NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Final appraisal determination

Vedolizumab for treating moderately to severely active Crohn's disease after prior therapy

This guidance was developed using the single technology appraisal (STA) process.

1 Guidance

1.1 Vedolizumab is recommended as an option for treating moderately to severely active Crohn's disease only if:

- a tumour necrosis factor-alpha inhibitor has failed (that is the disease has responded inadequately or has lost response to treatment) or
- a tumour necrosis factor-alpha inhibitor cannot be tolerated or is contra-indicated.

Vedolizumab is recommended only if the company provides it with the discount agreed in the patient access scheme.

1.2 Vedolizumab should be given as a planned course of treatment until it stops working or surgery is needed, or until 12 months after the start of treatment, whichever is shorter. At 12 months, people should be reassessed to determine whether treatment should continue. Treatment should only continue if there is clear evidence of ongoing clinical benefit. For people in complete remission at 12 months, consider stopping vedolizumab, resuming treatment if there is a relapse. People who continue vedolizumab should be

reassessed at least every 12 months to decide whether continued treatment is justified.

1.3 People whose treatment with vedolizumab is not recommended in this NICE guidance, but was started within the NHS before this guidance was published, should be able to continue treatment until they and their NHS clinician consider it appropriate to stop.

2 The technology

2.1 Vedolizumab (Entyvio, Takeda UK) is a humanised IgG1 monoclonal antibody derived from a newly engineered cell line. It is targeted against α4β7 integrin, which is expressed on certain white blood cells. α4β7 integrin is responsible for recruiting these cells to inflamed bowel tissue. It is administered by intravenous infusion.

2.2 Vedolizumab has a marketing authorisation in the UK for 'the treatment of adult patients with moderately to severely active Crohn's disease who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy or a tumour necrosis factor-alpha antagonist'. The summary of product characteristics states that the recommended dosage of vedolizumab for treating Crohn's disease is 300 mg at 0, 2 and 6 weeks, then every 8 weeks thereafter. It further notes that people who have not shown a response may benefit from a dose at week 10. If no evidence of therapeutic benefit is seen by week 14, vedolizumab should not be continued.

2.3 Vedolizumab's summary of product characteristics lists nasopharyngitis (inflammation of the nose and throat), headache and arthralgia (joint pain) as very common adverse reactions. For full details of adverse reactions and contraindications, see the summary of product characteristics.

2.4 The list price of vedolizumab is £2050 per 300-mg vial (excluding VAT; British national formulary, accessed online July 2015). The company has agreed a patient access scheme with the Department of Health. This scheme provides a simple discount to the list price of vedolizumab with the discount applied at the point of purchase or invoice. The level of the discount is commercial in confidence. The Department of Health considered that this patient access scheme does not constitute an excessive administrative burden on the NHS.

3 The company's submission

The Appraisal Committee (section 8) considered evidence submitted by Takeda and a review of this evidence by the Evidence Review Group (ERG; section 9). In addition to its original submission, the company submitted a second version of its economic model at the clarification stage. The company also received permission to submit new evidence in response to the appraisal consultation document, which comprised further clinical evidence and a third version of its model.

Clinical effectiveness

3.1 The company's systematic review identified 2 randomised, double-blind, placebo-controlled trials of vedolizumab, GEMINI II and GEMINI III. No relevant non-randomised controlled trials providing clinical efficacy information were identified.

3.2 The company said the eligibility criteria for GEMINI II and GEMINI III were identical. Both trials enrolled adults with moderately to severely active Crohn's disease (Crohn's Disease Activity Index [CDAI] score 220–450) that had shown inadequate response to, loss of response to, or intolerance to at least 1 of the following: immunomodulators, TNF-alpha inhibitors or corticosteroids (outside the USA only) within the last 5 years.

Therapeutic doses of oral 5-aminosalicylates, oral corticosteroids, probiotics, anti-diarrhoeals, azathioprine or mercaptopurine, methotrexate and antibiotics were permitted. However, treatment with adalimumab within 30 days and with infliximab or certolizumab pegol within 60 days before enrolment was not permitted.

GEMINI II study design

3.3 GEMINI II compared the efficacy and safety of vedolizumab with placebo plus conventional therapy (oral prednisone or budesonide, immunosuppressive agents, mesalazine and antibiotics were permitted) for moderately to severely active Crohn's disease. It comprised an induction trial (weeks 0–6) and a maintenance trial (weeks 6–52), giving an overall study duration of 52 weeks.

3.4 In the blinded induction trial (cohort 1), patients received vedolizumab 300 mg intravenously (n=220) or placebo (n=148) at weeks 0 and 2. Randomisation was stratified by concomitant use of corticosteroids and immunosuppressive agents or previous use of TNF-alpha inhibitors, or both. The proportion of patients with previous exposure to TNF-alpha inhibitors was limited to 50% (50.5% and 48.6% in the vedolizumab and placebo arms respectively). To fulfil sample-size requirements for the maintenance trial, 748 additional patients were assigned treatment in an open-label group (cohort 2), of whom 747 patients received the same vedolizumab regimen as that in cohort 1 in the blinded induction trial. Before the trial, 58% of all patients had experienced failure of a TNF-alpha inhibitor.

3.5 In the maintenance trial, patients from both cohorts who had a clinical response with vedolizumab at week 6 (that is, a 70-point or greater decrease in the CDAI score; n=461) were randomly assigned to continue in a blinded fashion to receive vedolizumab every 8 weeks (n=154), vedolizumab every 4 weeks (n=154), or

placebo (n=153), for up to 52 weeks. Randomisation was stratified according to (1) participation in cohort 1 or 2 during induction, (2) concomitant use of corticosteroids and (3) concomitant use of immunosuppressive agents or previous use of TNF-alpha inhibitors, or both. Patients from either cohort whose disease did not have a clinical response at week 6 to vedolizumab induction therapy (n=412) received maintenance treatment with vedolizumab 300 mg every 4 weeks and were followed to week 52. Patients in the placebo group of cohort 1 who completed induction treatment (n=137) continued to receive placebo and were also followed to week 52. Of 814 patients who received vedolizumab in GEMINI II, 295 completed week 52 assessments and enrolled in GEMINI LTS (an ongoing, single-arm, open-label safety study).

3.6 In GEMINI II, the primary outcomes during induction at week 6 were clinical remission (CDAI score 150 points or less) and enhanced clinical response (a 100-point or greater decrease in the CDAI score). During maintenance, the primary outcome was clinical remission at week 52. Secondary outcomes included CDAI-100 response and corticosteroid-free remission at week 52. Safety outcomes were included and quality of life was assessed using the Inflammatory Bowel Disease Questionnaire (IBDQ), 36-Item Short Form Health Survey (SF-36), and European quality-of-life 5-domain scale (EQ-5D) questionnaires (at screening and before dosing at weeks 6, 30 and 52).

3.7 The main analyses in the induction study of GEMINI II used the intention-to-treat population, which included all patients in cohort 1 who were randomised and received at least 1 dose of blinded study drug (n=148 in the placebo arm and n=220 in the vedolizumab arm). The maintenance study analyses also used the intention-to-treat population, which included patients who received vedolizumab whose disease had a clinical response at week 6 and who were

then randomised to vedolizumab either every 4 or 8 weeks (n=154 each) or placebo (n=153). The non-intention-to-treat population in the maintenance study was included in the safety assessment, comprising 814 patients who had vedolizumab and 301 patients who received placebo. In addition to patients in the intention-to-treat population, it included patients who received placebo in the induction phase and remained on placebo for the maintenance phase, and patients whose disease did not respond to vedolizumab by week 6 of the induction study. The company also presented subgroup analyses for patients in whom a TNF-alpha inhibitor had failed (vedolizumab every 8 weeks [n=82], vedolizumab every 4 weeks [n=77] and placebo [n=78]) and for patients who had not had a TNF-alpha inhibitor before (vedolizumab every 8 weeks [n=66], vedolizumab every 4 weeks [n=71] and placebo [n=71]).

GEMINI III study design

3.8 GEMINI III was a 10-week study that evaluated the efficacy of vedolizumab compared with placebo. Patients were randomised to receive vedolizumab 300 mg (n=209) or placebo (n=207) at weeks 0, 2 and 6 and stratified according to whether they had previously had TNF-alpha inhibitor treatment (there were 315 patients in whom a TNF-alpha inhibitor had failed and 101 patients who had not had a TNF-alpha inhibitor before). They were also stratified according to concomitant use of (a) oral corticosteroids and (b) concomitant use of immunomodulators (mercaptopurine, azathioprine, or methotrexate). At the end of the study, patients could enrol in GEMINI LTS. All randomised patients received at least 1 dose of blinded study drug and were included in the intention-to-treat population.

3.9 The primary analysis of GEMINI III focused on the 315 patients in whom a TNF-alpha inhibitor had failed. The primary outcome was clinical remission at week 6 (CDAI score 150 points or less). A

secondary analysis evaluated the overall population including patients who had not had TNF-alpha inhibitor treatment before. Secondary outcomes included clinical remission at week 6 in the overall population, clinical remission at week 10 in the population in whom TNF-alpha inhibitor treatment had failed and in the overall population, sustained clinical remission (CDAI score 150 points or less at both week 6 and week 10) in the population in whom TNF-alpha inhibitor treatment had failed and in the overall population, and safety outcomes. Other outcomes included health-related quality of life, as shown by change from baseline in IBDQ, SF-36, and EQ-5D scores at weeks 6 and 10.

GEMINI II – induction phase results

3.10 The results for the primary outcomes of GEMINI II showed that at week 6, clinical remission rates (CDAI score 150 points or less) were significantly higher in patients having vedolizumab than in patients having placebo (14.5% [95% CI 9.9 to 19.2] and 6.8% [95% CI 2.7 to 10.8] respectively, p=0.02). There was no significant difference in enhanced clinical response (a 100-point or greater decrease in the CDAI score) at week 6 between the vedolizumab and placebo groups (31.4% [95% CI 25.2 to 37.5] compared with 25.7% [95% CI 18.6 to 32.7] respectively; p=0.23).

3.11 The company carried out pre-specified subgroup analyses for the primary outcomes, investigating the influence of baseline characteristics on treatment effect. For vedolizumab compared with placebo, the analyses showed a between-treatment difference in clinical remission at week 6 of 8.2% in the population of patients who had not had a TNF-alpha inhibitor before (no further details were provided). In the new evidence provided in response to the appraisal consultation document, the company reported that clinical remission at week 6 in the population in whom a TNF-alpha inhibitor had failed was experienced by 10.5% of patients receiving

vedolizumab and 4.3% of patients receiving placebo (95% CI for between-treatment difference −9.1 to 21.3). It further reported that an exploratory analysis in the same population showed that clinical remission at week 10 was experienced by 16.0% of patients in the vedolizumab group (n=5) and 8.6% of patients in the placebo group (n=70) (95% CI for between-treatment difference 0.2 to 14.6).

3.12 The company presented results for changes in health-related quality of life from baseline to week 6 in the vedolizumab group (n=211) and the placebo group (n=146) in the overall population. The company advised that a decrease of at least 0.3 points in the EQ-5D score represented a clinically meaningful improvement in health-related quality of life. Adjusted mean change in EQ-5D score from baseline was −0.5 (95% CI −0.7 to −0.3) in patients who received vedolizumab and −0.3 (95% CI −0.5 to −0.0) in patients who received placebo, giving a difference in adjusted change of –0.2 (95% CI −0.5 to 0.1). The company did not present the change in EQ-5D scores for the population who had not had a TNF-alpha inhibitor before. In the new evidence submitted in response to the appraisal consultation document, the company reported that there was no clinically meaningful decrease in EQ-5D score from baseline to week 6 with vedolizumab in the population in whom a TNF-alpha inhibitor had failed. The adjusted mean change in EQ-5D score (95% CI) was −0.2 (−0.5 to 0.1) in the vedolizumab group (n=103) and −0.2 (−0.6 to 0.1) in the placebo group (n=69).

GEMINI II – maintenance phase results

3.13 In the intention-to-treat population in the maintenance study, there were statistically significantly higher rates of clinical remission (CDAI score 150 points or less) at week 52 in patients who received vedolizumab every 8 weeks or every 4 weeks, compared with patients who received placebo. In patients receiving vedolizumab every 8 weeks, the treatment difference from placebo

was 17.4% (95% CI 7.3 to 27.5, p=0.0007) and in patients receiving vedolizumab every 4 weeks, it was 14.7% (95% CI 4.6 to 24.7, p=0.0042).

3.14 Clinical remission rates were higher for patients who had vedolizumab every 4 or 8 weeks compared with those who had placebo regardless of prior TNF-alpha inhibitor use.

- In the population in whom a TNF-alpha inhibitor had failed, the remission rate was 28.0% in the vedolizumab every 8 weeks group and 12.8% in the placebo group, giving a treatment difference of 15.2% (95% CI 3.0 to 27.5). The remission rate in the vedolizumab every 4 weeks group was 27.3%, giving a treatment difference of 14.5% (95% CI 2.0 to 26.9) compared with placebo.
- In the population who had not had a TNF-alpha inhibitor before, the remission rate was 51.5% in the vedolizumab every 8 weeks group and 26.8% in the placebo group, giving a treatment difference of 24.8% (95% CI 8.9 to 40.6). The remission rate in the vedolizumab every 4 weeks group was 46.5%, giving a treatment difference of 19.7% (95% CI 4.2 to 35.2) compared with placebo.

3.15 The company presented results for the changes in health-related quality of life from baseline to week 52 in the groups having vedolizumab every 8 weeks (n=79), vedolizumab every 4 weeks (n=92) and placebo (n=81) in the overall population. Adjusted mean change in EQ-5D score from baseline was −1.5 (95% CI −1.8 to −1.2) in patients having vedolizumab every 8 weeks, −1.4 (95% CI −1.7 to −1.1) in patients having vedolizumab every 4 weeks and −1.0 (95% CI −1.3 to −0.7) in patients having placebo. The mean difference in adjusted change from baseline compared with placebo was –0.5 (95% CI −0.9 to −0.1) for vedolizumab every

8 weeks and –0.4 (95% CI −0.8 to 0.0) for vedolizumab every 4 weeks. The company did not provide results for the changes in health-related quality of life from baseline to week 52 for the sub-populations who had not had a TNF-alpha inhibitor before or in whom a TNF-alpha inhibitor had failed.

GEMINI III

3.16 For the primary outcome of clinical remission at week 6 in people for whom TNF-alpha inhibitor treatment had failed, no statistically significant difference was seen between the vedolizumab (15.2% [95% CI 9.6 to 20.8]) and placebo (12.1% [95% CI 7.0 to 17.2]) groups (p=0.433). However, an exploratory analysis found a higher proportion of these patients had clinical remission at week 10 with vedolizumab than with placebo (26.6% [95% CI 19.7 to 33.5]) compared with 12.1% [95% CI 7.0 to 17.2] p=0.0012 [nominal p-value]).

3.17 An exploratory analysis of the overall population including patients who had not had a TNF-alpha inhibitor before showed that clinical remission occurred in a higher proportion of patients having vedolizumab than placebo at week 6 (19.1% [95% CI 13.8 to 24.5] compared with 12.1% [95% CI 7.6 to 16.5], p=0.0478 [nominal p-value]) and week 10 (28.7% [95% CI 22.6 to 34.8] compared with 13.0% [95% CI 8.5 to 17.6], p<0.0001 [nominal p-value]).

3.18 The company provided results for changes in health-related quality of life from baseline to weeks 6 and 10 for patients in the population in whom a TNF-alpha inhibitor had failed. At week 6, adjusted mean change in EQ-5D score was −0.4 (95% CI −0.6 to −0.2) in patients who had vedolizumab (n=158) and −0.1 (95% CI −0.3 to 0.1) in patients who had placebo (n=149), giving a mean difference in adjusted change from baseline of −0.2 (95% CI −0.5 to 0.1). At week 10, adjusted mean change in EQ-5D score

was −0.6 (95% CI −0.8 to −0.4) in patients who had vedolizumab (n=152) and −0.1 (95% CI −0.4 to 0.1) in patients who had placebo (n=143), giving a mean difference in adjusted change from baseline of −0.5 (95% CI −0.8 to −0.2). Similar results were seen in the overall study population.

Pooled analysis of GEMINI II and III

3.19 In the new evidence submitted in response to the appraisal consultation document, the company described a post-hoc analysis of pooled data from GEMINI II and III in the population in whom a TNF-alpha inhibitor had failed. At 6 weeks there was no statistically significant difference between the vedolizumab group (n=263) and placebo group (n=227) in the proportion of patients in clinical remission (13.3% compared with 9.7% respectively, p=0.157). However, at 10 weeks the proportion of patients in clinical remission was statistically significantly higher with vedolizumab than with placebo (21.7% compared with 11.0%, p=0.0008).

Adverse effects of treatment

3.20 In the 52-week GEMINI II study, 706 patients (87%) taking vedolizumab and 246 patients (82%) taking placebo had an adverse event. A higher proportion of patients had a serious adverse event in the vedolizumab group compared with the placebo group (24.4% and 15.3% respectively). Serious infection affected 45 patients (5.5%) taking vedolizumab and 9 patients (3.0%) taking placebo. The most common adverse event was exacerbation of Crohn's disease, which occurred in 164 patients (20.1%) in the vedolizumab group and 65 patients (21.6%) in the placebo group.

3.21 In the 10-week GEMINI III study, 117 patients (56%) taking vedolizumab and 124 patients (60%) taking placebo had an adverse event. Serious adverse events occurred in 13 patients

(6%) taking vedolizumab and 16 patients (8%) taking placebo. Less than 1% of patients taking vedolizumab and 0% of patients taking placebo had a serious infection. Common adverse events were Crohn's disease (3% of patients taking vedolizumab and 10% of patients taking placebo), headache (5% of patients taking vedolizumab and 7% of patients taking placebo), nausea (6% of patients taking vedolizumab and 2% of patients taking placebo) and fever (3% of patients taking vedolizumab and 6% taking placebo).

3.22 The company's submission also included safety data from 3 additional sources: GEMINI LTS, a pooled safety analysis of GEMINI I (ulcerative colitis) and GEMINI II, and an integrated safety analysis of 6 randomised placebo-controlled trials of vedolizumab in inflammatory bowel disease (ulcerative colitis and Crohn's disease). The company noted that no cases of progressive multifocal leukoencephalopathy had been identified in any of the safety populations.

GEMINI LTS

3.23 In its new evidence submitted in response to the appraisal consultation document, the company included analyses from GEMINI LTS of clinical remission for up to 2 years. Clinical remission was defined as a Harvey–Bradshaw Index score of 4 points or less. In the population in whom TNF-alpha inhibitors had failed (n=136), clinical remission was 52% at week 52, 56% at week 80 and 51% at week 104.

Network meta-analyses

3.24 In the absence of direct trial evidence, the company carried out a systematic review and network meta-analyses to calculate relative treatment effects for vedolizumab compared with other biological therapies (adalimumab and infliximab) for treating moderate to severe Crohn's disease. Depending on available data, the

company compared outcomes for clinical remission (CDAI score of 150 or less), clinical response (drop in CDAI score of 70 or greater), enhanced clinical response (drop in CDAI score of 100 or greater) and discontinuation because of adverse events.

3.25 In its main submission, the company identified 10 studies providing information on vedolizumab, infliximab and adalimumab and included 6 of these in its primary analysis:

- CLASSIC I, an induction study that compared adalimumab with placebo in patients who had not had TNF-alpha inhibitor treatment before.
- Sandborn (2007), an induction study that compared adalimumab with placebo in patients who had previously had TNF-alpha inhibitor treatment.
- Targan (1997), an induction study that compared infliximab with placebo in patients who had not had TNF-alpha inhibitor treatment before.
- ACCENT I, a maintenance study that compared infliximab with placebo in patients who had not had TNF-alpha inhibitor treatment before.
- GEMINI II, an induction and maintenance study that compared vedolizumab with placebo in patients who had not had TNF-alpha inhibitor treatment before and patients in whom previous TNF-alpha inhibitor treatment had failed.
- GEMINI III, an induction study that compared vedolizumab with placebo in patients who had not had TNF-alpha inhibitor treatment before and patients in whom previous TNF-alpha inhibitor treatment had failed.

It considered that the 4 other studies were not comparable because they did not provide enough detail about whether patients had previously had TNF-alpha inhibitor treatment or not (Watanabe

[2012], CHARM and EXTEND) or because patients were re-randomised based on remission rather than on response (CLASSIC II).

3.26 The company conducted Bayesian fixed-effects and random-effects analyses for the following groups in the primary analysis in its main submission:

- patients having induction treatment who had not had TNF-alpha inhibitors before
- patients having maintenance treatment who had not had TNF-alpha inhibitors before
- patients having induction treatment who had previously had TNF-alpha inhibitor treatment.

The company stated that it validated its Bayesian analyses by running equivalent frequentist models, and that the point estimates and credible intervals closely matched.

3.27 In its main submission, the company advised that it was not able to provide all relevant indirect comparisons, and that caution should be used when interpreting some results because of data limitations:

- It was not possible to construct a network for maintenance treatment in the population who had previously had TNF-alpha inhibitor treatment (see section 3.28 for further discussion).
- In the network of patients in whom a TNF-alpha inhibitor had failed and who were having induction treatment:
  - none of the trials included infliximab
  - the vedolizumab studies included patients whose disease had previously responded inadequately (that is, primary non-response), had lost response or became intolerant to a TNF-alpha inhibitor whereas the comparator study included

only those whose disease lost response or became intolerant to TNF-alpha inhibitor therapy.

- Results for the 'mixed' population (that is, all patients regardless of TNF-alpha inhibitor status) were provided as a secondary analysis. The company noted that the placebo response rates in GEMINI II were inexplicably higher than in the other studies and considered that this could bias the results against vedolizumab. It therefore considered it more appropriate to use the subgroup analyses, rather than the whole population ones that may be affected by confounding factors.

3.28 The company used results for clinical remission and clinical response from a network meta-analysis for a population who had not had a TNF-alpha inhibitor before in its economic analyses. The company submission did not state if the results were from the Bayesian or frequentist analyses. In the new evidence submitted in response to the appraisal consultation document, the company further considered a network meta-analysis comparing vedolizumab with adalimumab and infliximab in the population for whom a TNF-alpha inhibitor had failed. It noted a lack of comparable evidence to inform a comparison of vedolizumab and the other biological agents. No evidence was identified from the systematic review for infliximab and the adalimumab trials were in patients who had experienced secondary failure of a TNF-alpha inhibitor, whereas GEMINI II and III enrolled patients who had experienced primary failure (inadequate or lack of response) or secondary failure (loss of response). It concluded that this meant the patients in the vedolizumab trials would be expected to have a poorer prognosis, and that the results from the network meta-analysis would be likely to underestimate the relative effectiveness of vedolizumab. The company stated that its network meta-analyses for induction and maintenance showed no significant difference in clinical response (induction only) or clinical remission.

However, the company did not use these results in its third economic model because it considered that it was not a robust comparison (see section 3.47).

ERG comments

3.29 The ERG considered the company's methods for performing the clinical effectiveness systematic review to be largely appropriate and was satisfied that all relevant vedolizumab studies had been included in the company's submission.

3.30 The ERG noted that GEMINI II and III assessed response in the induction phase at 6 weeks, which did not correspond with the recommended dosage in vedolizumab's summary of product characteristics. It considered assessment at 6 weeks to be earlier than in routine clinical practice in England, after receiving expert clinical advice that this is typically done 10–14 weeks after starting treatment.

3.31 The ERG noted high discontinuation rates in the maintenance phase of GEMINI II, which it considered could limit the robustness of the efficacy and safety data and undermine external validity. It highlighted that high rates of discontinuation were seen across all treatment groups (58% [89/153 patients] in the placebo arm, 53% [81/154 patients] in the vedolizumab every 8 weeks arm and 47% [72/154 patients] in the vedolizumab every 4 weeks arm).

3.32 The ERG had concerns about the generalisability and treatment duration of the GEMINI II and III trial populations to the population who would be expected to have vedolizumab in clinical practice in England:

- There was a large number of US-based study sites but apparently few UK-based study sites.

- In the USA, failure of either an immunomodulator (mercaptopurine or azathioprine) or a TNF-alpha inhibitor was required, but failure of corticosteroids alone was sufficient for study entry outside the USA.
- The ERG received clinical advice that the concomitant conventional therapy used in the GEMINI trials may not wholly reflect that used in clinical practice in England.
- Response in the induction phase was assessed earlier than in clinical practice in England (see section 3.30).
- The long-term efficacy and safety of vedolizumab is unknown because treatment duration in GEMINI II was 52 weeks, followed by enrolment in the ongoing GEMINI LTS study.

3.33 The ERG was satisfied that all relevant studies had been identified for potential inclusion in the network meta-analysis, apart from data from a trial by Watanabe et al. (2012) for the induction period in patients who had not had a TNF-alpha inhibitor before. It believed that the impact of this exclusion was likely to be relatively small because it was a small study (n=57).

3.34 The ERG noted that, although the company had stated that Bayesian and frequentist fixed-and random-effects models were conducted, not all models were reported within the company submission. The ERG considered that the results of the network meta-analyses may underestimate the uncertainty in treatment effects because fixed-effects models were used, and that there was clear evidence of heterogeneity among the trials included in the network meta-analyses.

3.35 The ERG noted that there was variation between studies in the inclusion of patients with strictures and a lack of clarity around the proportion of patients with fistulising disease. It also noted that the studies did not include patients with the upper range of severe

disease (CDAI score greater than 450). The ERG concluded that the generalisability of the results to these groups of patients was unclear.

3.36 The ERG noted that the mixed population analysis (presented by the company as a secondary analysis) included trials with different proportions of characteristics thought to affect the outcome of treatment (that is, the proportion of patients in whom a TNF-alpha inhibitor had failed), making it difficult to interpret the results and to generalise to any particular population.

3.37 The ERG considered that the network meta-analysis for the population who had previously had TNF-alpha inhibitor treatment may have overestimated efficacy for adalimumab, because patients whose disease had a primary non-response to TNF-alpha inhibitor therapy were excluded from the adalimumab study but not the vedolizumab studies. It agreed with the company that the analysis would not give a robust assessment of comparative treatment effects because of differences in patient populations.

3.38 The ERG concluded that the analysis for the population who had not had a TNF-alpha inhibitor before was the best match to patients presenting after conventional therapy failed in clinical practice in England. It also concluded the following about the company's network meta-analysis results:

- Induction phase:
  - When assessing response during induction, the ERG preferred using 10-week data to 6-week data because it had received clinical expert advice that response is typically assessed at 10–14 weeks in clinical practice, and because of the recommended dosage in the marketing authorisation.
  - If the Targan et al. (1997) study comparing infliximab with placebo as an induction therapy was included in the network

(which the ERG considered to be appropriate), treatment with infliximab led to significantly higher rates of clinical response and clinical remission than vedolizumab.

  - Regardless of the inclusion of Targan et al., there was insufficient evidence to conclude that there is a difference in efficacy between vedolizumab and adalimumab.

- Maintenance phase:
  - All of the presented analyses had limitations, for example patients entered the maintenance phase after their treatment response was assessed earlier than commonly done in clinical practice, which the ERG considered could affect estimates of efficacy and limit generalisation to patients whose disease takes longer to respond.
  - The ERG noted that the company's submission did not present maintenance data including adalimumab in its primary analyses because it had excluded the CLASSIC II study (which compared adalimumab with placebo). The ERG extracted this information from a supporting reference document provided by the company, and considered that the network meta-analyses including CLASSIC II showed that the relative efficacy of vedolizumab and adalimumab was uncertain and that it was likely that vedolizumab was less effective than infliximab.
  - In the network meta-analysis excluding CLASSIC II, vedolizumab appeared significantly better than infliximab for discontinuations because of adverse events (though the ERG advised that this should be interpreted with reference to the numbers who discontinued for each treatment in the induction period). The ERG noted that the statistical significance of the difference in response between vedolizumab and infliximab 5 mg was not reported by the company, and that there was no

statistically significant difference in remission between vedolizumab and infliximab.

- The ERG was not convinced by the company's argument for excluding CLASSIC II, and believed that networks including and excluding it should be examined. It also considered that a better approach could have been to use a random-effects analysis to formally consider heterogeneity, and that it may have been valid to consider that no network was possible because of clinical heterogeneity.

3.39 The ERG noted that the trial of vedolizumab maintenance therapy (GEMINI II) was not of sufficient size or duration to estimate the risk of uncommon adverse events.

Company's economic model

3.40 The company submitted a de novo economic model that compared vedolizumab with conventional non-biological therapy and with TNF-alpha inhibitors in patients with moderately to severely active Crohn's disease. In its response to clarification, the company provided a second model that addressed some of the issues and uncertainties that had been identified. In its new evidence submitted in response to the appraisal consultation document, the company provided a third model that focused only on the subgroup of patients for whom a TNF-alpha inhibitor had failed and addressed some of the Committee's concerns expressed in the appraisal consultation document.

3.41 The company used a 2-part model to capture the different phases of treatment in the clinical trials: a decision tree for the induction phase (6 weeks) and a Markov model (as a cohort transition model) for the maintenance phase. The Markov model was largely consistent with a previous model by Bodger et al. (2009) that compared infliximab and adalimumab for treating Crohn's disease.

It had a cycle length of 8 weeks (with half-cycle correction). The time horizon was 10 years in the first 2 models and lifetime in the third model. A discount rate of 3.5% was applied to costs and health benefits and the analysis was conducted from an NHS perspective (the company explained that personal social services were expected to be minimal in this population).

3.42 In the induction phase, patients started treatment with vedolizumab, infliximab, adalimumab or conventional non-biological therapy to induce a response (defined as a drop of at least 70 points of the CDAI score). In its first 2 models, the company chose a 6-week induction phase to be consistent with the vedolizumab clinical trials. In its third model, the company changed the induction period so response was assessed at 10 weeks, which was consistent with the marketing authorisations and clinical practice. Conventional non-biological therapy comprised aminosalicylates, corticosteroids and immunomodulators. Standard doses were assumed and the treatment mix was based upon the report of the UK Inflammatory Bowel Disease Audit Steering Group (Royal College of Physicians 2013).

3.43 Patients who entered the induction phase on a biological therapy (that is, all treatments except conventional non-biological therapy) and whose disease responded to treatment entered the Markov model for maintenance therapy and continued to receive biological therapy (unless they had stopped treatment because of adverse events). If their condition did not respond, or if they had stopped treatment because of adverse events, they switched to conventional non-biological therapy. Patients who entered the induction phase on conventional non-biological therapy could respond to treatment and enter the Markov model for conventional non-biological therapy. If their condition did not respond, they were assumed to remain in the moderate–severe disease health state for

the remainder of the model time horizon or until surgery. Regardless of response status at the end of the induction phase, patients taking conventional non-biological therapy remained on this treatment for the remainder of the model time horizon.

3.44 The modelled health states in the Markov model for maintenance therapy were remission (CDAI score 150 or less), mild (CDAI score 150–220), moderate–severe (CDAI score 220–600), surgery and death. Patients could transition between each of the 4 disease severity health states (remission, mild, moderate–severe, and surgery) or experience death. It was assumed that treatment with a biological therapy was limited to 1 year, when patients switched to conventional non-biological therapy (applied correctly in the second and third models at cycle 7 [week 54] instead of at cycle 6 [week 46], as in the original model). If patients were having biological therapy, they could stop treatment because of loss of response or adverse events (whereas conventional non-biological therapy was assumed to continue until surgery or the end of the model's time horizon). In the moderate–severe health state, patients stopped treatment after 1 year because of lack of response and switched to conventional non-biological therapy or surgery. After surgery, patients could transition to active treatment in a CDAI-based health state or remain in the surgery health state. The model used an age- and sex-specific mortality risk, which was adjusted for time spent in each health state.

Population

3.45 In its first 2 models, the company defined 3 patient groups with moderately to severely active Crohn's disease (CDAI score 220–450 points) who had an inadequate response with, lost response to, or are intolerant to either conventional non-biological therapy or TNF alpha inhibitors.

- In the first model, the populations were:
  - the mixed population (included patients who had not had a TNF-alpha inhibitor before and patients in whom a TNF-alpha inhibitor had failed, representing the intention-to-treat trial populations)
  - the population who had not had a TNF-alpha inhibitor before
  - the population in whom a TNF-alpha inhibitor had failed (both primary failure and secondary failure).
- The second model additionally included data for the subgroups defined by both prior use of TNF-alpha inhibitors and severity of disease at baseline.

The company compared vedolizumab with conventional non-biological therapy in all of these populations, but said it compared vedolizumab with the other biological treatments only in patients who had not had a TNF-alpha inhibitor before because of limitations in the data (see section 3.27).

3.46 In its third model, the company focused entirely on the subgroup of patients with moderately to severely active Crohn's disease (CDAI score 220–450 points) in whom a TNF-alpha inhibitor had failed (primary and secondary failure). The cost-effectiveness results for this subgroup generated using the third model superseded those generated using the first 2 models. The company's third model evaluated the cost effectiveness of vedolizumab compared with conventional non-biological therapy. Because of data limitations, the company carried out an exploratory cost comparison of vedolizumab compared with other biological treatments (adalimumab and infliximab), which included scenario analyses on comparator dose escalation based on clinical expert opinion.

Clinical parameters and transition probabilities

3.47 Treatment efficacy included response and remission data for the induction phase, and the probability of being in remission or having mild disease at the end of 1 year (the maintenance phase of the GEMINI II study):

- For comparisons between vedolizumab and conventional non-biological therapy in the mixed population and in the population in whom a TNF-alpha inhibitor had failed, the company used head-to-head results of GEMINI II and GEMINI III to estimate treatment efficacy.
- For the comparisons between vedolizumab and the other biological therapies (infliximab and adalimumab) in the population who had not had a TNF-alpha inhibitor before, the clinical parameters for vedolizumab in the company's second model were wholly derived from the network meta-analyses provided in the company's clarification response. These superseded the original analyses in which the clinical parameters for infliximab and adalimumab were derived from the network meta-analyses and those for vedolizumab (and conventional non-biological therapy) were derived from GEMINI II and III.
- For the exploratory comparison between vedolizumab and the other biological therapies (infliximab and adalimumab) using the third model with the population in whom a TNF-alpha inhibitor had failed, the company did not use the results of its network meta-analysis. Instead, it assumed equal efficacy and did not quantify differences in adverse-event profiles. Vedolizumab transition probabilities from the network meta-analysis were applied to adalimumab and infliximab.

The company's economic model defined response as a decrease in CDAI score of 70 or more from baseline and remission as a CDAI

score of 150 or less. The company assumed that, for all treatments, there was an equal percentage of patients whose disease responded but who did not move out of the moderate–severe health state.

3.48 The company estimated the efficacy of each treatment by estimating odds ratios using response and remission data from the network meta-analyses (the population who had not had a TNF-alpha inhibitor before) or from pooled trial data (the mixed population and the population in whom TNF-alpha inhibitor treatment had failed). In the population who had not had a TNF-alpha inhibitor before, infliximab data were derived from ACCENT I, separate from the network meta-analysis. The company said this was because the trial by Targan et al. in the network meta-analysis had a small sample size and did not use a standard infliximab dosage. The odds ratios were then used to estimate the percentage of patients in each health state at the end of the induction and maintenance periods. In the first 2 models, the probability of surgery was assumed to be the same across the different patient populations in the induction phase (2.03%) and maintenance phase (2.7%). In the third model, the probability of surgery was 0.83% during induction then 1.11% in each 8-week cycle, and the probability of further surgery in the next cycle was 1.11% (amended from 33.75% in the previous models).

3.49 The company based the model's starting annual mortality rate on all-cause mortality for the UK general population (0.0015). In the first 2 models, relative mortality risks were assumed to be 1.3 for mild disease, 2.3 for moderate–severe disease and 3.2 for surgery. In the third model,the relative risk of mortality was assumed to be 1 in all health states. Patients in remission were assumed to have the same mortality risk as the general UK population.

Adverse events and surgical complications

3.50 The company selected adverse events to be used in its economic model based on clinical expert opinion. They included serious infection, lymphoma, acute hypersensitivity reactions and skin reactions. In the first 2 models, the probability of each adverse event occurring with each treatment was estimated from clinical trial data included in the network meta-analyses. In the third model, the probabilities of adverse events associated with vedolizumab and conventional non-biological therapy were obtained from summaries of product characteristics. Surgical complications in the model were also based on clinical expert opinion and the probabilities of these occurring were estimated from pooled data from a systematic literature review on surgical intervention. Annual probabilities of discontinuing biological treatment owing to adverse events were derived from clinical trials.

Utility values

3.51 The company's base case used the observed EQ-5D scores from GEMINI II and GEMINI III. The company assumed a utility value for the surgery state that was equal to that for the moderate–severe health state because patients in GEMINI II and GEMINI III were not followed for surgery. The utility values used in the model were 0.820 for remission, 0.730 for mild disease and 0.570 for both moderate–severe disease and surgery. The company applied disutilities from published literature for adverse events.

Costs

3.52 Treatment acquisition costs, including the estimated doses and unit costs for conventional non-biological therapy, were taken from the British national formulary (2013; costs updated to November 2014 edition in the third model). The patient access scheme was applied to the cost of vedolizumab as a simple discount on the list price (the level of the discount is confidential; the discount was increased

in the third model). In the first 2 models, administration costs of £308 per administration in the maintenance phase and £616 in the induction phase were included for vedolizumab and infliximab (adalimumab did not have any administration costs). In the third model, a cost of £328 per administration of vedolizumab was applied in both phases.

3.53 In the first 2 models, health-state costs were taken from Bodger et al. (2009) and inflated to 2012 prices. The health-state costs were £110 for remission, £313 for mild disease, £490 for moderate–severe disease and £10,581 for surgery, which included surgical complications. In the third model, health-state costs were estimated based on a survey of 8 clinical experts in England. These were £100 for remission, £205 for mild disease, £1239 for moderate–severe disease and £7293 for surgery (including post-surgery complications).

3.54 In all models, surgery-related complication costs were estimated by applying NHS reference costs to resource use as reported by the company's clinical experts. The company estimated costs of adverse events as weighted averages according to the NHS reference costs and assumed that all affected patients were hospitalised.

Company base-case results

3.55 Using its second model submitted in response to clarification, which replaced the original model, the company provided base-case results incorporating the original vedolizumab patient access scheme for 3 populations:

- In the mixed population, vedolizumab was associated with greater costs and quality-adjusted life years (QALYs) than conventional non-biological therapy, giving an incremental cost-

effectiveness ratio (ICER) of £62,903 per QALY gained (incremental costs £8338; incremental QALYs 0.1334).

- In the population in whom a TNF-alpha inhibitor had failed, vedolizumab was associated with greater costs and QALYs than conventional non-biological therapy, giving an ICER of £98,452 per QALY gained (incremental costs £8615; incremental QALYs 0.0875). Note that the ICER for this population is superseded by that generated using the company's third model (see section 3.56).
- In the population who had not had a TNF-alpha inhibitor before, using vedolizumab trial data, vedolizumab was associated with greater incremental costs and QALYs than conventional non-biological therapy, giving an ICER of £22,718 per QALY gained (incremental costs £6402; incremental QALYs 0.282).
- In the population who had not had a TNF-alpha inhibitor before, using network meta-analysis data, vedolizumab was associated with lower QALYs and costs than infliximab, giving an ICER for infliximab compared with vedolizumab of £26,580 per QALY gained (incremental costs £917; incremental QALYs 0.034). Vedolizumab was associated with greater costs and QALYs than adalimumab, giving an ICER for vedolizumab compared with adalimumab of £758,344 per QALY gained (incremental costs £3497; incremental QALYs 0.005).

3.56 Using its third model submitted as new evidence in response to the appraisal consultation document, the company provided base-case results incorporating the updated vedolizumab patient access scheme for the population in whom a TNF-alpha inhibitor had failed. The deterministic ICER for vedolizumab compared with conventional non-biological therapy was £21,620 per QALY gained (incremental costs and QALYs are confidential) and the probabilistic ICER was £27,428 per QALY gained (95% CI −7883 to 82,947). The probability of vedolizumab being cost effective

compared with conventional non-biological therapy at a maximum acceptable ICER of £30,000 was 61%.

Company sensitivity and scenario analyses

3.57 For the original model, the company concluded that its deterministic sensitivity analyses showed that the key drivers were transition probabilities (in particular for remission), health state costs and utility values. It noted that assuming a longer time horizon in the original model made vedolizumab more cost effective in all populations. The company did not provide any sensitivity analyses using its second model.

3.58 For its third model, the company concluded that its deterministic sensitivity analyses showed that the ICER was most sensitive to variation in the conventional non-biological therapy arm transition probabilities from the moderately–severely active state and from the remission state. It noted that the scenario analyses showed that the ICER increased most markedly when the time horizon was reduced to 10 years, when enhanced response was used to determine progression to maintenance treatment (100-CDAI point decrease), when response was assessed at week 6 and when efficacy data were based solely on GEMINI II. The company's scenario analyses showed that continuing treatment with vedolizumab for 2 or 3 years (instead of 1 year) increased the ICER to £24,695 and £26,207 per QALY gained respectively.

Company exploratory analyses

3.59 The company carried out exploratory cost-comparison analyses that compared vedolizumab with adalimumab and with infliximab in the population in whom a TNF-alpha inhibitor had failed. The results of these analyses are confidential.

ERG comments on the company main submission

3.60 In its original report, the ERG critiqued the company's cost-effectiveness evidence, focusing on the second model provided at the clarification stage. The ERG was largely satisfied with the company's explanation about why it chose its model structure (adapted from Bodger et al.). However, the ERG considered the quality of the company's model to be generally poor, unnecessarily complex in its implementation and lacking detail on the sources of inputs and the derivation of the transition matrices.

Model structure

3.61 The ERG expressed concerns about the structure of the company's model in 4 main areas:

- It did not capture that Crohn's disease is a relapsing and remitting condition (that is, patients may experience spontaneous exacerbations and improvements). The company's model assumed that patients whose disease did not respond to conventional non-biological therapy at week 6 remained in the non-responder state and had moderate to severe Crohn's disease until death or surgery, which it considered to be overly pessimistic.
- Surgery was modelled as a single health state, which may be overly simplistic because subsequent surgery is likely to depend on the type of initial surgery. However, the ERG recognised the possible lack of data in this area and believed that the impact on results would be minimal.
- There were difficulties associated with parameterising the company's chosen structure, including how the transition probabilities were derived and how the model predictions were calibrated (see section 3.77).

- Some of the key structural assumptions that influenced the derivation of transition probabilities were considered debatable. These included:
  - Patients whose disease did not respond were assumed to have moderate to severe disease, which the ERG considered to be inappropriate. This is because these patients could have a drop in CDAI score of less than 70 points that would mean their disease would be reclassified as mild.
  - Except for continuing biological treatment after induction, no distinction was made between patients with moderate to severe Crohn's disease whose disease responded and patients whose disease did not respond. The ERG believed that outcomes would be likely to differ between these groups.
  - The definition of response was taken from the clinical trials, which may have limited relevance to clinical practice in England (because CDAI scores are not routinely used).
  - The same treatment duration was assumed for all therapies for the induction phase (6 weeks), which led to discrepancies in costing, cycle length and efficacy in the company's model.
  - All patients still having TNF-alpha inhibitor therapy at approximately 1 year were assumed to switch to conventional non-biological therapy. Based on the recommendations in the NICE technology appraisal guidance on infliximab (review) and adalimumab for the treatment of Crohn's disease, the ERG considered that a discontinuation rule may be appropriate for patients in remission, but not for patients whose disease is not in stable clinical remission.
  - It was assumed that there was no increase in relapse after withdrawal of biological treatment in patients in the remission or mild disease health states, which was not aligned with clinical expert opinion received by the ERG.

- The efficacy of conventional non-biological therapy was assumed to be independent of previous biological treatment (that is, conventional non-biological therapy was equally effective in patients who had previously had biological treatment as those who had not). The ERG considered that this would be unlikely.
- Discontinuation owing to lack of efficacy during the maintenance phase was not included in the company's economic model. Based on its interpretation of the data from the GEMINI trials, the ERG believed this should be incorporated.

3.62 The ERG noted that the duration of induction with the biological therapies was not always in line with UK licensing and clinical practice, meaning not all studies delivered a full induction dose in the second model:

- Vedolizumab was given in 2 doses at weeks 0 and 2 with assessment at week 6. The ERG considered it more appropriate to follow the marketing authorisation more closely by using the induction regimen from GEMINI III (that is, doses at weeks 0, 2 and 6 with assessment at week 10).
- Adalimumab was administered at 80 mg at week 0, then 40 mg at weeks 2, 4 and 6 with assessment at week 6. The ERG considered it preferable to administer 80 mg at week 0 and 40 mg at week 2, with assessment at week 4, which was more consistent with adalimumab's marketing authorisation.

The ERG believed that 3 doses of vedolizumab should be used during induction, rather than the 2 assumed in the company's base case, which would increase the treatment cost. It considered that 3 doses of adalimumab 40 mg should be given in the induction phase rather than 5 doses, which would decrease the cost. The

ERG was satisfied with the infliximab induction regimen used in the company's model because this reflected the marketing authorisation and the efficacy data used in the company's model.

Population

3.63 The ERG was unclear how results from the mixed intention-to-treat population could be interpreted. It believed that patients who had previously had TNF-alpha inhibitors and those who had never had TNF-alpha inhibitors are 2 distinct patient groups with different characteristics and likelihood of responding to treatment. It considered that the results from the 2 groups should be interpreted separately.

3.64 The ERG was satisfied that analyses according to disease severity could potentially be informative, despite not being defined in the NICE final scope. However, the ERG was unable to confirm the results of these analyses because it could not verify the calibrated transition probabilities and it was unsure how the clinical data had been estimated in the company's second model.

Clinical parameters

3.65 The ERG noted that the company had provided limited details on the network meta-analyses used in its second economic model for the population who had not had a TNF-alpha inhibitor before; how the vedolizumab clinical trial data had been pooled; and how the discontinuation rates because of adverse events had been calculated. Although the ERG recognised that the Targan trial comparing infliximab with placebo had limitations, it believed that it should have been included in the network meta-analysis for infliximab and used in the base case, potentially adjusting for small sample size. The ERG noted that the company had instead used data for infliximab from the placebo-controlled ACCENT I trial (separate from the network meta-analysis), but had not discussed

the trial's limitations. The ERG noted that including data for adalimumab from the trial by Watanabe et al. as well as CLASSIC I in the primary analysis would likely increase the probabilities of remission and response for adalimumab.

3.66 The ERG was unclear from the company submission and the publication by Bodger et al. how the transition probabilities for patients having surgery had been calculated. It considered the values used by the company for transitioning from surgery to surgery in the next cycle to be high (33.75%), and was not satisfied by the company's explanation. The ERG was not able to predict how correcting the transition matrix for movement between states after surgery would affect the ICERs.

3.67 The ERG expressed concerns about the assumptions about mortality used in the company's model because of a lack of detail in the company's submission. It noted that because mortality is conditional on the current health states in the company's model, the model predicts greater survival for patients who had biological therapy compared with patients who had conventional non-biological therapy. However, the study by Lichtenstein et al (2009), used by the company in its model, suggests no statistical differences in the excess mortality rates according to disease severity at baseline, or in mortality between patients who did or did not receive infliximab. The ERG stressed that no increased mortality rate was observed in patients taking placebo in GEMINI II. Given the lack of evidence of a differential mortality rate between treatments, the ERG believed that the same excess risk mortality should be applied to all Crohn's disease health states.

3.68 The ERG considered that the inclusion of adverse events and their impact on costs and health-related quality of life was flawed. It was unclear if all or only grade 3 or 4 adverse events had been included and noted the selection was based on the opinion of 2 clinical

experts. It found the calculations from the company to be simplistic and likely to be incorrect because they did not account for trial duration. Moreover, the ERG was unsure why data from the network meta-analysis for the incidence of serious adverse events were not used in the company's model.

Utility values

3.69 The ERG was largely satisfied with the company's approach to estimating utility values for the different health states in its model, but had some concerns:

- The same utility value was used for patients with moderate to severe disease, regardless of any response to treatment. The ERG considered that this was unlikely to be true because it implied that response (that is, improvement in symptoms) does not improve health.
- The company had assumed an equal utility value for patients having surgery as those with moderate to severe Crohn's disease. Although the ERG recognised that the GEMINI trials could not inform utility value estimates for surgery, it was unsure that the company's assumption was appropriate because the aim of surgery is to improve quality of life.
- Limited details were provided by the company regarding its approach to adjusting utility weights. However, the ERG anticipated that any impact on the ICERs would be minimal.

Cost-effectiveness results

3.70 The ERG noted that the company presented pairwise comparisons rather than a fully incremental analysis for the group who had not had a TNF-alpha inhibitor before and that it had not provided updated cost-effectiveness estimates for all of the patient groups covered by the original model. The ERG therefore extracted this information from the company's second model. In the group who

had not had a TNF-alpha inhibitor before, a fully incremental analysis gave the following ICERs:

- £19,705 per QALY gained (incremental costs £4146; incremental QALYs 0.2104) for adalimumab compared with conventional non-biological therapy
- £112,882 per QALY gained (incremental costs £4414; incremental QALYs 0.0391) for infliximab compared with adalimumab
- £758,344 per QALY gained for vedolizumab compared with adalimumab (incremental costs £3497; incremental QALYs 0.005), which was greater than that for infliximab compared with vedolizumab (£26,580 per QALY gained [incremental costs £917; incremental QALYs 0.034]). This meant that vedolizumab was subject to extended dominance.

The ERG could not confirm the results of subgroup analyses according to disease severity for the population who had not had a TNF-alpha inhibitor before because it was unclear how the data for patients who received infliximab and adalimumab had been estimated in the company's second model.

3.71 In the mixed population and in the population in whom a TNF-alpha inhibitor had failed, vedolizumab was associated with greater costs and QALYs compared with conventional non-biological therapy in subgroups according to disease severity, with ICERs of:

- £21,064 per QALY gained (incremental costs £6447; incremental QALYs 0.3061) for the mixed population with moderate disease
- £77,382 per QALY gained (incremental costs £7840; incremental QALYs 0.1013) for the mixed population with severe disease

- £55,201 per QALY gained (incremental costs £7909; incremental QALYs 0.1433) for the population in whom a TNF-alpha inhibitor had failed with moderate disease
- £134,330 per QALY gained (incremental costs £7926; incremental QALYs 0.0590) for the population in whom a TNF-alpha inhibitor had failed with severe disease.

The ERG was concerned that the number of patients with moderate to severe disease regularly did not match the number of patients with moderate disease plus the number of patients with severe disease. It also had concerns about the validity of the calibrated transition probabilities.

3.72 The ERG noted that the company had not re-run its deterministic sensitivity analyses using the second model. The ERG considered that the parameters that had the largest impact on the ICER would not change between the 2 versions of the model submitted. It agreed with the company that the key drivers of the ICER included many of the transition probabilities, health state costs and utility values. It considered the ranges used by the company for its deterministic and probabilistic sensitivity analyses to be somewhat arbitrary for most input parameters.

3.73 Using the company's second model, the ERG reported the results from the scenario analyses presented in the original company submission. It noted that the ICER was sensitive to all the scenarios considered, especially the time horizon and health state utility values.

3.74 For transparency, the ERG extracted the probabilistic ICERs using the second version of the company's model and noted that these were consistent with the deterministic ICERs. In a fully incremental comparison, the ERG reported that the probability of vedolizumab being cost effective was less than 1% at a maximum acceptable

ICER of £20,000 per QALY gained for all 3 populations in the base case of the company's second model. The probability of cost effectiveness increased to about 2% at a maximum acceptable ICER of £30,000 per QALY gained.

3.75 The ERG had concerns about the validity of the predictions made by the company's model, including several discrepancies between the results generated using the model and those from the clinical trials. It noted that for the population in whom a TNF-alpha inhibitor had failed, the company's model seemed to under-predict the proportion of patients having conventional non-biological therapy who were in remission.

ERG comments on the company's new evidence submitted in response to the appraisal consultation document

3.76 The ERG was satisfied with the most recent company changes made to the model, but continued to have concerns about the third model's structure and parameterisation. The remaining issues included:

- Potential omission of key aspects of the condition (for example, its relapsing–remitting nature).
- Simplifying and debatable assumptions regarding surgery.
- Difficulty associated with parameterising the model's structure (notably deriving transition matrices).
- Debatable assumptions including ending maintenance at 1 year, omitting discontinuation due to lack of efficacy and that patients receiving conventional non-biological therapy whose disease did not respond during the induction phase remain in the moderate–severe health state (and are not able to improve).

The ERG noted that, as with previous versions, the third model tended to under-predict the proportion of patients receiving conventional non-biological therapy who were in remission.

ERG exploratory analyses

3.77 Because of its concerns about the model structure, the ERG was not able to provide a robust ICER for vedolizumab using any of the company's 3 models. The ERG was unclear whether the vedolizumab's cost effectiveness would improve or deteriorate after addressing the structural issues.

3.78 The ERG carried out a stepwise verification of the individual changes made in building the company's third model, and identified those that had the greatest impact on the ICER for vedolizumab compared with conventional non-biological therapy in the population in whom a TNF-alpha inhibitor had failed. It found that increasing the time horizon from 10 years to lifetime reduced the ICER from the previous base-case ICER of £98,452 per QALY gained to £57,481 per QALY gained. When the health-state costs were derived from a survey of 8 clinical experts based in England rather than the literature (Bodger et al., 2009), the ICER dropped from £56,223 to £35,154 per QALY gained.

3.79 Full details of all the evidence are in the committee papers.

4 Consideration of the evidence

The Appraisal Committee reviewed the data available on the clinical and cost effectiveness of vedolizumab, having considered evidence on the nature of Crohn's disease and the value placed on the benefits of vedolizumab by people with the condition, those who represent them, and clinical experts. It also took into account the effective use of NHS resources.

4.1 The Committee discussed the impact of moderately to severely active Crohn's disease and its treatment on people who have the disease. It heard from the patient experts that disease symptoms were debilitating and that experiencing and maintaining remission were vital to maximising quality of life, including improved social life and employment prospects. For example, it noted the experience of 1 patient expert that ongoing symptoms for over 10 years without remission had forced a career change. The Committee also understood from the patient experts that disease symptoms could have a wide-ranging and devastating impact in areas such as mental health, relationships, and personal and social development (which it heard was especially important to younger people during their formative years). It heard that patients dreaded loss of remission and further flare-ups of the disease because of the major impact these have on their lives. It also heard that it was very important to have other treatment options if their treatment stopped working to enable them to regain remission. It further heard from the patient experts that they would prefer to avoid long-term corticosteroid use because of the associated side effects. The clinical and patient experts agreed that they preferred to manage Crohn's disease medically rather than surgically wherever possible. The Committee concluded that a further drug treatment that improves symptoms or brings the disease into remission would be highly valued by patients.

4.2 The Committee discussed the treatment pathway for Crohn's disease, including unmet clinical need. It heard from the clinical experts that the NICE clinical guideline on Crohn's disease was largely followed in clinical practice, with patients first having conventional non-biological treatment. It heard that the clinical experts valued using TNF-alpha inhibitors after conventional non-biological treatment failed because considerable clinical experience had been gained in using these treatments. It heard from the

clinical experts that after a TNF-alpha inhibitor failed (or if it was unsuitable), the treatment pathway was less clear. The Committee heard that patients may then switch to an alternative TNF-alpha inhibitor, enter a clinical trial (if available), or try less proven options (such as antibiotics), and that surgery would be considered in these circumstances only if no other options remained. The Committee noted that, according to its marketing authorisation, vedolizumab may be used after conventional non-biological therapy or TNF-alpha inhibitors have failed. It heard from the clinical experts that, in clinical practice, vedolizumab would mainly be used after TNF-alpha inhibitors have failed because there is extensive experience with using TNF-alpha inhibitors. It noted that the responses to consultation stressed that there is an extremely high unmet clinical need in people who have exhausted all of the current proven medical treatment options. The clinical experts confirmed that vedolizumab would be most useful in clinical practice for these patients, particularly those who had experienced treatment failure with 2 TNF-alpha inhibitors. The Committee concluded that the need for an additional treatment for Crohn's disease was greatest in people whose treatment options were limited, such as those whose disease had either failed to respond to, or lost response to TNF-alpha inhibitors, or for whom they were unsuitable.

Clinical effectiveness

4.3 The Committee discussed the generalisability of the GEMINI II and III trial populations. The Committee noted that, although a patient's Crohn's Disease Activity Index (CDAI) score could potentially exceed 600, the maximum CDAI score permitted at trial entry was 450. It heard from the clinical experts that only a few patients with a CDAI score greater than 450 were seen in routine clinical practice and consequently considered that the spectrum of disease severity in patients in GEMINI II and III (CDAI score 220–

450) was broadly comparable to that seen in clinical practice. The Committee concluded that the clinical characteristics of the populations in GEMINI II and III were generalisable to the population likely to have vedolizumab in clinical practice in England.

4.4 The Committee discussed which patient groups in the GEMINI II and III trials most closely matched the population likely to have vedolizumab in clinical practice in England. It was aware that the ERG had suggested that the clinical efficacy of vedolizumab may be different in those who had received previous TNF-alpha inhibitor treatment compared with those who had not, making it difficult to interpret the results from the intention-to-treat mixed populations, which included both patients who had not previously had TNF-alpha inhibitor treatment and those in whom a TNF-alpha inhibitor had failed. The Committee noted that 58% of patients in GEMINI II and 76% of patients in GEMINI III had experienced failure of a TNF-alpha inhibitor. The Committee concluded that for the purposes of its decision-making, it would be appropriate to evaluate vedolizumab in 2 distinct populations: those who had not had a TNF-alpha inhibitor before and those in whom a TNF-alpha inhibitor had failed.

4.5 The Committee discussed the induction regimens used in GEMINI II and III, including the timing of response assessment. It heard from the clinical experts that clinical trials used stringent definitions of response and remission, including the timing of assessment, whereas this was more flexible in clinical practice. The Committee noted that in the marketing authorisation for vedolizumab, the recommended induction dosing regimen is 300 mg at weeks 0, 2 and 6, corresponding with the induction regimen in GEMINI III, whereas patients in GEMINI II had vedolizumab at weeks 0 and 2. The Committee also noted that,

although response was assessed at 6 weeks in GEMINI II and III, the marketing authorisation allowed for an additional dose to be given at week 10 (with the first dose of maintenance treatment administered at week 14). The Committee further noted that the company's new evidence submitted in response to the appraisal consultation document stated that, in patients in whom a TNF-alpha inhibitor had failed, a numerically higher percentage of patients experienced remission at 10 weeks than at 6 weeks. The Committee heard from the clinical experts that they considered 6 weeks to be too early to discontinue treatment if a response had not been observed, and that induction response would generally be assessed later than this in clinical practice. It further heard that assessing remission rates at 6 weeks, may therefore not have reflected vedolizumab's true clinical efficacy, and that in clinical practice a patient was likely to have 4 doses before a decision was made to discontinue treatment because of a lack of response. The Committee concluded that assessing response at 6 weeks, as in GEMINI II and III, would not detect all patients whose disease would respond to induction treatment, and that using data from later time points in the trials could potentially increase the efficacy estimates for vedolizumab.

4.6 The Committee considered the clinical effectiveness of vedolizumab compared with placebo during induction treatment in GEMINI II and III. It noted that vedolizumab was more effective in inducing remission at 6 weeks than placebo in the intention-to-treat mixed population in GEMINI II and III. It further noted that the remission rate in GEMINI II was numerically higher with vedolizumab than placebo in both the population who had not had a TNF-alpha inhibitor before and the population in whom a TNF-alpha inhibitor had failed. The Committee was aware that in GEMINI III the primary outcome was remission at 6 weeks in the population in whom a TNF-alpha inhibitor had failed (76% of the

trial population). It noted that although the primary outcome in GEMINI III had not been met, an exploratory secondary analysis at 10 weeks did show a statistically significant benefit with vedolizumab in this group. The Committee concluded that for induction, vedolizumab improved clinical remission rates compared with placebo in the whole population, people who had never had TNF-alpha inhibitors and people in whom TNF-alpha inhibitor treatment had failed.

4.7 The Committee considered the clinical effectiveness of vedolizumab compared with placebo for maintenance treatment. It noted that the evidence was based only on the results of GEMINI II because GEMINI III did not include a maintenance phase. The Committee heard from the clinical experts that, in Crohn's disease, long-term maintenance of remission was the primary goal of treatment. The Committee noted, however, that the duration of GEMINI II was 52 weeks, that only limited data up to 104 weeks were available from GEMINI LTS. At 52 weeks, vedolizumab showed higher remission rates than placebo in the intention-to-treat population, in the population who had not had a TNF-alpha inhibitor before and in the population in whom a TNF-alpha inhibitor had failed. The Committee noted that the absolute remission rate was lower with both vedolizumab and placebo in the population in whom a TNF-alpha inhibitor had failed than in the population who had not had a TNF-alpha inhibitor before (see section 3.14). However, it observed that despite this lower absolute rate, there was a similar relative treatment effect and the remission rate was approximately twice as high with vedolizumab than with placebo in both of these populations. The Committee heard from the clinical experts that this reflected the fact that these patients had more established and difficult to treat disease, and that even a reduced absolute treatment effect would be perceived as highly beneficial in these patients because of the lack of other treatment options. Although it

was uncertain about longer-term effects, the Committee concluded that, for maintaining remission up to 52 weeks, vedolizumab was significantly better than placebo in the whole population, in the population who had not had a TNF-alpha inhibitor before and in the population in whom a TNF-alpha inhibitor had failed.

4.8 The Committee discussed the impact of vedolizumab on health-related quality of life, and was encouraged that the company had included self-reported quality of life in its clinical trials. It focused on the EQ-5D results because these were used to generate utility values for the company's economic model. It noted that the company's submission stated that a decrease in EQ-5D score of at least 0.3 points was considered a clinically meaningful improvement, and that all results had been presented as decreases in score compared with baseline values. However, the Committee was aware that an improvement in quality of life would be reflected by an increase in EQ-5D score. The company was unable to explain why it considered that vedolizumab improved quality of life, as assessed by EQ-5D, yet the scores decreased. The Committee concluded that, although the results using other assessment methods in GEMINI II and III suggested that vedolizumab could improve quality of life, it was unable to conclude what its effect was using EQ-5D scores.

4.9 The Committee discussed the safety of vedolizumab. It heard from the clinical experts that vedolizumab was thought to have a more favourable adverse-event profile than other biological treatments (such as TNF-alpha inhibitors) because of its gut-specific effect and fewer systemic side effects. The Committee understood that this could be a consequence of its selective mechanism of action, because the α4β7 integrin is expressed only on gut-selective T-helper lymphocytes. The selective effect might be a particular advantage in some people for whom a TNF-alpha inhibitor is

contraindicated. The Committee was aware that progressive multifocal leukoencephalopathy, a fatal condition affecting the brain has been seen with natalizumab, an antibody that inhibits α4-integrin. It was aware that because vedolizumab also inhibits a α4-integrin, the incidence of progressive multifocal leukoencephalopathy in people who have treatment with vedolizumab is being closely monitored, but noted that there have been no reports to date. It concluded that fewer systemic effects may be an advantage of this therapy over existing biological treatments for some patients, particularly if TNF-alpha inhibitor treatment was contraindicated.

4.10 The Committee considered the validity and usefulness of the company's network meta-analyses that compared vedolizumab with adalimumab and infliximab. It agreed with the ERG and the company that it was more meaningful to evaluate the clinical effectiveness of vedolizumab according to TNF-alpha inhibitor status rather than the overall population, because the results for this population were difficult to interpret and generalise (see sections 3.27, 3.36 and 3.37). It also agreed with the company and the ERG that the network meta-analyses of the results for the population in whom a TNF-alpha inhibitor had failed had a number of serious flaws. The Committee accepted that the trial populations in the induction phase in the vedolizumab and adalimumab studies were not comparable, and that it had not been possible to present an analysis for the maintenance phase. The Committee concluded that the network meta-analyses for the population in whom a TNF-alpha inhibitor had failed would not inform its decision-making, but that it should further consider the network meta-analysis in the population who had not had a TNF-alpha inhibitor before.

4.11 The Committee considered the validity and usefulness of the company's network meta-analyses that were provided at the

clarification stage for the population who had not had a TNF-alpha inhibitor before. It considered the population in the network meta-analyses to be broadly generalisable to the population presenting after conventional non-biological therapy has failed in clinical practice in England. However, it had several concerns:

- It heard from the clinical experts that placebo response rates varied considerably in Crohn's disease trials and that a meta-analysis comparing all treatments against a common placebo response had limitations.
- It noted that the company's submission included insufficient detail about how the network meta-analyses had been conducted and was unclear how the results presented in the clinical section related to those used in the company's economic model.
- The Committee noted the ERG's assertion that a random-effects model should be used instead of a fixed-effects model, and considered that this would be more appropriate.
- It had several specific concerns about the induction analysis:
  - The Committee was aware of the ERG's concerns about trials that had been excluded, particularly Targan et al, and was not fully satisfied by the justification given by the company.
  - The Committee noted that the analyses included the dose of adalimumab recommended in the summary of product characteristics but heard from the clinical experts that a higher loading dose is more often used in clinical practice and it was unclear how this would affect the estimates of clinical efficacy of adalimumab.
  - The Committee was aware that the company had used the 6-week time points to assess induction response but noted the clinical experts' view that this was too early to evaluate the response for at least some of the treatments.

- The Committee also had concerns about the network meta-analyses for the maintenance phase:
  - No primary analyses had been presented that included adalimumab (although it was aware that the company's model included data that estimated the relative treatment effect of vedolizumab and adalimumab).

The Committee concluded that the combination of these factors meant there was considerable uncertainty in the company's network meta-analyses results for the population who had not had a TNF-alpha inhibitor before.

Cost effectiveness

4.12 The Committee took into account the ERG's and the company's view that it was more meaningful to evaluate vedolizumab according to TNF-alpha inhibitor status rather than the mixed population, because the results of the mixed analysis were difficult to interpret and generalise to the NHS (see sections 4.3 and 4.10).

4.13 The Committee considered the structure of the company's economic model. It noted the ERG's concerns about several structural assumptions, including that the relapsing and remitting nature of the disease had not been captured, the simplistic approach to modelling surgery, the assumption that all patients whose condition did not respond to treatment had moderate to severe disease for the full duration of the model and that there would be no difference in outcomes between patients whose disease responded and patients whose disease did not respond in the moderate to severe health state. The Committee heard from the clinical experts that, for patients in whom multiple lines of therapy failed, the assumption of long-term continuation in the moderate to severe state as in the company's model was not unreasonable. The Committee concluded that the number of concerns raised by the

ERG meant that it was uncertain if the model was structurally sound but that, overall, it was acceptable to inform its decision-making.

4.14 The Committee discussed the dosing assumptions used during induction in the company's economic models. It noted that in the first 2 models the same treatment duration (6 weeks) was assumed for all therapies for the induction phase and that these were aligned with the clinical trials, rather than the marketing authorisations or clinical practice. The Committee appreciated the difficulty in aligning the model with clinical practice when this did not necessarily correspond with the trial data, but considered the assumption of 6 weeks for all therapies was a weakness because it did not give an accurate estimate of costs and clinical outcomes in clinical practice. However, the Committee was satisfied with the dosing assumptions used in the company's third model, and concluded that evaluating response at 10 weeks was appropriate.

4.15 The Committee discussed the assumptions related to treatment continuation in the company's economic models. It addressed the assumption that biological treatment would stop in all patients at 1 year. The Committee heard from the clinical experts that they would try to withdraw biological therapy after 1 year where remission had been achieved, but that treatment would be continued if there was a high risk of relapse or surgery, and that treatment would be likely to continue if the patient had not experienced remission but was gaining a benefit from treatment (that is, an improvement in symptoms). In general, it heard that they would not wish people to continue on a treatment if it was not needed. The Committee concluded that the company's approach to discontinuing biological therapy after 1 year of maintenance treatment was not unreasonable, but that treatment duration in clinical practice could be longer.

4.16 The Committee considered how the company had modelled long-term clinical effectiveness of treatment. It recalled that the NICE reference case specifies that the time horizon should be long enough to capture all associated costs and benefits. Because Crohn's disease is a chronic condition, the Committee took the view that a lifetime horizon (as in the third model) was more appropriate than a 10-year time horizon (as in the first 2 models). However, it had concerns about the relationship between the constant transition probabilities and the time horizon adopted. The Committee considered that the long-term benefit was difficult to predict, particularly when treatment was stopped at 1 year, and if overestimated, this would be amplified with a longer time horizon. The Committee concluded that there were uncertainties in the long-term modelling of clinical effectiveness of vedolizumab and its comparators, but that the company's approach in the third model (which adopted a lifetime horizon as recommended in the NICE reference case) was reasonable given the evidence available.

4.17 The Committee discussed the modelling of long-term adverse effects of conventional non-biological treatment and surgery. It was aware that repeated use of high-dose oral corticosteroids (which may be the mainstay of treatment when other medical options, including TNF-alpha inhibitors, have failed) was associated with a range of long-term adverse effects such as diabetes and osteoporosis. The Committee considered that these could have a substantial impact on health-related quality of life and could be associated with significant costs (for example, treating hip fractures and complications of diabetes). It considered that these would potentially have a greater impact when adopting a lifetime time horizon rather than 10 years. It noted that disutilities and costs for treatment of long-term adverse effects of corticosteroids had not been included in the company models. The Committee also noted consultation comments about costs of treating any unwanted

consequences of surgery, such as infertility. The Committee concluded that the company's model did not capture all the costs and disutilities associated with existing treatments for Crohn's disease, and that incorporating them would be likely to reduce the ICER in favour of vedolizumab.

4.18 The Committee discussed how surgery had been modelled. It was concerned that, in the first 2 models, around one-third of patients who had surgery would remain in the surgical health state in the next cycle, and noted that the associated health state costs were considerable (around £10,000 per cycle). The Committee considered that this was an unreasonably high proportion, and the modification made in the company's third model, which reduced the repeat surgery rate to 1%, was more appropriate. The Committee concluded that it was likely that the company's first 2 models overestimated the proportion of patients having repeated surgery, which would have the effect of overestimating the total costs for these patients, but that the proportion and costs used in the third model were acceptable.

4.19 The Committee discussed the clinical parameters used in the company's economic models. For the population in whom a TNF-alpha inhibitor had failed, it accepted that the network meta-analyses did not permit a robust comparison with the other biological therapies (see section 4.10). For the population who had not had a TNF-alpha inhibitor before, it noted that the company's submission was unclear about which results from the network meta-analyses had informed the model. The Committee was aware that the ERG considered that the results of the company's network meta-analysis could underestimate the uncertainty in treatment effects, because fixed-effects models had been used despite evidence of heterogeneity among the clinical trials in the network. The Committee noted the ERG's concerns about the lack of

information about how the GEMINI II and III results had been pooled, and the estimated rates of discontinuation because of adverse events. It noted further concerns that the company had used data from the single-arm ACCENT-I trial but excluded the placebo-controlled Targan trial when estimating the relative treatment effect of vedolizumab compared with infliximab during induction, and that the company had not discussed the limitations associated with ACCENT-I. It recalled that the ERG was not satisfied with the company's approach to modelling maintenance data for vedolizumab compared with adalimumab because CLASSIC II had been excluded (see section 3.38). Lastly, the Committee agreed that, in the absence of data suggesting otherwise, it should be assumed that the mortality rate is the same for all treatments and that the same risk should be applied to all Crohn's disease health states (as in the third model). The Committee concluded that, for the population in whom a TNF-alpha inhibitor had failed, it was appropriate to derive the clinical parameters from the GEMINI II and III clinical trial results for the comparison of vedolizumab with conventional non-biological therapy. The Committee further concluded that the clinical parameters derived from the network meta-analyses for the population who had not had a TNF-alpha inhibitor before were subject to considerable uncertainty.

4.20 The Committee discussed how the health-state costs had been modelled by the company. It noted that the costs had been based on those reported in Bodger et al. (2009) in the first 2 models, but that they had been updated in the third model so that they were derived from a survey of 8 clinical experts based in England. It noted that amending the health-state costs had significantly reduced the ICER for vedolizumab compared with conventional non-biological therapy in patients in whom aTNF-alpha treatment had failed (see section 3.78). The Committee heard from the

company that many of the costs described by Bodger et al. were out of date because of advances in diagnostic tests and monitoring in Crohn's disease. The clinical experts at the Committee meeting agreed with the company and stated that the heath-state costs derived from the survey seemed reasonable because managing Crohn's disease in patients in whom TNF-alpha inhibitors have failed uses very significant NHS resources. The Committee concluded that the health-state costs used in the third model were more likely to reflect current NHS practice in England.

4.21 The Committee considered the cost-effectiveness results for vedolizumab compared with conventional non-biological treatment and TNF-alpha inhibitors in people with moderately to severely active Crohn's disease. It noted that the ERG did not consider the company's ICERs to be robust because of the model's structural issues, and that consequently the ERG had not conducted exploratory analyses. The Committee noted the ERG's concerns but, in the absence of alternative estimates, decided to consider the company's analyses in further deta.

4.22 The Committee discussed the cost-effectiveness results for vedolizumab compared with conventional non-biological treatment and TNF-alpha inhibitors in the population who had not had a TNF-alpha inhibitor before. It noted that vedolizumab was subject to extended dominance and that it had been excluded from the ERG's fully incremental analysis (see section 3.70). It noted that the probability of vedolizumab being the most cost-effective treatment option at £20,000 per QALY gained was less than 1%. The Committee concluded that vedolizumab was not cost effective compared with TNF-alpha inhibitors for treating moderately to severely active Crohn's disease in the population who had not had a TNF-alpha inhibitor before.

4.23 The Committee considered the cost-effectiveness results for vedolizumab compared with conventional non-biological treatment and TNF-alpha inhibitors in the population in whom a TNF-alpha inhibitor had failed. It acknowledged the opinion of the clinical experts that these were the people for whom access to a new agent would be of most value because of the very limited treatment options available to them (see section 4.2). It considered therefore that the most appropriate comparator in this population was conventional non-biological therapy. It noted that the ICER for vedolizumab compared with conventional non-biological therapy generated using the company's third model (that is, the model that incorporated many of the Committee's preferences) was £21,600 per QALY gained. The Committee was aware that in clinical practice, vedolizumab treatment could be longer than the 1-year duration assumed in the company's base case (see section 4.15) and noted that extending vedolizumab's treatment duration to 2 or 3 years increased the ICER to £24,700 and £26,200 per QALY gained respectively. The Committee considered that the ICERs for vedolizumab compared with conventional non-biological treatment were within the range that would normally be considered a cost-effective use of NHS resources (that is, £20,000–30,000 per QALY gained) but it was concerned about the uncertainty in the modelling of long-term treatment effects of vedolizumab (see section 4.16) and some of the structural assumptions (see section 4.13). However, it recalled that disutilities and costs for treatment of long-term adverse effects of corticosteroids and surgical complications had not been included in the company models (see section 4.17), which would reduce the ICER, and noted the substantial unmet need in this population of patients. It concluded, on balance, that vedolizumab could be considered a cost-effective use of NHS resources and should be recommended for people in whom TNF-alpha inhibitor treatment has failed. It also concluded that

when treatment duration is greater than 1 year, it would be important to identify the people who would continue to derive ongoing clinical benefit. Therefore it concluded that, in the absence of treatment failure, people receiving vedolizumab should be reassessed at 12 months and then at least annually to see if continued treatment is justified.

4.24 The Committee gave further consideration to whether there was a subgroup of patients who cannot take TNF-alpha inhibitors because they are contraindicated or not tolerated, and in whom vedolizumab would provide the only medical alternative to conventional non-biological therapy. The Committee appreciated the high unmet clinical need in this group (see section 4.2) for whom there would otherwise be no biological treatment options, and concluded that when TNF-alpha inhibitors were contraindicated or not tolerated, it was reasonable for vedolizumab to be prescribed.

4.25 The Committee contemplated whether vedolizumab could be considered an innovative technology. It noted that vedolizumab has a different mechanism of action to other drug treatments for Crohn's disease, and the clinical experts' opinion that the systemic side effects of treatment were lower with vedolizumab compared with other treatments (including TNF-alpha inhibitors). The Committee found it plausible that vedolizumab's gut-selective mechanism of action could result in a more favourable side-effect profile than other treatments (including other types of biological treatment) which have more systemic effects. It recalled that there were no head-to-head trials comparing vedolizumab and TNF-alpha inhibitors and that the company had not presented adverse-event data in the network meta-analyses in its submission. The Committee concluded that vedolizumab has a different mechanism of action to other drug treatments for Crohn's disease

and, in this regard, was innovative. As discussed in section 4.17, it also considered that the costs and disutilities associated with infertility and the long-term adverse effects of oral corticosteroid use had not been captured in the company's cost-effectiveness calculations.

4.26 The Committee discussed the potential equality issues raised during consultation. It was aware that surgery can have an impact on fertility and that surgery has particular issues for people following certain religious practices. The Committee heard from the clinical and patients experts that, in general, people with Crohn's disease preferred to manage their condition medically rather than surgically, and therefore considered that this did not represent an equality issue. The Committee considered that there were no equality issues associated with benefit according to age because vedolizumab brought treatment benefits to adults of all ages. The Committee concluded that there was no need to alter its recommendations because of any equality issues.

Summary of Appraisal Committee's key conclusions

TAXXX	**Appraisal title: Vedolizumab for treating moderately to severely active Crohn's disease after prior therapy**	**Section**
Key conclusion		
Vedolizumab is recommended as an option for treating moderately to severely active Crohn's disease only if: • a tumour necrosis factor-alpha inhibitor has failed (that is the disease has responded inadequately or has lost response to treatment) or • a tumour necrosis factor–alpha inhibitor cannot be tolerated or is		1.1

contra-indicated. Vedolizumab is recommended only if the company provides it with the discount agreed in the patient access scheme. The Committee considered that the incremental cost-effectiveness ratios (ICERs) for vedolizumab compared with conventional non-biological treatment in people in whom tumour necrosis factor (TNF) alpha treatment has failed were within the range that would normally be considered a cost-effective use of NHS resources (that is, £20,000–30,000 per quality adjusted life-year [QALY] gained). It concluded that vedolizumab should be recommended for people in whom TNF-alpha inhibitor treatment has failed; however, if treatment lasts for longer than 1 year, people receiving it should be reassessed at 12 months and then at least annually to see if continued treatment is justified. The Committee appreciated the high unmet clinical need in people for whom TNF-alpha inhibitors were contraindicated or not tolerated and concluded that it was reasonable for vedolizumab to be prescribed for this group. The Committee concluded that vedolizumab was not cost effective compared with TNF-alpha inhibitors for treating moderately to severely active Crohn's disease in the population who had not had a TNF-alpha inhibitor before.		4.23 4.24 4.22
Current practice		
Clinical need of patients, including the availability of alternative treatments	The Committee concluded that a further drug treatment that improves symptoms or brings the disease into remission would be highly valued by patients.	4.1, 4.2

The technology		
Proposed benefits of the technology How innovative is the technology in its potential to make a significant and substantial impact on health-related benefits?	Vedolizumab (Entyvio, Takeda UK) is a humanised IgG1 monoclonal antibody derived from a newly engineered cell line. It is targeted against α4β7 integrin, which is expressed on certain white blood cells. α4β7 integrin is responsible for recruiting these cells to inflamed bowel tissue. The Committee concluded that vedolizumab has a different mechanism of action to other drug treatments for Crohn's disease and, in this regard, was innovative.	2.1 4.25
What is the position of the treatment in the pathway of care for the condition?	The Committee noted that, according to its marketing authorisation, vedolizumab may be used after conventional non-biological therapy or TNF-alpha inhibitors have failed. It concluded that the need for an additional treatment was greatest in people whose treatment options were limited, such as those whose disease had either failed to respond to, or lost response to TNF-alpha inhibitors, or for whom they were unsuitable.	4.2
Adverse reactions	The Committee concluded that fewer systemic effects may be an advantage of vedolizumab over existing biological treatments for some patients, particularly if TNF-alpha inhibitor treatment was contraindicated.	4.9
Evidence for clinical effectiveness		

Availability, nature and quality of evidence	The Committee considered the clinical effectiveness of vedolizumab compared with placebo during induction and maintenance treatment in GEMINI II and III. It noted that the duration of GEMINI II was 52 weeks and that only limited data up to 104 weeks were available from GEMINI LTS. The Committee concluded that the network meta-analyses for the population in whom a TNF-alpha inhibitor had failed would not inform its decision-making. The Committee concluded that there was considerable uncertainty in the company's network meta-analyses results for the population who had not had a TNF-alpha inhibitor before.	4.6, 4.7 4.10 4.11
Relevance to general clinical practice in the NHS	The Committee concluded that the clinical characteristics of the populations in GEMINI II and III were generalisable to the population likely to have vedolizumab in clinical practice in England.	4.3
Uncertainties generated by the evidence	The Committee concluded that assessing response at 6 weeks, as in GEMINI II and III, would not detect all patients whose disease would respond to induction treatment.	4.5

Are there any clinically relevant subgroups for which there is evidence of differential effectiveness?	The Committee concluded that for the purposes of its decision-making, it would be appropriate to evaluate vedolizumab in 2 distinct populations: those who had not had a TNF-alpha inhibitor before and those in whom a TNF-alpha inhibitor had failed.	4.4
Estimate of the size of the clinical effectiveness including strength of supporting evidence	The Committee concluded that for induction, vedolizumab improved clinical remission rates compared with placebo in the whole population, and also in populations of people who had never had TNF-alpha inhibitors and in whom TNF-alpha inhibitor treatment had failed. Although it was uncertain about longer-term effects, the Committee concluded that, for maintaining remission up to 52 weeks, vedolizumab was significantly better than placebo in the whole population, in the population who had not had a TNF-alpha inhibitor before and in the population in whom a TNF-alpha inhibitor had failed.	4.6 4.7
Evidence for cost effectiveness		
Availability and nature of evidence	The Committee concluded that the number of concerns raised by the Evidence Review Group (ERG) meant that it was uncertain if the model was structurally sound but that, overall, it was acceptable to inform its decision-making.	4.13
Uncertainties around and plausibility of	The Committee considered that the dosing assumptions used in the company's first	4.14

assumptions and inputs in the economic model	2 models did not necessarily give an accurate estimate of costs and clinical outcomes in clinical practice. However, the Committee was satisfied with the dosing assumptions used in the company's third model, and concluded that evaluating response at 10 weeks was appropriate.	
	The Committee concluded that the company's approach to discontinuing biological therapy after 1 year of maintenance treatment was not unreasonable, but that treatment duration in clinical practice could be longer.	4.15
	The Committee concluded that there were uncertainties in the long-term modelling of clinical effectiveness of vedolizumab and its comparators, but that the company's approach in the third model (which adopted a lifetime horizon as recommended in the NICE reference case) was reasonable given the evidence available.	4.16
	The Committee concluded that the company's model did not capture all the costs and disutilities associated with existing treatments for Crohn's disease, and that incorporating them would be likely to reduce the ICER in favour of vedolizumab.	4.17
	The Committee concluded that it was likely that the company's first 2 models overestimated the proportion of patients having repeated surgery, which would have	4.18

	the effect of overestimating the total costs for these patients, but that the proportion and costs used in the third model were acceptable.	
Incorporation of health-related quality-of-life benefits and utility values Have any potential significant and substantial health-related benefits been identified that were not included in the economic model, and how have they been considered?	The Committee was encouraged that the company had included self-reported quality of life in its clinical trials and focused on the EQ-5D results because these were used to generate utility values for the company's economic model. The Committee was unable to form a conclusion on vedolizumab's effect on quality of life using EQ-5D scores because of uncertainty in how these had been reported. The Committee considered that the costs and disutilities associated with infertility and the long-term adverse effects of oral corticosteroid use had not been captured in the company's cost-effectiveness calculations.	4.8 4.25
Are there specific groups of people for whom the technology is particularly cost effective?	Not applicable.	
What are the key drivers of cost effectiveness?	The key drivers of the ICER included many of the transition probabilities, health state costs, utility values and time horizon.	3.57, 3.58, 3.73

Most likely cost-effectiveness estimate (given as an ICER)	Because the ERG had not conducted exploratory analyses, the Committee was only able to consider the company's ICERs. It considered the cost-effectiveness results for the different populations in turn: • In the population who had not had a TNF-alpha inhibitor before, vedolizumab was subject to extended dominance in the ERG's fully incremental analysis. • In the population in whom a TNF-alpha inhibitor had failed, the company's ICER for vedolizumab compared with conventional non-biological therapy was £21,600 per QALY gained.	4.21 4.22 4.23
Additional factors taken into account		
Patient access schemes (PPRS)	The company has agreed a patient access scheme with the Department of Health that would provide a simple discount to the list price of vedolizumab. The level of the discount is commercial in confidence.	2.4
End-of-life considerations	Not applicable.	
Equalities considerations and social value judgements	The Committee discussed the potential equality issues raised during consultation on the appraisal consultation document and concluded that they did not require its recommendations to be altered.	4.26

5 Implementation

5.1 Section 7(6) of the National Institute for Health and Care Excellence (Constitution and Functions) and the Health and Social Care Information Centre (Functions) Regulations 2013 requires clinical commissioning groups, NHS England and, with respect to their public health functions, local authorities to comply with the recommendations in this appraisal within 3 months of its date of publication.

5.2 The Welsh Assembly Minister for Health and Social Services has issued directions to the NHS in Wales on implementing NICE technology appraisal guidance. When a NICE technology appraisal recommends the use of a drug or treatment, or other technology, the NHS in Wales must usually provide funding and resources for it within 3 months of the guidance being published.

5.3 When NICE recommends a treatment 'as an option', the NHS must make sure it is available within the period set out in the paragraphs above. This means that, if a patient has Crohn's disease and the doctor responsible for their care thinks that vedolizumab is the right treatment, it should be available for use, in line with NICE's recommendations.

5.4 The Department of Health and Takeda have agreed that vedolizumab will be available to the NHS with a patient access scheme which makes it available with a discount. The size of the discount is commercial in confidence. It is the responsibility of the company to communicate details of the discount to the relevant NHS organisations. Any enquiries from NHS organisations about the patient access scheme should be directed to [NICE to add details at time of publication]

5.5 NICE has developed tools [link to www.nice.org.uk/guidance/TAXXX] to help organisations put this guidance into practice (listed below). [NICE to amend list as needed at time of publication]

- Slides highlighting key messages for local discussion.
- Costing template and report to estimate the national and local savings and costs associated with implementation.
- Implementation advice on how to put the guidance into practice and national initiatives that support this locally.
- A costing statement explaining the resource impact of this guidance.
- Audit support for monitoring local practice.

6 Related NICE guidance

Details are correct at the time the final appraisal determination goes out for appeal and will be removed when the final guidance is published. Further information is available on the NICE website.

Published

- Crohn's disease: management in adults, children and young people. NICE guideline CG152 (2012).
- Infliximab (review) and adalimumab for the treatment of Crohn's disease (including a review of technology appraisal guidance 40). NICE technology appraisal guidance 187 (2010).
- Extracorporeal photopheresis for Crohn's disease. NICE interventional procedure guidance 288 (2009).

7 Review of guidance

7.1 The guidance on this technology will be considered for review 3 years after publication of the guidance. The Guidance Executive will decide whether the technology should be reviewed based on

information gathered by NICE, and in consultation with consultees and commentators.

Jane Adam

Chair, Appraisal Committee

July 2015

8 Appraisal Committee members and NICE project team

Appraisal Committee members

The Appraisal Committees are standing advisory committees of NICE. Members are appointed for a 3-year term. A list of the Committee members who took part in the discussions for this appraisal appears below. There are 4 Appraisal Committees, each with a chair and vice chair. Each Appraisal Committee meets once a month, except in December when there are no meetings. Each Committee considers its own list of technologies, and ongoing topics are not moved between Committees.

Committee members are asked to declare any interests in the technology to be appraised. If it is considered there is a conflict of interest, the member is excluded from participating further in that appraisal.

The minutes of each Appraisal Committee meeting, which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

Dr Jane Adam (Chair)
Department of Diagnostic Radiology, St George's Hospital, London

Professor Iain Squire (Vice-Chair)
Consultant Physician, University Hospitals of Leicester

Dr Graham Ash
Consultant in General Adult Psychiatry, Lancashire Care NHS Foundation Trust

Dr Simon Bond
Senior Statistician, Cambridge Clinical Trials Unit

Dr Jeremy Braybrooke

Consultant Medical Oncologist, University Hospitals Bristol NHS Foundation Trust

Dr Gerardine Bryant

GP, Swadlincote, Derbyshire

Dr Andrew England

Senior Lecturer, Directorate of Radiography, University of Salford

Dr Peter Heywood

Consultant Neurologist, Frenchay Hospital, Bristol

Dr Louise Longworth

Reader in Health Economics, HERG, Brunel University

Dr Anne McCune

Consultant Hepatologist, University Hospitals Bristol NHS Foundation Trust

Professor John McMurray

Professor of Medical Cardiology, University of Glasgow

Dr Alec Miners

Senior Lecturer in Health Economics, London School of Hygiene and Tropical Medicine

Dr Mohit Misra

GP, Queen Elizabeth Hospital, London

Sarah Parry

CNS Paediatric Pain Management, Bristol Royal Hospital for Children

Pamela Rees

Lay member

Dr Ann Richardson

Lay member

Mr Stephen Sharp

Senior Statistician, University of Cambridge MRC Epidemiology Unit

Dr Brian Shine

Consultant Chemical Pathologist, John Radcliffe Hospital, Oxford

Dr Peter Sims

GP, Devon

Dr Eldon Spackman

Research Fellow, Centre for Health Economics, University of York

David Thomson

Lay member

Dr Nerys Woolacott

Senior Research Fellow, Centre for Health Economics, University of York

Professor Olivia Wu

Professor of Health Technology Assessment, University of Glasgow

NICE project team

Each technology appraisal is assigned to a team consisting of 1 or more health technology analysts (who act as technical leads for the appraisal), a technical adviser and a project manager.

Linda Landells

Technical Lead

Zoe Charles

Technical Adviser

Bijal Joshi

Project Manager

9 Sources of evidence considered by the Committee

A. The Evidence Review Group (ERG) report for this appraisal was prepared by the School of Health and Related Research:

- Rafia R, Scope A, Harnan S et al. Vedolizumab for the treatment of adults with moderately to severely active Crohn's disease: A Single Technology Appraisal, October 2014

B. The following organisations accepted the invitation to participate in this appraisal as consultees and commentators. They were invited to comment on the draft scope, the ERG report and the appraisal consultation document (ACD). Organisations listed in I were also invited to make written submissions. Organisations listed in II and III had the opportunity to make written submissions. Organisations listed in I, II and III also have the opportunity to appeal against the final appraisal determination.

I. Company:

- Takeda UK (vedolizumab)

II. Professional/expert and patient/carer groups:

- British Society of Gastroenterology
- Crohn's and Colitis UK
- Royal College of Physicians
- United Kingdom Clinical Pharmacy (UKCPA)

III. Other consultees:

- Department of Health
- NHS England
- Welsh Government

- Department of Health, Social Services and Public Safety for Northern Ireland
- Healthcare Improvement Scotland

IV. Commentator organisations (did not provide written evidence and without the right of appeal):

- AbbVie (adalimumab)
- Department of Health, Social Services and Public Safety for Northern Ireland
- Healthcare Improvement Scotland
- Merck Sharp & Dohme (infliximab)
- Napp Pharmaceuticals (prednisolone)
- National Institute for Health Research Health Technology Assessment Programme
- School of Health and Related Research Sheffield (ScHARR)

C. The following individuals were selected from clinical expert and patient expert nominations from the consultees and commentators. They gave their expert personal view on vedolizumab by attending the initial Committee discussion and providing a written statement to the Committee. They were also invited to comment on the ACD.

- Dr Seamus Murphy, Consultant Gastroenterologist, nominated by organisation and representing British Society of Gastroenterology – clinical expert
- Dr Jeremy Sanderson, Consultant Gastroenterologist, nominated by organisation and representing British Society of Gastroenterology – clinical expert
- Paula Battersby nominated by organisation and representing Crohn's and Colitis UK – patient expert
- Andy Phillips nominated by organisation and representing Crohn's and Colitis UK – patient expert.

D. Representatives from the following company attended Committee meetings. They contributed only when asked by the Committee chair to clarify specific issues and comment on factual accuracy.

- Takeda UK