Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy (including a review of TA140 and TA262): Clinical effectiveness systematic review and economic model

Abbvie – Assessment Group response to fact-check

Page #	Text	Suggested text or action	AG response
33	Five hundred forty-two participants were randomised to two groups including placebo	**Five hundred and eighteen** patients entered the study, of which 258 were randomised to adalimumab 160/80/40 mg and 260 randomised to placebo.	The manufacturer is correct. This will be amended in the monograph
34	The two induction adalimumab groups (one licenced dose and the other unlicensed) were combined as one active treatment group for outcomes at 52 weeks.	The ULTRA 1 open label extension combined all 3 arms post 8-weeks to give outcome measures at week 52.	Not a factual inaccuracy. The text in the assessment report relates to Suzuki *et al* not ULTRA1.
34	ULTRA3 was the 52-week open label extension study to ULTRA1 and ULTRA2.	ULTRA 3 was the **156-**week open label extension study of ULTRA1 and ULTRA2	The manufacturer is correct. This will be amended in the monograph
34	Overall, 1228 patients who were responders to golimumab induction therapy in two previous golimumab induction therapy trials (including PURSUIT-SC) were randomised to licenced maintenance doses of 50mg golimumab, 100mg golimumab or placebo.	Out of the 1228 patients who enrolled in PURSUIT, 464 were randomised to receive placebo (n=156) Golimumab 50 mg (n=154) or Golimumab 100 mg (n=154). Seven hundred and sixty four patients were not randomised: 129 were placebo non-responders, 230 were placebo induction non responders and 405 were golimumab induction non-responders.	The manufacturer is correct. This will be amended in the monograph
39	ULTRA1 and ULTRA2 (currently ongoing)	ULTRA1 and 2 have completed. ULTRA3 which is the follow- on study is ongoing in a number of countries	The text implies that ULTRA3 is currently ongoing.
39	ACT 2 inclusion/exclusion criteria same as ACT 1	ACT2 patients need only have failed 5-ASA as a minimum. Active ulcerative colitis with a Mayo Score of 6-12 points and endoscopy subscore of	The manufacturer is correct. There are slight differences in the inclusion criteria. This will be amended in the monograph

		2-3 Concurrent treatment with at least oral corticosteroids, immunosuppressants OR 5-ASA But patients could enrol if they failed to tolerate or respond to corticosteroids, azathioprine, 6-MP OR 5-ASA	
64	A *post hoc* analysis of ULTRA2 data at week 52 demonstrated that mean days in clinical response (134.58 vs. 94.55, *p*<0.001) and mean days in clinical remission were also greater for adalimumab-treated patients (85.32 vs. 52.87, *p*<0.001).	A *post hoc* analysis of ULTRA2 data at week 52, **which included patients from the placebo arm who switched to adalimumab**, demonstrated that mean days in clinical response (134.58 vs. 94.55, *p*<0.001) and mean days in clinical remission were also greater for adalimumab-treated patients (85.32 vs. 52.87, *p*<0.001).	This is fine and will be amended in the monograph
64	Similarly, among patients who had received no prior anti-TNF-α treatment, greater proportions of patients in the adalimumab 160/80mg induction group reached clinical response (50.5% vs. 38.6%, p <0.005) and clinical remission (16.5% vs. 11.0%, p <0.05) at week 8 than placebo-treated patients.	Similarly, among patients who had received no prior anti-TNF-α treatment, greater proportions of patients in the adalimumab 160/80mg induction group reached clinical response (**59.3%** vs. 38.6%, p <0.005) and clinical remission (**21.3%** vs. 11.0%, p <0.05) at week 8 than placebo-treated patients.	This is a typo which will be amended in the monograph. The values used in the NMA are correct.
65	The open-label extension study ULTRA3 presented the proportions of patients who continued to receive adalimumab and were in clinical response (42.6%) and remission (25.6%) at week 52.	The open-label extension study ULTRA3 presented the 4-year efficacy and safety results of 588 patients from ULTRA1 AND ULTRA2 who were followed.[i]. Of the 588 patients who entered the ULTRA3 extension study, 52.2% (307/588) were in remission upon entry according to partial Mayo Scores. Partial Mayo Scores were calculated at each study visit and at week 156, 46.4% (273/588) of patients had achieved clinical remission	We will amend the text for the monograph.
84	Acute infusion reactions occurred in similar numbers of patients in both treatment arms	Acute infusion reactions occurred in similar numbers of patients in both treatment arms through	This is a typo which will be amended in the monograph.

	through week 54 of ACT1 (infliximab 5mg/kg 2/121, 9.9%; placebo 13/121, 10.7%).	week 54 of ACT1 (infliximab 5mg/kg **12**/121, 9.9%; placebo 13/121, 10.7%).	
212	It is noteworthy that the difference in utility for the post-surgery state and the active UC state in the selected utility values within the AbbVie model (0.61-0.55=0.06) is smaller than that observed within other EQ-5D UC valuation studies (e.g. Woehl et al[109] estimated this difference to be ~0.71– 0.41= 0.30).	The incorrect reference number is assigned to Woehl in the text (109) as the quoted reference does not relate to UC nor does it provide utility values. The correct reference is 162.	This is a misreferenced study. We will amend this in the monograph.
219	Table 65 Health state utilities for all presurgical and post-surgical states Woehl *et al*[109]		This is a misreferenced study. We will amend this in the monograph.
225	Of the 10 potentially relevant EQ-5D studies, those reported by Woehl *et a*[109] …		This is a misreferenced study. We will amend this in the monograph.
225	Woehl et al[109] collected EQ-5D utility scores from 180 patients with active UC.		This is a misreferenced study. We will amend this in the monograph.
227	The utility score for post-surgery was based on the mean value reported by Woehl et al[109] …		This is a misreferenced study. We will amend this in the monograph.
245	Table 85 Source of health utilities Woehl et al[160] (chronic pouchitis valued using Arseneau et al110)	The incorrect reference number is assigned to Woehl in the text. The correct reference is 162.	This is a misreferenced study. We will amend this in the monograph.
247	Where considered appropriate by the physician, patients, family members and/or carers require training for the administration of subcutaneous injections. This training is associated with additional resource use and costs.	Where considered appropriate by the physician, patients, family members and/or carers require training for the administration of subcutaneous injections. This training is associated with additional resource use and costs, **but is provided free with adalimumab through AbbVie Care.**	Irrespective of who bears the costs, they are still costs.
284	High risk - although ITT analysis was undertaken, there was a high level of attrition and an imbalance between treatment groups (PBO, 50%; ADA, 59%)	It is not clear to AbbVie to what the 50% and 59% in the quoted text relates to. In the ULTRA2 study, 246 patients were randomised to placebo of which 131 completed the study (53%).	This is a typo and will be amended in the monograph. In the placebo arm, 115 of the 260 patients (44.2%) originally randomised to placebo discontinued. In the adalimumab 160/80/40

		In the adalimumab arm, 154 (62%) patients completed the study out of 248 who were randomised.	arm, 94 of the 258 patients (36.4%) originally randomised to adalimumab discontinued. As such, ULTRA2 would still be classed by the AG as high risk for attrition bias.

MSD – Assessment Group response to fact-check

Page	Issue/factual inaccuracy	AG response
4	The AG has suggested a potential bias arising from the PURSUIT design as those patients who received golimumab in PURSUIT-Maintenance were re-randomised golimumab-induction-responders (distinct from the treat-through designs of ACT and ULTRA where both responders and non-responders went on to receive active drug in the maintenance phase). However, one must also be conscious of another major difference in study design in terms of how both continuity of response and loss of response were defined. The primary endpoint of PURSUIT-Maintenance was <u>continuous clinical response</u> throughout 54 weeks, i.e. for a patient to achieve the primary endpoint, they were required to be in a state of clinical response at each 4-weekly assessment (not only at 2 or 3 time points), and a loss of response at any individual time point (detected via partial Mayo scoring and confirmed by colonoscopy) constituted treatment failure. This is in contrast to the designs of ACT and ULTRA. In ACT, patients were assessed for response at weeks 8, 30, and 54 (ACT 1 only), and patients in clinical response at each of these time points met the criteria for <u>sustained clinical response</u>. In ULTRA2, patients were assessed for response at weeks 8, 32, and 52, and a patient in clinical response at both weeks 8 and 52 met the criteria for <u>sustained clinical response</u>. In ULTRA2 inadequate response was defined as: (1) partial Mayo score equal to or above baseline score on 2 consecutive visits at least 14 days apart (for patients with a partial Mayo score of 4-7 at baseline); (2) partial Mayo score ≥7 on 2 consecutive visits at least 14 days apart (for patients with a partial Mayo score of 8 or 9 at baseline) - one could make the argument that it was therefore more difficult for a patient to achieve the primary endpoint in PURSUIT.	This is not a matter of factual inaccuracy. It is another source of potential heterogeneity.

5	Data for infliximab biosimilars are included in the AG report, however, their inclusion is inappropriate because at the time of the technology appraisal guidance being issued biosimilars will not yet be available for use in UK clinical practice (prior to the Remicade patent expiry). In addition, there are no clinical data relating to the UC indication to support considerations of clinical effectiveness, and no list price to support considerations of cost-effectiveness.	This is not a matter of factual inaccuracy. See comments from Celltrion.
10	The AG report states that: "*The risk of relapse and disease flares is increased by poor adherence to medication regimens*". We note that the 4-weekly administration of golimumab and the 8-weekly administration of infliximab in the maintenance setting could be expected to facilitate better adherence, compared with medicines that need to be administered more frequently.	This is not a matter of factual inaccuracy
14	The economic burden of UC is reported in Euros; it would be useful to convert this value to GBP.	This is not a matter of factual inaccuracy
25	Consistency in reviewer inclusion/exclusion decisions was assessed during first passing (title and abstract screening) but not for second passing (full-text screening). It is therefore unclear whether all inclusion decisions were consistent between reviewers.	Two reviewers closely discussed inclusion decisions at the full text stage.
32	The AG acknowledges that: "*None of the included RCTs applied Truelove and Witts' disease severity criteria*". This was expected and was discussed at the scoping workshop. Using the Mayo score to classify disease severity may be more appropriate given the evidence base.	Agreed.
33	It is inappropriate to include UC-SUCCESS in the AG systematic review as the population deviates from that specified in the final scope: the patients in UC SUCCESS were required to be either azathioprine-naïve or free from azathioprine treatment for at least 3 months before enrollment (only ~10% of patients had prior immunomodulator exposure at baseline so this represents a very different population to that in the TNF-α inhibitors licences and the scope of this appraisal).	This was a borderline inclusion and is stated as such. This study was not included in the NMA.
95	In Figure 14, it is not acknowledged that infliximab is statistically significantly superior to adalimumab in clinical response/remission in induction (-0.52; 95% CrI: -10.03, 0.00) (in Figure 41 a credible interval meeting 0.00 is stated to be a significant result).	This is not factually incorrect.

105	The AG acknowledges the limited data in the NMA and used a random effects model with an informative prior distribution for between-study heterogeneity (whereas we use a fixed effects model for the maintenance NMA). We note that where there is only one trial per connection in the network, the choice of the prior distribution for the between-study heterogeneity is directly incorporated in the posterior distributions of the treatment effects and is not updated based on the data. As such, the results of the NMA (i.e. 95% CrI and associated rankings) are arguably more influenced by the choice of prior distribution than by actual trial data. In this context we would like to emphasise the very wide intervals for remission in Table 16 and 17.	This is not factually incorrect. A fixed effects model assumed that trials are homogeneous, which is unlikely to be the case using our dataset. We have used a random effects model to reflect that trials are heterogeneous. We have selected a weak prior so that the results are in a plausible range.
156	The AG report states that: there is "*a potential inconclusive benefit for intervention-treated patients compared with placebo*". It is not clear why this conclusion has been reached given the data from Sandborn *et al*. 2009 which indicates a 41% reduction in the risk of colectomy with infliximab relative to placebo (p. 74 of the AG report).	This is not factually incorrect.
177	The MSD NMA includes data for unlicensed doses from the pivotal trials, however, only odds ratios relating to licensed doses were included in the MSD cost-effectiveness model.	This is not factually incorrect.
180	The possible transitions in the model were limited by the available data for all competing TNF-α inhibitors. The duration of drug maintenance treatment is reflected with the time spent in the health states 'Response (pre-colectomy; maintenance)' and 'Remission (pre-colectomy; maintenance)' and is based on efficacy data. Upon losing remission patients transition to the health state 'Response (pre-colectomy; maintenance)'. Upon losing response, patients transition to the health state 'Relapse (relapse management)'. To ensure that the modeled duration of treatment with a TNF-α inhibitor matched the observed discontinuation rates of biologic treatment in the trial for golimumab, we had to assume that patients losing remission do not transition straight to 'Relapse (relapse management)'; the corresponding transition probability was set to 0. However, we would like to note that the model structure/programming would allow such transitions, which was the reason for the corresponding arrow in the figure.	This is not factually incorrect. The value was zero within the model which means that this transition is not possible.

184	The description of the golimumab induction dosing regimen in Table 46 is incorrect. Patients do not receive 2x200mg then 2x100mg; rather patients receive 2x100mg at week 0 and 1x100mg at week 2. Given that the cost per cycle is calculated for 8 weeks (patients will receive 2x100mg at week 0, 1x100mg at week 2, and 1x50mg or 100mg at week 8), the cost of £3051.88 is correct when the PAS for golimumab is applied (which it has not been elsewhere in the AG report).	This is a typo. We will correct it for the monograph.
189	The AG report states that: "*The DSAs indicate that the cost-effectiveness results for golimumab versus colectomy are sensitive to the utility values for pre-colectomy remission and pre-colectomy relapse*". This is not the conclusion presented in the tornado diagrams in the MSD submission.	This is a typo. We will amend to "post-colectomy remission" only.
214	The population considered in the AG model comprises those who have failed at least one prior therapy "*as reflected in the RCTs*": this does not reflect the RCT populations. Further, the licences for golimumab, infliximab, and adalimumab refer to 'moderately to severely active ulcerative colitis in adult patients who have had an inadequate response to conventional therapy including corticosteroids and 6-mercaptopurine (6-MP) or azathioprine (AZA), or who are intolerant to or have medical contraindications for such therapies' - meaning that the relevant population should include failure of two prior therapies.	Not a factual inaccuracy.
215	In the AG model, patients in the conventional treatment group and those who have previously achieved but subsequently lost response to biologic therapy are assumed to continue receiving conventional therapy irrespective of whether response or remission is achieved. The rationale for continuing treatments which have been ineffective is unclear.	Not a factual inaccuracy.
219	In Table 65 the evidence source for pre- and post-surgical utility states is given as Woehl *et al*. (reference 109). The same reference was (rightly) criticised in the MSD submission (p. 176, p.181 of the AG report); it was referenced by Tsai *et al*. but does not contain the relevant information. Incorrect referencing of the Woehl study also occurs on p. 225 and p.227 of the AG report.	Agreed. This will be amended in the monograph.
220	The transition probabilities in Table 66 indicate that after patients fail on biologic treatment, they go onto conventional therapy. At this point the AG uses the placebo transitions and permits patients to move to remission/response as a consequence of being placed on conventional therapy despite the patient having already failed both biologic and conventional therapy.	Not a factual inaccuracy. This is based on the observed transitions in the trials.

222	The AG identified 6 potentially relevant studies for colectomy rates (Table 67), selecting Solberg *et al*. (2009) primarily due to its size. The cumulative colectomy rate after 10 years was 9.8% (95% CI: 7.4%, 12.4%). However, the population comprises newly-diagnosed UC patients who may be different to a moderate to severe population initiating treatment with a biologic. In fact, the average disease duration at baseline in the trials forming the basis for the economic model is ~6 years (Table 8, page 53). Furthermore, the Solberg study enrolled patients between 1990 and 1994, which suggests that the patients will not have received biologics during follow-up. As such, we have concerns around the use of this study to determine colectomy rates. Although limited data were reported from included trials for the outcome of time to surgical intervention, Sandborn *et al*. (2009) combined data from ACT1/2 and reported that the cumulative incidence of colectomy through 54 weeks was higher for the placebo group (17%) than for the infliximab group (10%). These infliximab estimates correspond to a cumulative incidence of colectomy of 19% at 2 years, 41% at 5 years, and 65% at 10 years, and a 6-month incidence rate of 0.053. This 0.053 rate is ~10 times larger than the 6-month rate of 0.0051 used by the AG. Overall, it appears that the colectomy rate has been underestimated by the AG. Our model assumed that within 1 year, 15%-20% patients are in the colectomy/post-colectomy health states. This was criticised for being higher than the proportion of patients undergoing surgery in RCTs (colectomy rates in the RCTs included in the systematic review were 0.7% to 5.8% in individual trial arms at ~1 year). However, reference to the RCT rates further underlines that the estimate used by the AG is too low.	We agree that the colectomy rate is a source of uncertainty. However the rates reported by Solberg are in line with expert opinion received by the assessment group.
228	In Table 72 the induction dosing regimens are not described correctly. When the PAS is applied to the golimumab cost, the cost for the induction phase is £2288.91 and the cost for the maintenance phase is £4976.34 irrespective of whether 50mg or 100mg golimumab is received.	Not a factual inaccuracy. At the request of NICE, the PAS was not included in the original AG analysis.

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