment comparison that was carried out to estimate the relative treatment effect and safety of vedolizumab compared with the biologic therapies, infliximab, adalimumab and golimumab. The company updated its search protocol to include surgery and ciclosporin but stated that none of the studies identified were appropriate to include in a meta-analysis. It further stated that as ciclosporin was predominantly used for acute rather than chronic treatment it did not consider it an appropriate comparator. The identified studies which were used in meta-analyses included:

- ULTRA 1, ULTRA 2 and Suzuki et al which compared adalimumab with placebo;
- ACT 1 and ACT 2 which compared infliximab with placebo
- PURSUIT-SC/M which compared golimumab with placebo
- and GEMINI I which compared vedolizumab with placebo.

4.19 The company noted that there were differences between the studies. The duration of the studies varied between 6 and 8 weeks for the induction phase and between 52 and 54 weeks for the maintenance phase of treatment. Only GEMINI and ULTRA 2 included people who had received prior TNF-alpha inhibitors. GEMINI I, included people in whom treatment with TNF alpha inhibitors had failed, whereas ULTRA-2, included people who had shown a loss of response to, or an intolerance of another TNF alpha inhibitor, prior to starting adalimumab. The company commented that people in whom prior treatment with a TNF alpha inhibitor had failed may be less likely to have a successful response to subsequent treatment than people who had lost a response to, or had shown intolerance of a TNF alpha inhibitor. Another difference between the trials was how people were randomised after the induction phase. In GEMINI and PURSUIT-M, people were re-randomised if they responded to treatment during the induction phase, before entering the maintenance phase of trial. In all of the other trials people were randomised at baseline (prior to induction

treatment) only and continued to be followed during the maintenance phase in their assigned study arm regardless of whether they showed a response to treatment in the induction phases. See tables 27 and 28 pages 115 to 122 of company's submission for a summary of the studies.

4.20 The induction phase and maintenance phase data were synthesised separately by the company. The company presented data from a fixed effect model for a population who had not previously received a TNF alpha inhibitor, a population who had received a TNF alpha inhibitor which had failed and the whole population (using data from the ITT population in GEMINI for vedolizumab). No data were available for the efficacy of golimumab or infliximab in people who had received prior TNF alpha inhibitors and as such these comparators were not included in the meta-analysis in the population in which a TNF alpha inhibitor had failed. The company stated that its primary analyses were the subgroup analyses. grouped by prior TNF alpha inhibitor experience, because the patient populations differed between the studies and the proportion of people who had and had not have previous treatment with TNF alpha inhibitors may affect the results.

4.21 The company presented the odds ratios, estimated from the mixed treatment comparison, for vedolizumab compared with placebo, and two dosing regimens for adalimumab, golimumab and infliximab, for the population who had not had a TNF alpha inhibitor. There were no statistically significant differences in the odds of clinical response or remission during induction and maintenance treatment between vedolizumab and adalimumab used at its licensed dose (160 mg at week 0, 80 mg at weeks 2, 4 and 6 and 40 mg every other week thereafter) or golimumab used at its licensed dose (200 mg week 0, 100 mg week 2, 50 or 100 mg every 4 weeks thereafter) because the confidence intervals crossed 1. Compared with infliximab used at its licensed dose (5 mg/kg weeks 0,2,6 and every 8 weeks thereafter), there were no statistically significant differences in the odds of clinical response or remission during induction treatment but vedolizumab was associated with statistically

significantly better odds of achieving clinical remission than infliximab. In the population for whom TNF alpha inhibitors had failed there was no statistically significant difference in the odds of clinical remission with vedolizumab compared with adalimumab. The results are summarised in tables 9 and 10 for the population who had not, and who had received prior TNF alpha inhibitors respectively. The results of the mixed treatment comparison for the whole population are not presented here as these data were not subsequently used in the company's economic model (these are summarised in tables 37- 39 pages 133- 135 company's submission).

Table 9 Odds ratios for vedolizumab relative to each comparator in the subgroup who had not received TNF alpha inhibitors. Tables 31, 32, 33 company's submission pages 130-132

Outcome Measured	**Placebo**	**Ada 80/40[a]**	**Ada 160/80[a]**	**Gol 200/100[b]**	**Gol 400/200[b]**	**Inflix 5 mg/kg[c]**	**Inflix 10 mg/kg[c]**
Induction							
Clinical response vs. vedo	2.62 (1.69, 4.2)	1.88 (1.09, 3.40)	1.48 (0.90, 2.50)	1.04 (0.58, 1.80)	0.91 (0.51, 1.6)	0.64 (0.36, 1.2)	0.69 (0.39, 1.3)
Clinical remission vs. vedo	3.67 (1.67, 9.1)	3.00 (1.15, 8.30)	2.09 (0.88, 5.7)	1.05 (0.39, 3.1)	1.11 (0.4, 3.1)	0.72 (0.29, 1.9)	0.97 (0.39, 2.6)
Maintenance							
	Placebo	**Ada 40mg[d]**		**Gol 50mg[e]**	**Gol 100mg[e]**	**Inflix 5 mg/kg[f]**	**Inflix 10 mg/kg[f]**
Clinical remission vs. 4-wkly vedo	3.83 (1.94, 8.98)	2.28 (0.88, 6.18)		2.26 (0.96, 5.71)	2.16 (0.92, 5.38)	3.16 (1.05,8.88)	2.46 (0.86, 7.15)
Clinical remission vs. 8-wkly vedo	3.67 (1.85, 7.88)	2.14 (0.81, 5.82)		2.1 (0.9, 5.32)	2.01 (0.82, 4.97)	2.93 (1.03, 8.46)	2.3 (0.78, 6.69)

Ada: adalimumab; Inflix: Infiximab; Gol: golimumab; Vedo: vedolizumab
[a]dose at week0 /dose at weeks 2,4,6; [b]dose at week0 /dose at weeks 2; [c]dose at weeks 0, 2,6; [d]Every other week;[e] every 4 weeks;[f] every 8 weeks.
An odds ratio >1 favours vedolizumab; an odds ratio of <1 favours comparator

Table 10 odds ratio for vedolizumab relative to each comparator in the subgroup who have had a prior TNF alpha inhibitor. Tables 34- 36, page 133 company's submission

Induction		
	Placebo	**Adalimumab 160/80[a]**
Clinical response vs. vedolizumab	2.5 (1.2, 5.5)	1.7 (0.7, 4.4)
Clinical remission vs. vedolizumab	-3.7 (0.9, 28)	2.7 (0.4, 24)
Maintenance		
	Placebo	**Adalimumab 40 mg[b]**
Clinical remission vs. 4-weekly vedolizumab	11 (2.62, 76.0)	3.06 (0.34, 31.0)
Clinical remission vs. 8-weekly vedolizumab	12 (3.14, 78)	3.4 (0.4, 33)

[a]dose at week0 /dose at weeks 2,4,6; [b]Every other week

An odds ratio >1 favours vedolizumab; an odds ratio of <1 favours comparator

ERG comments

4.22 The ERG stated that the results from the network meta-analyses should be interpreted with caution as they were based on a fixed effects model rather than a random effects model (a fixed effect model assumes that the average result from each trial should be the same, a random effects model assumes that the average result from each trial may differ, but the average of the trial results would be the 'true result'). The ERG highlighted that there were considerable differences between the trials included in the network meta-analyses and that a random effects model would explicitly model these differences and capture the uncertainty in the true treatment effect, whereas a fixed effect model would underestimate the uncertainty.

4.23 The ERG considered the trials included in the network meta-analysis. It noted that the trials had different follow up times, and different study designs. The ERG agreed with the company that the difference in study duration during the maintenance phase would not have a large impact on the results. The ERG considered the difference in the study designs and noted that GEMINI I and PURSUIT-M only included people who had responded to induction treatment in the maintenance phase of the trials, and that they were re-randomised at the start of the maintenance phase. It noted that to enable comparison with adalimumab and infliximab the company had accounted for this by adjusting the results of the other trials

(ULTRA2, Suzuki 2014 and ACT1), to assume that responders at the end of the induction phase were the same as responders at the end of the maintenance phase. The ERG stated that the people who had not responded to treatment at the end of the induction phase may respond during the maintenance phase and therefore using the responders at the end of the maintenance phase may provide an overestimate of responders. The impact of this is likely to be different on different treatment arms and therefore the impact on relative treatment effect is unclear. The ERG stated therefore that it was not clear whether the results in GEMINI or PURSUIT-M over- or under-estimated the relative treatment effect of vedolizumab relative to the comparators in the maintenance phase.

4.24 The ERG noted that the company had presented separate network meta-analyses for people who had or had not received prior treatment with TNF-alpha inhibitors, without providing a full rationale for this approach. The ERG stated that the disadvantage of doing separate analyses by subgroup is that the possibility of an interaction between treatment and subgroup cannot be explored, and that this should be explored using meta-regression. The company stated in response to clarification questions that performing a meta-regression was not appropriate because there were an insufficient number of trials included in the networks. The ERG stated that in the absence of a meta-regression analysis the company should present the predictive distribution of mean treatment effect which incorporates extra uncertainty due to potential differences between studies.

4.25 The ERG noted that the company had modelled response and remission separately and stated that the results for clinical remission and response should be interpreted with caution because these results were estimated without considering dependence or correlation between these 2 outcomes.

4.26 The ERG noted that there were numerous errors in the company's reporting of the numbers of patients experiencing response and remission across the treatment arms in ULTRA 2, PURSUIT-M and Suzuki et al. It stated that it had not checked all other outcomes presented by the company and it was unclear if the data used for the other outcome measures were correct.

Adverse effects of treatment

4.27 The company presented adverse event data from:

- GEMINI I,
- 2 further placebo controlled clinical trials of vedolizumab in people with Crohn's disease (GEMINI II and III), and
- interim safety data from a single arm extension study evaluating the long-term safety of vedolizumab, in people with ulcerative colitis or Crohn's disease, beyond 12 months of treatment.

The safety population was defined as people who had received at least one dose of the study drug. The adverse events occurring in more than 5% of patients in the GEMINI I safety population, and rates of serious adverse events and serious infections are described in table 11. The interim results of the extension study showed that drug-related adverse events were similar between people with ulcerative colitis and people with Crohn's disease, with the most common being headache (6%), nasopharyngitis (4%), nausea (4%), arthralgia (4%), upper respiratory infection (3%), and fatigue (3%) and that adverse events were similar to those observed in GEMINI I and II, in which people received treatment for up to12 months. The company pooled safety data from all of the vedolizumab trials (including 2 further phase II trials) in both people with ulcerative colitis and Crohn's disease. The most common serious adverse events were exacerbation of ulcerative colitis (or Crohn's disease), abdominal pain, and anal abscess (estimated less than 2 patients experiencing an event per 1000 patient years). No cases of PML (progressive multifocal leukoencephalopathy) were reported (June 2013).

Table 11 Adverse events occurring in ≥5% of patients receiving vedolizumab in GEMINI (table 44 company submission page 142)

Event	Placebo N=275 n (%)	Vedolizumab N=620 n (%)
Headache	28 (10.2)	80 (12.9)
Ulcerative colitis	58 (21.1)	97 (15.6)
Nasopharyngitis	26 (9.5)	80 (12.9)
Upper respiratory tract infection	21 (7.6)	52 (8.4)
Arthralgia	25 (9.1)	56 (9.0)
Nausea	19 (6.9)	38 (6.1)
Abdominal pain	10 (3.6)	35 (5.6)
Anaemia	16 (5.8)	35 (5.6)
Fatigue	10 (3.6)	33 (5.3)
Cough	13 (4.7)	36 (5.8)
Any serious adverse event	37 (13.5)	77 (12.4)
Any serious infection†	8 (2.9)	12 (1.9)
Any cancer	3 (1.1)‡	1 (0.2)§

† A serious infection was defined as a serious adverse event of infection according to the MedDRA, version 15, criteria.

‡ colon cancer, transitional-cell carcinoma and squamous-cell carcinoma of the skin occurred in 1 patient each in the placebo group

§ Colon cancer occurred in 1 patient in the VEDO group.

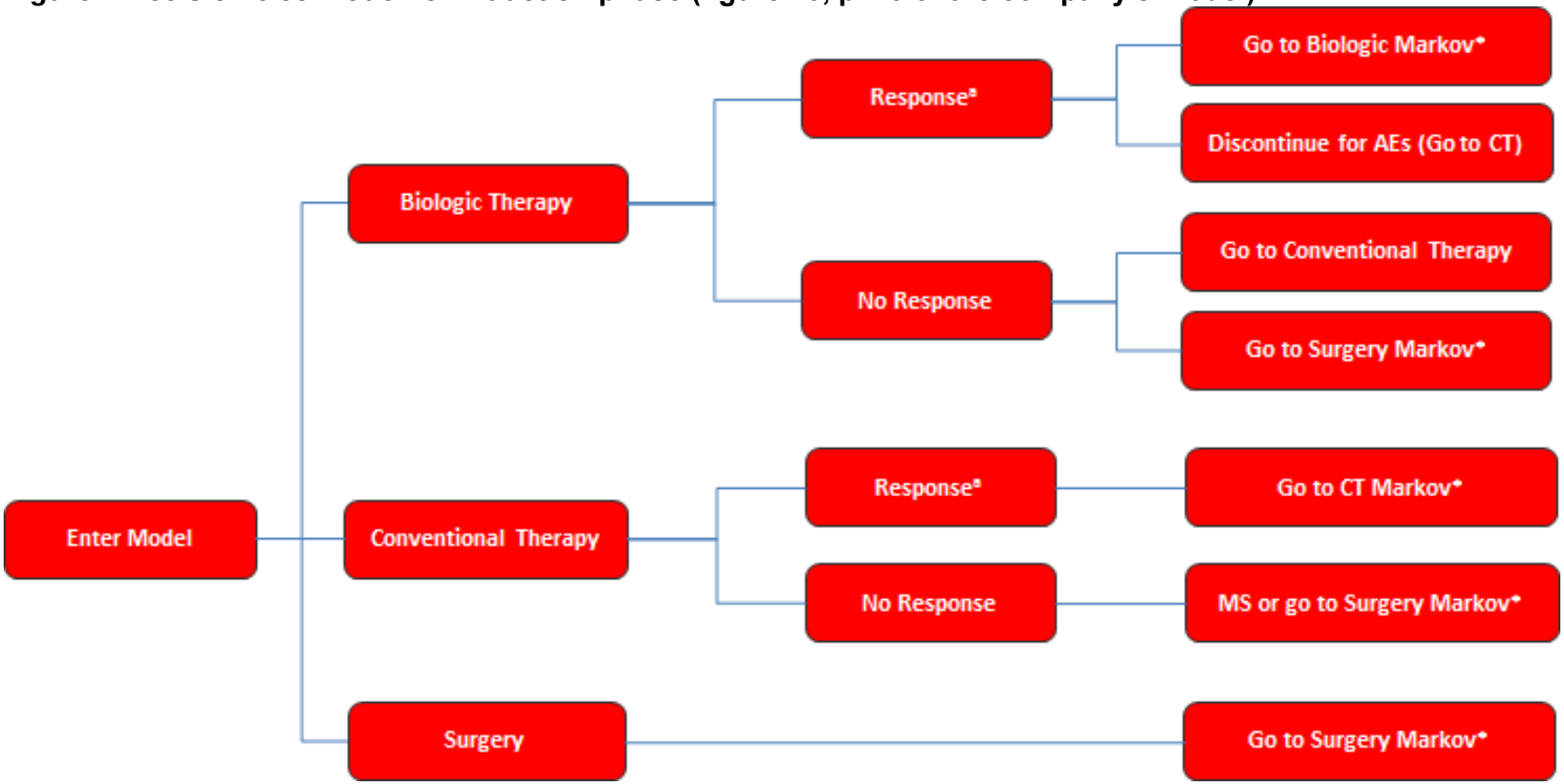
ERG comments

4.28 The ERG noted that there was an absence of long-term safety data, and that only limited data on deaths were presented in the company submission. However, the ERG noted that the EPAR concluded that no deaths following studies of vedolizumab could be ascribed with any reasonable degree of certainty to vedolizumab. The ERG stated that as of June 2013, across treatment arms, 1 person had died in GEMINI I, 5 people had died in a placebo controlled trial of vedolizumab for Crohn's disease (GEMINI II), and 7 people in the extension study of the GEMINI trials. The ERG heard from the company, in response to clarification questions, that none of these deaths were considered to be related to vedolizumab by the study investigators.

5 Cost-effectiveness evidence

5.1 The company identified 2 relevant cost-effectiveness studies; however neither study presented results for a life time horizon. The company developed a new model of the induction and the maintenance phases of treatment with vedolizumab and its comparators. A decision tree structure was used to model the induction phase of treatment (Figure 2). The induction phase was assumed to be 6 weeks. The criterion for response was a drop in Mayo score of 3 or more. People who had a response remained on their assigned treatment in the maintenance phase. People who did not have a response, or who discontinued a biologic treatment (vedolizumab, adalimumab, infliximab or golimumab) because of an adverse event were assumed to have conventional therapy in the maintenance phase. The maintenance phase of the model had a Markov structure (Figure 3) similar to those developed for technology appraisal 140 (2008) Infliximab for subacute manifestations of ulcerative colitis, and a published cost utility analysis of infliximab compared with conventional therapy (Tsai et al.) People entered the maintenance phase in one of 3 disease severity health states (defined according to Mayo scores: ‘remission’ [Mayo score of 0 to 2]; ‘mild’ [Mayo score of 3 to 5]; and ‘moderate to severe’ [Mayo score of 6 to 12]) or the ‘surgery’ health state; depending on response at the end of the induction phase. In addition, the model contained health states for, ‘postsurgical remission’, ‘postsurgical complications’, ‘people who had discontinued treatment’ and ‘death’. The model considered the costs and health benefits from the perspective of the NHS and were discounted by 3.5% per year over a time horizon of 10 years. The cycle length for the maintenance phase was 8 weeks, which the company stated was likely to be sufficient time for the Mayo scores to be relatively stable.

Figure 2 Decision-tree model for induction phase (figure 10, p173 of the company's model)

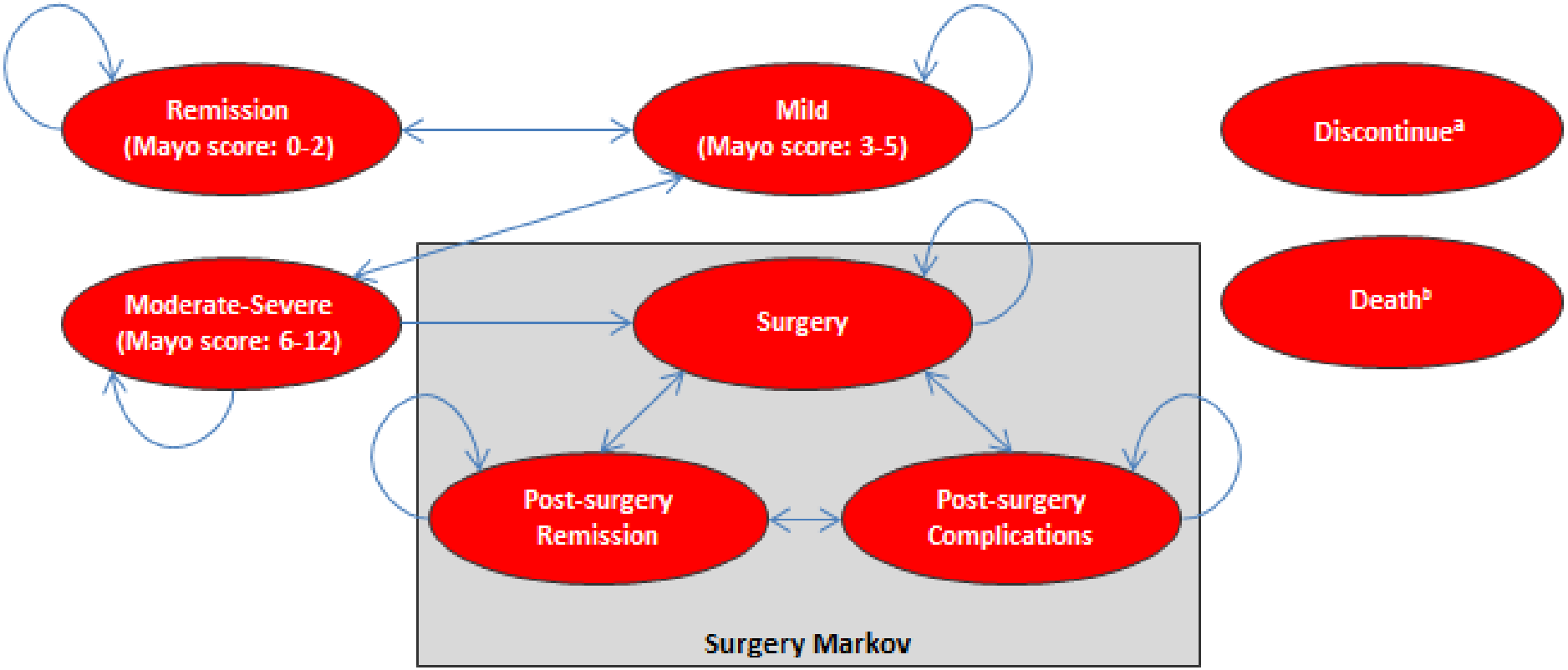


AE, adverse event; CT, conventional therapy; MS, moderate to severe.

[a] Response is defined as a drop in Mayo score of 3 points or more. This includes patients who also achieve remission, as remission is a subset of response. Remission is defined as a Mayo score less than 3.

* The 'biologic Markov' and 'conventional Markov' structures for biologic therapies and conventional therapies are the same but the probability of moving between health states differs. The surgery Markov is a subset of the larger Markov.

Figure 3 Markov model for maintenance phase (figure 11, p176 of the company's model)



[a] Reasons for discontinuation include lack of response and adverse events. Discontinuation due to adverse events is applicable only to responders on biologic treatments, because non responders on biologics switch to conventional therapy and continue receiving such until the end of the model's time horizon or until the patients require surgery.

[b] Patients may transition to death from any health state during any cycle.

ERG comments

5.2 The ERG noted that a 10–year time horizon was used for the company's base case, but it was not clear whether all relevant health gains and costs would be captured within that time. The ERG stated that running the model over a lifetime horizon was preferable, but noted that the clinical trial data only assessed outcomes up to 54 weeks and extrapolating data to a lifetime horizon would be subject to considerable uncertainty.

Model details

Population and comparators

5.3 The company's analysis was presented for 3 populations:

- The whole (mixed) population, including people who had prior anti-TNF inhibitor therapy and those who had not (TNF alpha inhibitor naïve);
- people who are TNF alpha inhibitor naïve only, and;
- people in whom prior TNF alpha inhibitor treatment had failed.

For all 3 analyses, the comparators included conventional therapies (a combination of aminosalicylates, immunomodulators and corticosteroids) and surgery. TNF alpha inhibitors (infliximab, adalimumab and golimumab) were only included as comparators for the subgroup of people who were TNF alpha inhibitor naïve.

5.4 In the model it was assumed that response to induction treatment would be assessed at 6 weeks based on when it was assessed in GEMINI I. The company noted that the trials for infliximab and adalimumab measured response at week 8, but for the purposes of the modelling it was assumed that response at week 6 would be equivalent to those observed at week 8. The number of doses people received during the induction phase was also assumed to be the same as the clinical trials on which the efficacy estimates were made. This meant that people receiving vedolizumab or golimumab had 2 doses (at week 0 and week 2), people receiving adalimumab had 4 doses (at weeks 0, 2, 4 and 6) and people receiving

infliximab had 3 doses (at weeks 0, 2 and 6). The company tested a scenario in which response was tested at week 10 after 3 doses of vedolizumab, which the company stated may reflect clinical practice where the decision to continue with treatment is made later (see section 5.29). It was assumed that people receiving vedolizumab or TNF alpha inhibitors were concomitantly treated with conventional therapy, but at a lower dosage than if conventional therapy was their only treatment. See table 54, p181 of the company's submission for more information on the treatment regimens for the intervention and comparators.

5.5 Table 12 summarises the sources of efficacy data used in the model according to the population and treatment. For the whole population and population in whom prior treatment withTNF alpha inhibitors had failed, efficacy data (response and remission) from GEMINI I were used. For the population who had not received prior treatment with a TNF alpha inhibitor, the company used efficacy estimates (response and remission) from its mixed treatment comparison. The estimates of the probability of having surgery, having remission after surgery, and having complications following surgery, were obtained from published literature.

Table 12 Populations, comparators and sources of efficacy evidence used in company's health economic analysis

Population	Interventions compared	Source of efficacy data
(1) Mixed (ITT population)	• Vedolizumab • Conventional non-biologic therapies (5-ASAs, immunomodulators, corticosteroids) • Surgery	**Induction:** Observed outcomes from GEMINI1 used to inform probabilities of response/remission **Maintenance:** Model fitted to probability of achieving response/remission observed in the GEMINI1 trial
(2) Anti-TNF-α naïve population	• Vedolizumab • Infliximab • Adalimumab • Golimumab • Conventional non-biologic therapies (5-ASAs, immunomodulators, corticosteroids) • Surgery	**Induction:** company's network meta-analysis (NMA) used to inform probabilities of response/remission for each option **Maintenance:** Model maintenance transition matrix fitted against 1-year probabilities of response/remission predicted by company's NMA
(3) Anti-TNF-α failure population	• Vedolizumab • Conventional non-biologic therapies (5-ASAs, immunomodulators, corticosteroids) • Surgery	**Induction:** Observed outcomes from GEMINI1 used to inform probabilities of response/remission **Maintenance:** Model fitted to probability of response/remission observed in the GEMINI1 trial

Model transitions and discontinuations

5.6 To derive the probability of moving between, or remaining in the remission, mild, or moderate to severe health states during the maintenance phase, the company used a calibration approach. The company used data from GEMINI I on the proportion of people in remission, or with moderate to severe ulcerative colitis at the end of the induction treatment (6 weeks), and the proportion of people responding, or in remission, at the end of the maintenance period (52 weeks) to determine estimates of the probability of moving between the health states during the first year of maintenance. These transition probabilities were then assumed to remain constant over time and therefore were

applied to each subsequent year in the model. To calculate the estimates and calibrate the model the company applied the following constraints

- No more than 99.5% of people would remain in remission in each week cycle
- No more than 20% of people with mild disease would enter remission
- More people would remain in the mild health state than enter the moderate to severe health state; more people would remain in the moderate to severe health state than move to the mild health state
- People would not move directly from remission to the moderate to severe health state and vice versa
- The sum of the transition probabilities would equal 1

5.7 In the model, people could progress to receive surgery. People could move into the surgery health state if they didn't respond to induction treatment, or if they had moderate to severe ulcerative colitis during the maintenance phase. Once in the surgery and post-surgery health states, treatment was discontinued for the remainder of the person's lifetime. It was assumed that 40% of people having surgery would have a proctocolectomy with ileostomy and 60% would have subtotal proctocolectomy with pouch formation with or without loop ileostomy. Following surgery, some people subsequently experienced postsurgical complications, required additional surgeries, or remained in postsurgical remission. The company derived the transition probabilities from surgery and the post surgery health states from a review of published literature.

Table 13 transition probabilities used in the model for people who had surgery (table 65 page 193 company's submission)

	Surgery	**Post surgery Remission**	**Post surgery Complications**
Surgery	0.050	0.450	0.500
Post surgery remission	0.050	0.777	0.173
Post surgery complications	0.050	0.245	0.705

5.8 In the model it was assumed some people would discontinue treatment with vedolizumab, adalimumab, infliximab or golimumab. Treatment was discontinued if people had not had a response by the end of the induction phase or if there were adverse events at any time. The data for discontinuations and for adverse event rates were obtained from the relevant clinical trials for each treatment (see tables 60 and 61 of the company's submission, p189 and 190, respectively). For people who continued on treatment, the treatment duration with biologic therapy (vedolizumab, infliximab, adalimumab or golimumab) was assumed to be a maximum of 1 year, after which people switched to conventional therapy. People who had conventional therapy were assumed to only discontinue treatment if they needed surgery. The Markov model did not include the option of discontinuing treatment temporarily, because of a lack of data on treatment breaks for all comparators. The company stated that the clinical trials results would capture the effect of any temporary discontinuation.

Table 14 probabilities of adverse events by treatment (table 61 page 190 company submission)

Adverse Event	Vedolizumab[a]	Infliximab	Adalimumab[c]	Golimumab	Conventional Therapy
Serious infection	1.15%	2.48%	0.83%	1.37%	1.98%
Tuberculosis	0.00%	0.00%	0.00%	0.00%	0.08%
Lymphoma	0.00%	0.00%	0.00%	0.00%	0.00%
Acute hypersensitivity reactions	0.00%	1.65%	0.00%	0.30%	0.29%
Skin Reactions	0.29%	11.57%	0.00%	5.03%	3.20%

5.9 During the maintenance phase of the model people could die whilst in any health state at any time. The probability of dying was estimated using age- and sex-specific all-cause mortality from the UK (Office for National Statistics, 2011). This was adjusted for disease severity, surgery, post surgery remission, and complications to incorporate an increased risk of mortality associated with moderate to severe disease, and surgery (see

table 67 page 195 company's submission for relative mortality risk by health state).

ERG comments

5.10 The ERG commented on the company's use of a calibration approach to estimate transition probabilities in the maintenance phase. It noted that patient level data for people with remission, mild, or moderate to severe disease during the maintenance phase would be available for GEMINI I. However, it may not be available to the company for the adalimumab, golimumab, and infliximab trials included in the network meta-analysis. It commented that the assumptions and constraints used in the calibration calculations, including using a different starting matrix for biologic therapies and conventional therapies, were arbitrary. It commented that using a calibration process to fit 7 unknown parameters to 2 known data points meant that over fitting may have occurred. The ERG commented that there would be many possible combinations of transition probabilities that could fit the 1 year data points for response and remission. It also noted that the calibration process did not account for people who achieved response but whose disease symptoms remained moderate to severe.

5.11 The ERG commented on the plausibility of the assumptions relating to the transition probabilities between the surgery and post surgery health states. It noted that the company had converted 6-month estimates for repeat surgery, and complications following surgery, to an 8-weekly probability (assuming a constant rate), and then had applied these probabilities for the full time horizon. However, the ERG stated that the probability of repeat surgery and complications would be expected to be greater in the first 12 months after surgery, rather than remaining constant indefinitely. It also noted that the company's estimate of entering remission after having a post-surgery complication was based on an estimate for 1 type of complication only (pouch leaks) and it was unclear how the probability related to annual risk. Overall the ERG considered that

the company's assumptions would overestimate the probability of undergoing surgical procedures, and the time spent in the post-surgical complications state, which would result in an inflation of the costs and a reduction of the health gains associated with surgery.

5.12 The ERG commented that the marketing authorisations for vedolizumab, infliximab, golimumab or adalimumab do not stipulate if or when people responding to therapy should stop treatment. It noted that the company assumed that people who were responding to these biologic treatments would take them for 1 year and then switch to conventional therapy. The ERG stated that it was unclear whether in clinical practice biologic therapy would be stopped when a patient is gaining clinical benefit from it. The ERG commented that it was also assumed in the model that people would continue to receive biologic maintenance therapy up to 1 year, even if they lost response after the induction period and stated that this 'continuation rule' was unlikely to be clinically realistic.

Utility values

5.13 To estimate utility values for the health states in the model the company conducted a post-hoc analysis of EQ-5D data from the maintenance phase of GEMINI I. It used the combined data from people who had received vedolizumab or placebo, and from all time points at which data was collected. The scores were grouped according to whether they were in remission (Mayo score 0-2), had mild disease (Mayo score 3-5) or had moderate to severe disease (Mayo score 6-12). Surgery outcomes were not assessed in GEMINI I, and utility values associated with the surgery and post surgery health states were taken from Punekar 2010. This study reported EQ-5D data collected from UK patients with UK tariffs applied to the EQ-5D scores. The company conducted sensitivity analyses that used utility values from Punekar 2010 (for all health states), Tsai et al, or Arseneau et al. The Arsenau et al study used time trade off and visual analogue methods to derive utility values (see section 5.29).

Table 15 Mean EQ-5D scores in the maintenance phase of GEMINI I by mayo score used in the company's base case model (table 69, p228 and table 73, p240 of the company's submission)

	Remission (SD)	Mild (SD)	Moderate to Severe (SD)	Surgery	Post surgery remission	Post surgery comp.
Company base case	0.86 (0.188)	0.80 (0.184)	0.68 (0.224)	0.42	0.60	0.42
Utility values used in company's sensitivity analyses						
Punekar and Hawkins	0.88	Assumed 0.76	0.42	0.42	0.60	0.42
Arseneau et al.	0.79	Assumed 0.76	0.32	0.32	0.63	0.49
Tsai et al.	0.88	0.76	0.42	0.61	0.61	0.55

SD, standard deviation; comp: complications

5.14 The company also assigned utility decrements for selected adverse events. The rates of adverse events experienced were obtained from the clinical trials. The utility decrements were sourced from published studies. The disutility values associated with treating TB and lymphoma were not used in the company's model because the incidence rate of these events was assumed to be zero.

Table 16 utility decrements associated with adverse events in the company model (table 34 p87 ERG report)

Adverse event	Disutility value	Source	Elicitation methods
Serious infection	-0.520	Brown *et al*	Proxy utility values derived from 180 nurses using standard gamble (SG) methods. Reported utility of 0.48 converted to disutility assuming baseline of perfect health.
TB	-0.550	Porco *et al*	Elicitation method unclear. Reported utility of 0.45 converted to disutility assuming baseline of perfect health.
Malignancy (including lymphoma)	-0.195	Hornberger *et al*	Elicitation method unclear. Reported utility of 0.805 for follicular lymphoma (pre-progression) converted to disutility assuming baseline of perfect health.
Acute hypersensitivity reactions	-0.110	Beusterien *et al*	Cross-sectional SG using members of the general public. Disutility directly estimated as part of analysis.
Skin site reactions	-0.030	Beusterien *et al*[59]	Cross-sectional SG using members of the general public. Disutility directly estimated as part of analysis.

ERG comments

5.15 The ERG considered it was appropriate to use EQ-5D data from the GEMINI I trial to determine the utility associated with the disease severity health states in the model. However, the ERG noted that this approach did not differentiate between people who did or did not have a response to treatment with moderate to severe ulcerative colitis.

5.16 The ERG noted that in the company's model it was assumed that the utility value for people in post-surgical remission was lower (0.60) than the utility value for people with moderate to severe ulcerative colitis (0.68), reflecting worse quality of life. The ERG considered that this was not plausible because it doesn't represent any benefit from surgery. Therefore the ERG questioned why anyone would go through surgery if their expectation was a lower quality of life afterwards.

5.17 The ERG was unable to verify that utility values for surgery, post surgery remission, and post surgery complications from Punekar and Hawkins were for people with ulcerative colitis. The ERG commented that the Punekar and Hawkins paper cited the source of utility values as a study of the epidemiology and costs of Crohn's disease. The ERG identified a different health utility study of people with ulcerative colitis reporting utility values for remission, response, moderate to severe ulcerative colitis and post-surgery (Woehl et al). It noted that the values for people who had undergone surgery presented in Woehl et al were much higher than those reported in Punekar and Hawkins. In addition, the values for the pre-surgery states were slightly different. The ERG considered that the company's assumptions surrounding surgery and post-surgery health state utility values would underestimate the health gains for people undergoing surgery and favoured drug therapies over surgery.

5.18 The ERG commented on the utility decrements applied in the company's model, for people who experienced adverse events. The ERG noted that the disutility values were calculated by deriving the utility for someone experiencing an adverse event, and subtracting this value from 1. The

ERG commented that this would over estimate disutility as it assumes the person would otherwise have a utility of 1, that is, have perfect health. The ERG commented that disutility may also have been captured to some extent because the utility values for each disease severity health state were derived from the GEMINI I trial (therefore any adverse events experienced during the trial should be included). However, the ERG noted that the EQ-5D scores from GEMINI I had not been assessed separately for people who received treatment with vedolizumab or placebo.

5.19 The ERG commented on the probability of experiencing an adverse event in the model. The ERG noted that the estimates of adverse events with conventional therapy were derived from a pooled analysis of the placebo arms of trials of vedolizumab, adalimumab, infliximab and golimumab. It noted that in these trials people in the placebo arm had received a placebo transfusion or injection, which would not normally be given as part of conventional therapy. The ERG stated that it was not clear whether skin reactions experienced on conventional therapy may be infusion site rashes as a consequence of placebo delivery rather than as a reaction to the conventional therapy itself.

Costs and resource-use

5.20 The company has agreed a patient access scheme with the Department of Health for Vedolizumab. The discounted price of [redacted] per vial dose was used in the model (The NHS list price without the discount is £2050 per vial). The model used the NHS list price for adalimumab; golimumab and infliximab (see table 17). The company stated that only a 50 mg dose of golimumab is available in the UK (please note, however, that the BNF lists a 100 mg pre-filled pen, and Merck Sharpe and Dohme, the manufacturer of golimumb, have agreed a patient access scheme where the 100 mg dose is available to the NHS for the same price as the 50 mg dose (£762.97)). The company estimated a weighted average cost of conventional therapy including a combination of aminosalicates, corticosteroids and immunosuppressants (azathioprine, mercaptopurine

and methotrexate). The proportion of each drug used was based on clinical expert opinion. The cost of conventional therapies were based on costs and dosing regimens in the BNF (2013) and was £204.80 for 8 weeks of treatment. The company assumed that the costs of conventional therapy would be halved if taken with vedolizumab, adalimumab, infliximab or golimumab rather than if conventional therapies were the only treatment a person received.

Table 17 Drug costs used in the model. (combined data from tables 77, 78 and 79 on pages 254 and 255 of company's submission)

	Dosing schedule induction	**Dosing schedule maintenance**	**Cost per vial**	**Cost per cycle**
Vedolizumab (intravenous)	1 x 300 mg at weeks 0+2	1 x 300 mg every 8 weeks	[redacted]	Induction: [redacted] Maintenance [redacted]
Infliximab (intravenous)	5mg/kg (4x100 mg) at weeks 0, 2, 6	4 x 100 mg once every 8 weeks	£419.62	Induction: £5,035.44 Maintenance: £1,678.48
Adalimumab subcutaneous	4 x 40 mg at week 0 2 x 40 mg week 2 1 x 40 mg at weeks 4+6	1 x 40 mg 4 times every 8 weeks	£352.14	Induction: £2,817.12 Maintenance: £1,408.56
Golimumab subcutaneous	4 x 50 mg at week 0 2 x 50 mg at week 2	1 x50 mg twice every 8 weeks	£762.97	Induction: £4,577.82 Maintenance: £1,525.94
Conventional therapies (% people using)	<u>Aminosalicyclates:</u> Balsalazide (13%) Mesalazine (13%) Olsalazine (13%) Sulfasalazine (13%) <u>Corticosteroids:</u> Budesonide (1%) Prednisolone (36%; topical) Immunomodulators Azathioprine (39%) Mercaptopurine (15%) Methotrexate (9%)		n/a	Maintenance (8 week cost) £204.80

5.21 Resource costs in the model included the costs of consultant visits, blood tests, and elective and emergency endoscopy, which were based on NHS

reference costs 2012/13. The cost of surgery was assumed to be £13, 577.27 (Buchanan et al 2011). The frequency of resource use in each health state was based on Tsai et al. An additional cost of £308 for intravenous infusion was applied to vedolizumab and infliximab at each administration visit (payment by results tariff 2012/13). Please see table 80 page 256 of the company's submission for a breakdown of health care resource use by health state. Costs of adverse events are summarised in table 18.

Table 18 costs of adverse events in company's model (table 81 page 258 company submission)

Adverse Event	Total Cost	Source
Serious infection	£1,470.00	NHS Reference Costs 2011/12. Average of 5 different types of serious infections: sepsis, pneumonia, urinary tract infection, respiratory infection, and bronchitis
Tuberculosis	£2,272.00	NHS Reference Costs 2011/12. Average of non-elective short-stay and long-stay tuberculosis
Lymphoma	£14,975.00	NICE (2003), NICE (2012), and NICE (2011). Average of lymphoma costs from three technological appraisals for rituximab (TA65, TA243, and TA226)
Hypersensitivity	£3,188.00	NHS Reference Costs 2011/12. Average of non-elective short-stay and long-stay pyrexia
Injection site reactions	£1,363.28	NHS Reference Costs 2011/12. Average of procedures associated with skin disorders

NHS, National Health Service; NICE, National Institute for Health and Care Excellence.

ERG comments

5.22 The ERG noted that the costs in the model for endoscopy, consultant visits, blood tests, and hospitalisations were based on 2006/7 NHS reference costs (cited in Tsai et al, and uplifted to current prices) rather than 2012/13 NHS Reference Costs, as stated by the company. The ERG commented that the actual 2012/13 NHS reference costs were much lower (with the exception of consultant visit costs). The ERG stated that this resulted in the model overestimating costs in the post-surgical complication health state (£12,470 per year rather than £9,109 per year) and the post-surgical remission health state (£3,048 per year rather than £1,447 per year).

5.23 The ERG commented on the costs included in the post-surgery health states. It stated that it was not clear whether costs associated with care of stoma (a surgical opening of the digestive tract in the abdomen to bypass the rectum), which would include nurse visits and consumables, were included. The ERG estimated that the costs of stoma care would be approximately £466 per year based on Buchanan et al.

5.24 The ERG noted that in the company's model, the costs of treating adverse events included the costs of either a short or long inpatient stay. The ERG considered that whilst some of the adverse events would require inpatient admission, it is unlikely that this is true for all adverse events, for example skin reactions. The ERG considered that the costs assumed for treating adverse events may overestimate the true cost.

5.25 The ERG noted that in the company's model, the costs associated with conventional non-biologic therapies were half of those incurred by people who are receiving conventional therapies only, and that this assumption was not justified.

5.26 The ERG stated that the company's model included the cost of topical rather than oral prednisolone. It noted that replacing the cost of topical prednisolone with that for oral prednisolone reduces the overall cost of conventional therapy but noted that this did not have a large impact on the ICER for vedolizumab.

Company's base case results and sensitivity analysis

5.27 The company presented deterministic base case results for the three populations it had modelled (see section 5.3). The company presented pairwise comparisons showing the incremental cost effectiveness ratio (ICER) for vedolizumab compared with each comparator separately, and did not present a fully incremental analysis.

- **For the whole population** vedolizumab dominated surgery (it was less costly and more effective). The ICER for vedolizumab compared with conventional therapy was £33,297 per QALY gained.

- **In the population who had not received prior TNF alpha inhibitors** vedolizumab dominated infliximab, golimumab and surgery. Vedolizumab was associated with an ICER of £6,634 per QALY gained when compared with adalimumab, and £4,862 per QALY gained when compared with conventional therapy.
- **In the population in whom treatment with prior TNF alpha inhibitors had failed** vedolizumab dominated surgery and was associated with an ICER of £64,999 per QALY gained when compared with conventional therapy.

Table 19: Company's base case. Deterministic ICERs, pairwise comparisons with vedolizumab (adapted from table 106 page 286 company submission)

	QALYs	Costs	Incremental QALYs	Incremental costs	Pairwise ICER (vedolizumab versus comparator)
Whole population					
Vedolizumab	5.55	£77,056	-	-	**-**
Conventional therapy	5.40	£71,925	0.15	£5,131	£33,297
Surgery	4.28	£107,831	1.27	-£30,775	Vedolizumab dominates
Population who have not received prior TNF alpha inhibitors					
Vedolizumab	5.90	£69,075	-	-	**-**
Infliximab	5.82	£73,952	0.08	-£4,877	Vedolizumab dominates
Golimumab	5.79	£70,387	0.11	-£1,312	Vedolizumab dominates
Adalimumab	5.76	£68,157	0.14	£918	£6,634
Conventional therapy	5.56	£67,406	0.34	£1,669	£4,862
Surgery	4.28	£107,831	1.67	-£38,756	Vedolizumab dominates
Population in whom prior treatment with TNF alpha inhibitors has failed					
Vedolizumab	5.46	£78,409	-	-	**-**
Conventional therapy	5.37	£72,570	0.09	£5,839	£64,999
Surgery	4.28	£107,831	1.182	-£29,422	Vedolizumab dominates

5.28 One way sensitivity analyses were conducted for each of the 3 populations. When comparing vedolizumab with surgery in all three populations, the variables with the largest impact on the ICER were probabilities of transitioning between surgery and post-surgical health states, health state costs and health state utilities (see figures 34, 39 and 41 of the company's submission). In all of the sensitivity analyses vedolizumab remained dominant to surgery. When vedolizumab was compared with conventional therapy, the parameters that had a larger impact on the ICER, for all 3 populations included:

- remission transition probabilities for conventional therapy and vedolizumab;
- the efficacy of vedolizumab and conventional therapy in the initial response period;
- and health state costs

The tornado diagrams for these sensitivity analyses are on pages 288 to 294 of the company's submission.

ERG comments

5.29 The ERG commented on the company's use of deterministic rather than probabilistic estimates of the ICERs for vedolizumab compared with its comparators. It noted that the company had performed a probabilistic sensitivity analysis but that this was not straightforward in the company's model due to limitations of the model structure which only allowed comparison of two comparators, rather than all comparators, at a time. The ERG stated that it was inappropriate to use deterministic rather than probabilistic estimates of the ICER, as this does not take into account non-linearity of the model.

Company's scenarios

5.30 The company presented 5 scenario analyses (see table 41, section 7.6.9 of the company's submission) which included altering the model time horizon (lifetime and 1 year, rather than 10 years), using alternative

sources of utility values (see section 5.13), excluding the excess mortality risk for ulcerative colitis, using 10-week response data rather than 6-week response data and extending the maximum duration of biologic treatment from 1 year to 3 years. Pairwise comparisons of vedolizumab with the comparators were reported for the the whole intention to treat population (table 20), the population in whom treatment with a TNF alpha inhibitor had failed (table 21) and the population who had not had prior treatment with a TNF alpha inhibitor (table 22).

Table 20 Scenario analyses presented by the company for the whole intention to treat population

Scenario	ICER vs. conventional therapy	ICER vs. surgery
Base case	£33,297	Vedolizumab dominates
1-year time horizon rather than 10-years	£188,640	Vedolizumab dominates
Lifetime horizon (63 years)	£20,599	Vedolizumab dominates
Utility values from Punekar *et al*	£17,857	£117,134*
Utility values from Arseneau *et al*	£18,008	£26,438*
Utility values from Tsai *et al*	£18,627	£46,733*
Excluding the assumption of an increased mortality risk associated with ulcerative colitis	£33,675	Vedolizumab dominates
Response assessment following induction at 10 weeks rather than 6 weeks	£31,414	Vedolizumab dominates
Maximum duration treatment with vedolizumab =3 years rather than 1 year	£39,575	Vedolizumab dominates

* Results are in the south-west quadrant of the cost-effectiveness plane: the ICER represents the money saved with vedolizumab for QALYs lost, compared with surgery

Table 21 Scenario analyses presented by the company for the population in whom prior treatment with TNF alpha inhibitors has failed

Scenario	ICER vs. conventional therapy	ICER vs. surgery
Base case	£64,999	Vedolizumab dominates
1-year time horizon rather than 10-years	£230,671	Vedolizumab dominates
Lifetime horizon (63 years)	£44,132	Vedolizumab dominates
Utility values from Punekar *et al*	£35,830	£67,866*
Utility values from Arseneau *et al*	£35,355	£22,164*
Utility values from Tsai *et al*	£37,589	£35,732*
Excluding the assumption of an increased mortality risk associated with ulcerative colitis	£66,025	Vedolizumab dominates
Response assessment following induction at 10 weeks rather than 6 weeks	£55,763	Vedolizumab dominates
Maximum time on treatment =3 years rather than 1 year	£55,149	Vedolizumab dominates

* Results are in the south-west quadrant of the cost-effectiveness plane: the ICER represents the money saved with vedolizumab for QALYs lost, compared with surgery

Table 22 Scenario analyses presented by the company for the population who have not had prior treatment with anti TNF alpha inhibitors

Scenario	ICER for vedolizumab versus comparator				
	Conventional therapy	**Infliximab**	**Adalimumab**	**Golimumab**	**Surgery**
Anti-TNF-α naïve population					
Base case	£4,862	Vedolizumab dominates	£6,634	Vedolizumab dominates	Vedolizumab dominates
1-year time horizon	£139,885	Vedolizumab dominates	£135,406	£51,918	Not reported
Lifetime horizon	Vedolizumab dominates	Vedolizumab dominates	Vedolizumab dominates	Vedolizumab dominates	Not reported
Utility values from Punekar et al	£2,469	Vedolizumab dominates	£3,342	Vedolizumab dominates	Not reported
Utility values from Arseneau et al	£2,375	Vedolizumab dominates	£3,190	Vedolizumab dominates	Not reported
Utility values from Tsai et al	£2,375	Vedolizumab dominates	£3,459	Vedolizumab dominates	Not reported
Excluding the assumption of an increased mortality risk associated with ulcerative colitis	£4,647	Vedolizumab dominates	£6,452	Vedolizumab dominates	Not reported
Response assessment following induction at 10 weeks rather than 6 weeks	£12,726	Vedolizumab dominates	£21,006	£6,916	Not reported
Maximum time on (biologic) treatment 3 years	£26,152	Vedolizumab dominates	£50,607	£15,548	Not reported

5.31 In the company's scenario analyses, the model was found to be sensitive to the time horizon, with longer time horizons reducing the ICER in all analyses. The company noted that in the base case, all people treated with a biologic were assumed in the model to switch to conventional therapy after 1 year. Therefore, the long-term effectiveness of vedolizumab was determined by the effect of vedolizumab treatment over 1 year on the distribution of people across the health states at the end of that year. That is, the long-term benefit of vedolizumab was derived from there being a higher proportion of people in better health states at the end of 1 year. The model was also sensitive to the utility weights that were

applied. The company noted that the base case utility value for people with moderate to severe disease was 0.68, but that the literature has used a value in the range of 0.3 to 0.4. Using the alternative utility values reduced the ICER for vedolizumab compared with conventional therapies or the other biologics.

ERG exploratory analyses

5.32 The ERG carried out the following analyses:

- **Scenario 1** Correction of an error in the model in which baseline values for infliximab, rather than conventional therapy, were used in the maintenance model, for people who had not received TNF alpha inhibitors and who were receiving conventional therapy.
- **Scenario 2** Utility values from Woehl et al were used in the model (see table 23)
- **Scenario 3** Utility values from Swinburn et al were used in the model (see table 23)

Table 23 Utility values used in ERG exploratory analyses

Health state	Company base case	Woehl (scenario 2)	Swinburn (scenario 3)
Remission	0.86	0.87	0.91
Mild	0.8	0.76	0.8
Moderate to severe	0.68	0.41	0.55
Surgery	0.42	0.41	0.55
Post-surgery remission	0.6	0.71	0.59
Post-surgery complications	0.42	0.54	0.42

In scenario 2 and 3 it was assumed that the utility associated with surgery was the same as having moderate to severe ulcerative colitis. It was also assumed that people with post surgery complications would have a utility decrement of 0.17 relative to people in post-surgery remission, to account for the complications (the 0.17 utility decrement was based on Arseneau et al).

- **Scenario 4** Different assumptions were applied to estimate the transition probabilities between the surgery and post-surgery health states. It was assumed that:
  - people would not have repeat surgery (because the cost estimates for surgery included the cost of repeat surgery),
  - people leaving the surgery health state were assumed to remain in the post-surgery complications state, or remission state, for the remainder of the modelled time horizon.
  - The probability of experiencing late complications was based on the probability of chronic pouchitis reported in Arai et al.

Table 24: ERG Amended transition matrix for surgery and post-surgery states

	Surgery	**Post-surgery remission**	**Post-surgery complications**
Surgery	0.00	0.95	0.05
Post-surgery remission	0.00	1.00	0.00
Post-surgery complications	0.00	0.00	1.00

- **Scenario 5** People can continue to receive biologic therapies beyond 1 year if they are responding, or are in remission, to those therapies.
- **Scenario 6** Costs of conventional therapies are the same if they are taken concomitantly with a biologic therapy or if conventional therapy is the only treatment a person receives.
- **Scenario 7** Use of NHS 2012/13 reference costs for health state resource cost estimates rather than the estimates reported in Tsai et al.
- **Scenario 8** Cost of stoma care were included in the post-surgery health states for the 40% of people whose surgical procedure was assumed to have been an ileostomy. Over a 6 month period people were assumed to have 1.5 nurse visits at a cost of £136.88 and need consumables costing £178.0
- In all scenarios, except for scenario 1, the ERG also assumed a life-time time horizon rather than a 10 year time horizon. The corrections in scenario 1 were also applied in all scenarios.

5.33 The ERG presented fully incremental results for the manufacturers base case (see table 44 ERG report page 99), and the ERG's scenarios, (tables 48 to 55, pages 112 to 117 ERG report [summarised in tables 25-27 in this premeeting briefing]). The impact these scenarios is as follows:

- In all scenarios, except scenario 2, vedolizumab was the most effective option (it had the greatest modelled QALYs).
- In scenario 2, in which utility values from Woehl et al were used, surgery became the most effective option, and vedolizumab was less effective and less costly than surgery in all 3 modelled populations .
- **In the whole population** scenarios 3, 6, 7 and 8 favoured vedolizumab because they resulted in an ICER for vedolizumab compared with the next most effective treatment option (conventional therapy) that was lower than the company's base case. Scenarios 4 and 5 did not favour vedolizumab because they resulted in an ICER for vedolizumab compared with conventional therapy which was greater than the company's base case ICER.
- **In the population who have not had prior treatment with TNF alpha inhibitors** scenarios 3, 6 and 8 favoured vedolizumab because vedolizumab dominated all treatment options when these scenarios were applied. Scenario 7 also favoured vedolizumab because the ICER for vedolizumab compared with the next most effective treatment option, adalimumab, was lowered from £6,634 (in the company base case) to £759 per QALY gained. Scenario 4 (amending transition matrix for surgery and post-surgery health states) had a different impact on the ICER depending upon the comparison. When comparing vedolizumab with conventional therapy or the TNF alpha inhibitors, scenario 4 favoured vedolizumab, resulting in vedolizumab dominating or extendedly dominating these treatment options.. Compared to surgery, however, scenario 4 did not favour vedolizumab because rather than dominating surgery (as in the company's base case), the ICER for vedolizumab compared with surgery was £20,449 per QALY gained. Scenario 5 (assuming that there was no maximum treatment

duration with biologic treatments) also did not favour vedolizumab because it resulted in the ICER for vedolizumab compared with adalimumab increasing from £6,634 per QALY gained in the company base case to £3,807,239 per QALY gained. However the modelled QALY difference between these two treatments in the ERG scenario was minimal.

- **In the population in whom treatment with prior TNF alpha inhibitors had failed** ERG scenarios 3, 5,6,7,8 favoured vedolizumab because the ICER for vedolizumab compared with conventional therapies was lower than the company's base case. Scenario 4 did not favour vedolizumab because the ICER for vedolizumab compared with conventional therapy increased from £64,999 per QALY gained in the company's base case to £73,931 per QALY gained.

Table 25 ERG scenario results for whole population, most effective compared with next most effective option (exluding dominated options)

Option	Comparison	QALYs	Costs	Incremental QALYs	Incremental costs	ICER
Company base case. (fully incremental results)	Vedolizumab	5.55	£77,056	0.15	£5,131	£33,297
	Conventional therapy	5.40	£71,925			
	Surgery	4.28	£107,831	1.27	-£30,775	Surgery dominated
1: Model correction	Vedolizumab	As company base case- only affects population who have not had prior TNF alpha inhibitor where infliximab was included as comparator				
	Conventional therapy					
2 Woehl utility values	Surgery	13.05	£248,631	2.21	£48,285	£21,881*
	Vedolizumab	10.84	£200,346	* Vedolizumab less effective and less costly than surgery		
3: Swinburn utility values	Vedolizumab	11.36	£200,346	0.27	£4,144	£15,267
	Conventional therapy	11.09	£196,202			
4: Amended surgery health state transition probabilities	Vedolizumab	13.29	£131,111	0.17	£7,478	£44,114
	Conventional therapy	13.12	£123,634			
5: No maximum treatment duration with biologics	Vedolizumab	12.85	£210,883	0.42	£14,681	£34,827
	Conventional therapy	12.43	£196,202			
6: No assumption of lower use of conventional therapies with biologics	Vedolizumab	12.63	£200,747	0.20	£4,544	£22,590
	Conventional therapy	12.43	£196,202			
7: 2012/13 NHS resource use costs	Vedolizumab	12.63	£151,516	0.20	£5,611	£27,893
	Conventional therapy	12.43	£145,905			
8: Costs of stoma care included	Vedolizumab	12.63	£204,395	0.20	£3,949	£19,630
	Conventional therapy	12.43	£200,447			

Table 26 ERG scenario results for the population who have not had prior treatment with TNF alpha inhibitors; most effective compared with next most effective option (excluding dominated options)..

Option	Comparison	QALYs	Costs	Incremental QALYs	Incremental costs	ICER
Company base case.	Vedolizumab	5.90	£69,075	0.14	£918	£6,634
	Infliximab	5.82	£73,952	-	-	Dominated
	Golimumab	5.79	£70,387	-	-	Dominated
	Adalimumab	5.76	£68,157	0.21	£751	£3,664
	Conventional therapy	5.56	£67,406	-	-	-
	Surgery	4.28	£107,831	-	-	Dominated
1: Model correction	Vedolizumab	5.90	£69,400	0.14	£909	£6,469
	Adalimumab	5.75	£68,492	n.b infliximab remains dominated		
2 Woehl utility values	Surgery	13.05	£248,631	1.15	£66,163	£57,725
	Vedolizumab	11.90	£182,468	Vedolizumab less effective and less costly than surgery		
3: Swinburn utility values	Vedolizumab	12.50	£182,468	Vedolizumab dominates all options		
4: Amended surgery health state transition probabilities	Vedolizumab	14.01	£125,340	1.70	£34,696	£20,449
	Surgery	12.31	£90,645	-	-	-
5: No maximum treatment duration with biologics	Vedolizumab	13.39	£207,168	0.00	£11,711	£3,807,239
	Adalimumab	13.39	£195,457	0.58	£8,065	£13,908
	Golimumab	13.10		-	-	-
6: No assumption of lower use of conventional therapies with biologics	Vedolizumab	13.47	£182,961	Vedolizumab dominates all options		
7: 2012/13 NHS resource use costs	Vedolizumab	13.47	£139,548	0.27	£203	£759
	Adalimumab	13.20	£139,727			
8: Costs of stoma care included	Vedolizumab	£13.47	£185,809	Vedolizumab dominates all options		

Table 27 ERG scenario results for the population in whom previous treatment with TNF alpha inhibitors has failed; most effective compared with next most effective option (excluding dominated options)).

Option	Comparison	QALYs	Costs	Incremental QALYs	Incremental costs	ICER
Company base case. (fully incremental results)	Vedolizumab	5.46	£78,409	0.09	£5,839	£64,999
	Conventional therapy	5.37	£72,570			
	Surgery	4.28	£107,831			Dominated
1: Model correction	Vedolizumab	Same as company base case- only affects population who have not had prior TNF alpha inhibitor where infliximab was included as comparator				
	Conventional therapy					
2 Woehl utility values	Surgery	13.05	£248,631	2.49	£51,599	£20,714
	Vedolizumab	10.56	£182,468	Vedolizumab less effective and less costly than surgery		
3: Swinburn utility values	Vedolizumab	11.20	£202,259	0.16	£5,227	£33,472
	Conventional therapy	11.04	£197,032			
4: Amended surgery health state transition probabilities	Vedolizumab	13.19	£131,271	0.10	£7,300	£73,931
	Conventional therapy	13.09	£123,971			
5: No maximum treatment duration with biologics	Vedolizumab	12.89	£212,963	0.49	£15,931	£32,524
	Conventional therapy	12.40	£197,032			
6: No assumption of lower use of conventional therapies with biologics	Vedolizumab	12.52	£202,608	0.12	£5,577	£47,087
	Conventional therapy	12.40	£197,032			
7: 2012/13 NHS resource use costs	Vedolizumab	12.52	£152,558	0.12	£6,072	£51,271
	Conventional therapy	12.40	£146,486			
8: Costs of stoma care included	Vedolizumab	12.52	£206,410	0.12	£5,105	£43,108
	Conventional therapy	12.40	£201,305			

5.34 The ERG combined all of its scenarios, except scenario 3 (utility values from Swinburn), in its exploratory base case. The results are presented for a lifetime time horizon. In all 3 populations all options are dominated by surgery (surgery is more effective and less costly). The ERG noted that surgery may not be an acceptable treatment option for all people. The ERG stated that if surgery is not an acceptable option:

- In the whole population the ICER for vedolizumab compared with conventional therapy is £53,084 per QALY gained
- In the population who have not had prior treatment with TNF alpha inhibitors vedolizumab is dominated by adalimumab
- In the population in whom treatment with a prior TNF alpha inhibitor has failed the ICER for vedolizumab compared with conventional therapy is £48,205 per QALY gained.

Table 28: ERG-preferred base case

Option	QALYs	Costs	Incremental QALYs	Incremental costs	ICER
Mixed ITT population					
Surgery	14.60	£65,204	-	-	dominating
Vedolizumab	11.78	£107,604	-	-	dominated
Conventional therapy	11.31	£82,940	-	-	dominated
Anti-TNF-α naïve population					
Surgery	14.60	£65,204	-	-	dominating
Adalimumab	12.39	£102,666	-	-	dominated
Vedolizumab	12.37	£115,240	-	-	dominated
Golimumab	12.05	£98,594	-	-	dominated
Infliximab	12.01	£102,916	-	-	dominated
Conventional therapy	11.73	£81,501	-	-	dominated
Anti-TNF-α failure population					
Surgery	14.60	£65,204	-	-	dominating
Vedolizumab	11.84	£110,025	-	-	dominated
Conventional therapy	11.28	£83,094	-	-	dominated

Innovation

5.35 Justifications for considering vedolizumab to be innovative:

- Vedolizumab has a different mechanism of action to all other licensed treatments for ulcerative colitis. The target of vedolizumab is α4β7 integrin, which is found in T helper lymphocytes. T helper lymphocytes are found in the gut, therefore, unlike other treatments it has a gut-specific action and has a low risk of effects outside of the gut.

6 Equalities issues

6.1 No issues were raised during the scoping process or by the company in its submission, the ERG or patient and professional groups.

7 Authors

Mary Hughes

Technical Lead

Melinda Goodall

Technical Adviser

with input from the Lead Team (Pamela Rees, Graham Ash and Eldon Spackman).

Appendix A: The European public assessment report for vedolizumab is available from:

http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/002782/WC500168530.pdf