National Institute for Health and Care Excellence
Centre for Health Technology Evaluation

Pro-forma Response

ERG report

Vedolizumab for treating moderately to severely active ulcerative colitis [ID691]

Issue 1 Inclusion of outcomes bases on relapse rates

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 1 The comparators considered within the systematic review, network meta-analysis (NMA) and health economic analysis are not consistent. Outcomes data on relapse rates were not presented in the MS.	No change is proposed based on Takeda's justification.	Takeda acknowledges that outcomes based on relapse rates are not included. Relapse was not a primary endpoint in GEMINI I and was not reported for most studies. Therefore any analyses would have been subject to considerable statistical uncertainty.	This is not a factual error. No amendment is necessary.

Issue 2 Use of incorrect data in the NMA

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 4: "The fixed effects NMA suggested that in the induction phase for anti-TNF-α naïve patients, infliximab provided the largest treatment effect on clinical response, remission and mucosal healing compared with placebo, and vedolizumab has the lowest rate of discontinuations due to AEs compared with placebo." This is no longer correct and requires updating.	Takeda proposes the highlighted statement should be amended as follows: *In the induction treatment of anti-TNF–naïve patients with UC, the efficacy of vedolizumab (in terms of response, remission, and mucosal healing) was not significantly different to that of other biologics. A benefit was seen in terms of the rate of discontinuation due to AEs; the rate for vedolizumab was statistically significantly better than that for adalimumab 160/80 mg.*	By implementing this change to the ERG's text, the report will then accurately reflect the results of the NMA.	This is not a factual error. If additional analyses have been undertaken these should be presented to NICE. The ERG has not seen any corrections.

Issue 3 Dependence/correlation between response and remission

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 5 The results of the NMA for clinical response and remission should be interpreted with further caution because these were estimated without considering the dependence/correlation between response and remission. Use of these results in the economic model ignores this dependence and may generate inappropriate samples for probabilistic sensitivity analysis (PSA).	No change is proposed based on Takeda's justification.	Takeda acknowledges the view of the ERG, and believes that the data were handled in the most appropriate way. Therefore the results of the NMA can be interpreted with greater confidence than suggested. The ERG appears to assume that the data were reported on an ordinal scale (neither response nor remission, response but not remission, both response and remission). However, all studies reported response and remission separately. It was therefore not possible to analyse the data in this way. Hence, Takeda is unable to provide such analyses.	This is not a factual error. The ERG acknowledges that ordered categorical data were not available to the manufacturer for all trials in the NMA, however the point remains that treating these as binomial does not adequately reflect the nature of the data. No changes have been made.

Issue 4 Anti-TNF failure versus anti-TNF experienced/exposed

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
In the executive summary (page 6) and in section 4.3, there is misinterpretation/misunderstanding reflected in the text with respect to the anti-TNF exposed/experienced versus	Takeda proposed the following amendment to the text where required: *In GEMINI 1 failure to anti-TNFs was defined as inadequate response (i.e. primary non-responders to induction therapy*	Takeda recognises that there is a need amend the text as it gives the impression that 'anti-TNF failure' is the same as 'anti-TNF exposed/experienced'; and it does not reflect the difference in primary	This not a factual error, however text has been amended for clarity as suggested by the manufacturer.

anti-TNF failure. This is particularly relevant in the discussion of comparability between the patient populations in GEMINI 1 study and ULTRA 2 study, the latter being the only available comparator study which included anti-TNF exposed patients.	*with anti-TNF therapy), loss of response (i.e. secondary non-response/loss of response anti-TNF over time following initial response) or patients were intolerable to anti-TNFs. This is different to ULTRA 2 study where previous use of anti-TNF agents other than adalimumab was permitted if the patient had discontinued its use due to a loss of response or intolerance to the agent for longer than 8 weeks (i.e. this study does not appear to have included primary non-responders to anti-TNFs).*	non-response/failure and secondary non-response/failure and therefore the manufacturer has no reasonable justification for the approach adopted in the NMA (the latter issue will be addressed separately below). In ULTRA 2 study, previous use of anti-TNF agents other than adalimumab was permitted if the patient had discontinued its use due to a loss of response or intolerance to the agent for longer than 8 weeks (Sandborn 2012, Reference 38 in the ERG report). Therefore this study defined anti-TNF 'failure' as patients who were specifically secondary non-responders (i.e. lost response at least after 8 week induction phase) or intolerability to an anti-TNF. This is in contrast to GEMINI 1 where failure to anti-TNFs was defined as inadequate response (i.e. primary non-responders to induction therapy with anti-TNF therapy), loss of response (i.e. secondary non-response/loss of response anti-TNF over time following initial response) or patients were intolerable to anti-TNFs. In addition, the inclusion criteria required that patients demonstrated anti-TNF failure over the previous 5-year period.	"In GEMINI 1, failure to anti-TNFs was defined as inadequate response (i.e. primary non-responders to induction therapy with anti-TNF therapy), loss of response (i.e. secondary non-response/loss of response anti-TNF over time following initial response) or patients were intolerable to anti-TNFs. Whilst in ULTRA 2, previous use of anti-TNF agents other than adalimumab was permitted if the patient had discontinued its use due to a loss of response or intolerance to the agent for longer than 8 weeks (i.e. this study does not appear to have included primary non-responders to anti-TNFs)"

		The ERG note that anti-TNF-experienced population may be a more difficult to treat population than those who are anti-TNF naïve but there is no reference to the type of anti-TNF failure. Takeda acknowledges this view, and is of the opinion that primary non-response (i.e. inadequate response during induction phase) is more difficult to treat than secondary non-response/loss of response: at baseline in GEMINI 1 study there were 46% (PL) and 54% (VDZ) with inadequate responses to anti-TNFs, 41% (PL) and 39% (VDZ) who had loss of response to anti-TNFs and 13% (PL) and 7% (VDZ) with tolerability issues to anti-TNFs. Therefore GEMINI 1 is the first phase 3 study to include primary and secondary failure to anti-TNFs in patients with moderate to severely active UC. Unlike the anti-TNF failure population; anti-TNF experienced patients included those patients who may have had a partial response or relapse following anti-TNF therapy. PL=placebo VDZ = vedolizumab	

Issue 5 Network diagram errors

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 59 Whilst network diagrams were not reported within the MS, these were provided by the manufacturer in response to a request for clarification from the ERG (see Figures 2-5). The ERG believes that there are mistakes in the diagrams provided by the manufacturer.	The footnotes assigned to the network diagrams require correction.	When preparing the response to clarification questions, some footnotes to the network diagrams were copied incorrectly. This was purely an editorial issue and should be corrected.	This is not a factual error in the ERG report. Changes have not been made – network diagrams are reported as provided by the manufacturer.

Issue 6 Use of incorrect data in the NMA

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
On page 62–63 of the ERG report, errors in data extraction of the trials relevant for the NMA are highlighted.	Takeda proposes to update the relevant NMA and the CEA model, and to forward to NICE as soon as is possible.	Takeda acknowledges the need to correct factual errors in the submission and the supporting documents.	This is not a factual error made by the ERG, no changes have been made.

Issue 7 Incorrect numbers

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 62 The ULTRA trial data for the anti-TNF-α experienced/failure population maintenance phase clinical remission reported in Table 142 in the MS1 report were incorrect as the N for placebo and adalimumab should be 29 and 36, respectively.	Update Table 142 with the correct data.	The ULTRA data used to run the MTC were correct, whereas Table 142 contains a typographical error that needs to be corrected. Takeda notes that the ERG highlighted other apparently incorrect numbers elsewhere. On review of these numbers, Takeda believes that these numbers are due to the methods used to calculate the data, based on the number of responders at the end of induction, rather than the N value itself. Therefore, Takeda does not believe that further corrections are required.	This is not a factual error in the ERG report, no changes have been made. Data are reported as presented within the manufacturer's submission. Errors identified by the ERG are discussed.

Issue 8 Approach adopted by the manufacturer in the NMA with respect to prior anti-TNF therapy

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
In the executive summary and specifically on page 63 of the report, reference is made to the manufacturer undertaking and presenting separate NMAs of anti-TNF-α naïve and anti-TNF-α experienced/failure subgroups but that no rationale for conducting such analysis on subgroups separately is provided.	Takeda suggests the amendments be made as follows: *The manufacturer undertook separate NMAs of anti-TNF-α naïve and anti-TNF-α experienced/failure subgroups. The rationale for this approach are based on the difference between anti-TNF-experienced/exposed patients (ULTRA 2) compared to anti-TNF-failure patients (GEMINI 1). GEMINI 1 included primary non-responders to induction therapy with anti-TNF therapy, secondary non-response/loss of response anti-TNF over time following initial response, or patients who were intolerable to anti-TNFs. In ULTRA 2 study only secondary non-responders to anti-TNFs (other than adalimumab) and patients with intolerance to anti-TNF for longer than 8 weeks (i.e. this study does not appear to have included primary non-responders to anti-TNFs).[Sandborn 2012]* *As a result of these differences, separate analyses were required for the anti-TNF naïve subgroup and anti-TNF failure as the ITT populations between the studies were not comparable and when looking further at subgroups by prior anti-TNF exposure, there are further issues regarding type of anti-TNF failure*.	Takeda acknowledges the issue of comparability, and highlights that the anti-TNF failure population is more difficult to treat than the anti-TNF experienced population. Therefore, conclusions from these analyses should be made with caution. The complete rationale for the approach of separate NMAs for the anti-TNF is explained in detail in Takeda's submission (page 123-125) and in the clarification response. Further discussion on the challenges in performing NMA are presented in the report provided to NICE (Takeda Data on File, 2014. An Updated Systematic Literature Review and Network Meta-analysis in Ulcerative Colitis. Final Report).	This is not a factual error and the ERG's point appears to have been misinterpreted. Our suggestion was to examine whether there is an interaction between subgroup and treatment effect. This would require subgroup to be a covariate in the NMA.

Issue 9 Calibration on NMA data

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 102 The manufacturer's use of calibration methods discards their empirical trial data.	The following statement should be included: *"The calibration was performed on NMA data to be consistent among comparators."*	Takeda acknowledges that transition probabilities were fitted to modelled health state distributions using an NMA. Takeda believes that this is justified given the nature of the available data and that the use of mixed methods for this calibration step would also have been inappropriate. Further, the ERG acknowledges in its report that a calibration would not be possible with other biologic comparators due to lack of patient-level data (pages 102 and 103).	This is not a factual error. The ERG report already acknowledges the issue regarding consistency of approach. However, it would still have been preferable, in the view of the ERG, to use the trial data directly. This point is reflected in Section 5.2.12 of the NICE Methods Guide.

Issue 10 Use of repeat surgery

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Pages 103 /104 / 111. The ERG report states that it is implausible that patients have repeat surgery (pages 103 and 104). However there are data presented by Loftus (which are limited to 6 months follow up) that show that patients have repeat surgery. The report also cites the	The transition matrix should allow for transition from surgery to surgery. We recommend that the ERG base case model should allow repeat surgery.	The data from the Loftus and Arai papers demonstrate that patients with UC do have repeat surgeries and the model should account for these.	The ERG is not arguing that patients cannot have further surgery after the initial procedure. However, it is important to be clear about why patients may require additional surgery. This may be due to the need for subsequent stages of IPAA or may be required to resolve

study by Arai et al., 2005 (page 111 of the ERG report). That study found that amongst 296 patients that underwent proctocolectomy, with a mean follow up of 4.38 years, there were 35 subsequent surgical treatments for complications.			complications of the primary surgery. The former is already captured in the use of surgical cost estimates based on Buchanan et al (this paper assumes 40% of patients undergo proctocolectomy with ileostomy, 60% undergo subtotal proctocolectomy). In the manufacturer's model, the use of a 0.05 transition probability from all post-surgery states back to surgery means that patients who enter surgery undergo 4.32 surgeries over 10-years (and many more over a lifetime horizon). This is clearly very different from the manufacturer's suggestion that 12% (35/296) patients will undergo repeat surgery for complications. The ERG believes that transitions back to the surgery state should <u>not</u> be included in the model.

Issue 11 Early post-surgical complications

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 104 / Technical Appendix The ERG report states that the	The transition matrix should allow for early complications from surgery, with a transition	The data from the Loftus, Mahadevan and Arai papers	This is not a factual error. The ERG agrees that patients can

model over-estimates the probability of early complications following surgery (the transition from surgery to post-surgical complications). However, given the data presented in studies by Loftus et al., Mahadevan et al., Arai et al. it is reasonable to model the transition from surgery to post-surgical complications: the literature shows that patients do have early complications. The transition is applied during the lifetime of the model, but only for the transition from one state to another (i.e. it is only applied for the 8 weeks after surgery). It is not a transition probability from and to the same state and it does not matter when it is applied in the model, it is still valid. As implemented in the ERG base case model, early and late complications are apparently grouped together. Patients can only transition from post-surgery remission to post-surgery complications. However, they cannot transition out of post-surgery complications to any other health state.	from surgery to post-surgery complications. We recommend that the ERG base case model should be amended to allow for early and late complications rather than grouping them together and assuming that once a patient has complications they will have complications for life.	demonstrate that early complications of surgery occur and the model should allow for those complications.	have early complications. The manufacturer's model only differentiates between early and late complications in terms of transitions between states, not in terms of differences in cost or HRQoL.

Issue 12 Late post-surgical complications

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 105 / 111 / Technical Appendix The ERG report states that the model over-estimates the probability of late complications from surgery (represented in the model by the transition probability from post-surgical remission to post-surgical complications). The ERG base case model assumes that the transition from post-surgery remission to post-surgery complications is 0.05. However, the ERG report cites the study by Arai et al., 2005 which found that that 49%, 9.9%, 6.8%, 6.9% and 3.7% of patients had complications in the first to fifth year following proctocolectomy. As implemented in the ERG base case model, early and late complications are apparently grouped together. Patients can only transition from post-surgery remission to post-surgery complications. However, they cannot transition out of post-surgery complications to any other	We recommend that the ERG base case model transition matrix should allow for a higher probability of transition from post-surgical remission to post-surgical complications. We recommend that the ERG base case model should be amended to allow for early and late complications rather than grouping them together and assuming that once a patient has complications they will have complications for life.	The data from the Loftus and Arai papers suggest a high rate of complications in the first 6 to 12 months and a higher rate of complications than would be captured by a transition probability of 0.05. Separating early and late complications allows for a more accurate assessment of the effect of surgery on costs and utilities.	This is not a factual error. The values used by the manufacturer relate to short time intervals following surgery and are applied on a repeated basis. The ERG does not believe that this is credible.

health state.			

Issue 13 NICE reference case

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 110 The manufacturer's model has been implemented partly in line with NICE's Reference Case … and the economic analysis is generally in line with the final NICE scope. Three deviations from the NICE Reference Case warrant more detailed discussion: these relate to (i) relevant comparators excluded from analyses of mixed ITT and anti-TNF-α naïve populations; (ii) use of a 10-year time horizon, and (iii) failure to undertake a fully incremental analysis.	Takeda recognises the observations of the ERG regarding the reference case, and is pleased that there the model is largely in line. The first deviation is noted and Takeda believes it to be justified specifically in the case of vedolizumab.	a) Regarding the choice of comparators, Takeda is in line with the comparators identified by NICE in the scoping meeting. b) Specifically for the mixed ITT population, anti-TNFs were excluded because the anti-TNF-naïve and -experienced populations were considered to be managed differently, based on their disease activity and treatment journey up to that point, based on NICE guidelines. c) Specifically for the anti-TNF-failure model, only one study in addition to GEMINI I involved patients who could be classified as such, but that study only recruited patients with secondary failure, so could not be considered comparable to the failure population in GEMINI I. d) Consequently, there were no data to include in the model rather than there being active exclusion of anti-TNF drugs from the model.	This is not a factual error.