National Institute for Health and Care Excellence

Multiple Technology Appraisal

Infliximab, Adalimumab and Golimumab for Treating Moderately to Severely Active Ulcerative Colitis after the Failure of Conventional Therapy

AbbVie's Response to the Assessment Report

5 August 2014

Executive Summary

AbbVie welcomes the opportunity to comment on the Assessment Group Report.

This document is structured into three sections, namely comments on the Assessment Group model, including factual inaccuracies (

Section **1**), response to the Assessment Group critique of the AbbVie model (

Section 2) and details about the changes made to the original AbbVie submission model comparing adalimumab to conventional non-anti-TNF therapy in responding to critiques from the Assessment Group (Section 3). A description of the submission model comparing to surgery will also be presented as will summary results for the two models.

AbbVie believes that there are a number of assumptions within the Assessment Group model that results in increased uncertainty around the true value of the incremental cost effectiveness ratio (ICER) for anti-Tumour Necrosis Factor (TNF) therapies and adalimumab in particular; these are briefly outlined below, and discussed in full in Section 1 under the subsections indicated

1.1 The model relies on assumptions around the definition and measurements of response, remission and non-response that underestimates the efficacy of adalimumab in UK clinical practice. Specifically, the adalimumab trials used the worst-ranking Mayo-score to measure response and employed the non-responder imputation (NRI) method. Using the worst-ranking Mayo-score, a method not employed by the other anti-TNF therapies, and defining patients who discontinued the study for any reason, had missing Mayo Scores or dose-escalated as "non-responders" results in a conservative estimation of the efficacy of adalimumab. Relying on these efficacy assumptions in the model results in an ICER that is not a true reflection of the cost-effectiveness of adalimumab.

1.2 The modelled assumption on dose escalation does not employ an escalation rate that is reflective of the population under evaluation and the model assumptions on non-responders does not account for the efficacy of adalimumab dose escalation. As such, the model considers the additional costs associated with dose escalation without simultaneously incorporating the benefit, thereby overestimating the cost of adalimumab and underestimating its efficacy, resulting in a high ICER for adalimumab.

1.3 The treatment sequence strategy assumed in the model does not reflect UK clinical practice. In particular, patients who fail anti-TNF therapy are not given the option to receive surgery and have to continue on other conventional therapies which have shown to be ineffective (as per the definition of the population in the decision problem scope). Consequently, the ICER favours surgery and conventional therapy over anti-TNFs.

1.4 The model uses probabilities of proceeding to surgery that may not be reflective of the patient population under evaluation. This results in an ICER that favours the surgical cohort as fewer patients on anti-TNFs and conventional treatment proceed to surgery, which in the model is associated with better outcomes.

1.5 The model does not fully account for the cost and disutilities of surgery-related transient and chronic complications, causing the cost of surgery to appear lower than it is in clinical practice and the outcomes to appear better than they are. It also does not consider the disutility of surgery itself. This results in an ICER that favours surgery over anti-TNFs and conventional treatment.

1.6 The cost of surgery is underestimated in the model, particularly the assumptions on the number of surgical procedures that will be required. There is also no consideration for the cost of long term care in patients who have had a stoma fitted. These cost underestimations favour the results in the surgical and conventional treatment cohorts as more patients in these cohorts proceed to surgery.

1.7 Hospitalisation risk in the model may not reflect the relative risk (RR) observed in clinical trials of adalimumab. Specifically, adalimumab reduced the risk by 30% for all cause hospitalisation compared to

placebo and by 50% for UC-related hospitalisation compared to placebo.[1] The risk reduction in the model was 30%, but is it is not clear how this was derived or if it relates to all-cause or UC-related hospitalisations. Employing a 30% risk reduction when it should be 50% increases the cost associated with adalimumab, resulting in a higher ICER compared to conventional treatment and surgery.

1.8 The cost and doses used for oral 5-aminosalicylates (5-ASAs) do not reflect UK clinical practice for moderate to severe UC patients in the UK. The underestimation of costs of 5-ASAs increases the ICER of anti-TNF therapy versus conventional therapy.

1.9 The model does not incorporate the cost of dying, and in particular, the cost of dying following surgery. Not incorporating a cost of death underestimates the true costs of the treatment options, especially if patients die following surgery. Omitting the cost associated with surgery related deaths worsens the ICER for conventional treatment and anti-TNFs against surgery.

AbbVie will also address the Assessment Group's criticism of the AbbVie model, providing further information and justification where necessary.

In response to the critiques from the Assessment Group, **AbbVie has also updated the original AbbVie model comparing adalimumab to conventional non-anti-TNF therapy using a life time horizon, and revised the comparator to surgery in a separate model, thereby allowing a comparison of adalimumab to surgery using a life time horizon**. A comparison to other anti-TNF treatments is not, however, undertaken due to limitations in the available evidence that would prevent a reliable comparison across different anti-TNF therapies.

The first model supplied with this response is an updated version of the original submission model, which focuses on comparing adalimumab to conventional therapy (referred to in this response document as the updated submission model), and the second model focuses on comparing adalimumab to surgery (referred to in this response document as the model comparing adalimumab to surgery). In the revised models, the ICER **was £23,027 comparing adalimumab to conventional therapy and £15,750 comparing adalimumab to surgery** at life time among moderate and severe UC patients who failed conventional therapy, respectively**.** These results are robust in deterministic sensitivity analyses, with ICER ranging from £18,218 to £ £25,594 per QALY gained comparing adalimumab to conventional therapy, and from £13,172 to £21,719 per QALY gained comparing adalimumab to surgery. The probability that adalimumab is cost-effective at the willingness to pay threshold of £30,000 per QALY gained is 73% compared to conventional therapy, and 94% compared to surgery.

Section 1. Comments of the Assessment Group Model

1.1 The model relies on assumptions around the definition and measurements of response, remission and non-response that underestimates the efficacy of adalimumab.

AbbVie believes that the assumptions and methods used in the model underestimate the efficacy of adalimumab, resulting in a high ICER:

1.1.1 The differences in Mayo scoring methods across the adalimumab, infliximab, and golimumab trials can cause the efficacy for adalimumab to be underestimated relative to the anti-TNFs. As will be discussed in detail in

1.1.2 Section 2, the differences in the "worst scoring method" used in the adalimumab trials and the "average scoring method" used in the other anti-TNF trials results in a decreased treatment efficacy for adalimumab, which used the worst scoring method, compared to efficacy results for infliximab and golimumab.

1.1.3 The NRI method used in the adalimumab trials to classify non-responders, does not properly account for the efficacy of adalimumab among patients who respond to adalimumab, but discontinued the trial due to non-efficacy reasons. The use of a response based model based on NRI from the adalimumab clinical trials (besides the fact it does not consider efficacy from dose escalation of adalimumab) underestimates the efficacy of adalimumab as NRI assumes that patients who discontinue therapy for any reason are non-responders. Patients in the ULTRA2 trial discontinued treatment for multiple reasons, including loss of efficacy, loss of follow up, and protocol violations. As stated in the AbbVie submission, among the 15 anti-TNF naïve patients who discontinued adalimumab in ULTRA2 for non-efficacy reasons (i.e., loss of follow up and protocol violations), 10 patients (66%) were at remission or response at the time of discontinuation. These patients would have been categorised as non-responders in the trial and would have discontinued adalimumab in the model, although in UK practice these patients might have achieved response or remission.

1.1.4 The efficacy estimates for adalimumab employed in the model do not consider the efficacy from adalimumab dose escalation. As discussed above and in Section 1.2 below, the NRI method taken from the trial and applied in the model assumes that all adalimumab-treated patients who dose-escalate are non-responders, which is contradictory to the clinical evidence.

1.1.5 The discontinuation rule of anti-TNF therapy in the model may also not adequately capture the benefits of adalimumab. The model assumes that after week 8, patients who relapse discontinue adalimumab and begin conventional treatment. It should be acknowledged that UC symptoms may wax and wane over time, even when patients are receiving effective therapy. Based on results from the ULTRA2 trial[3], 9 out of 14 UC patients (64.3%) who achieved response at week 8 and escalated to once-weekly dosing with adalimumab after relapsing on treatment, regained remission or response within the trial period following escalation (based on observed Mayo scores)

1.1.6 The study design in ULTRA2 allowed open-label rescue from week 12. Clinicians and patients were aware of the study design and that it was possible for patients with protocol-defined inadequate response to move from the blinded arm and receive open label adalimumab treatment. AbbVie felt that in a parallel group design study, it would be difficult for ethics committees and investigators to justify putting an already very ill patient on placebo for 52 weeks and therefore the company felt it was appropriate to include a rescue arm. In order to retain blinding, this option had to be offered for both arms of the study. To account for this, AbbVie employed the NRI analysis, so patients who dropped out of the trial as well as patients who completed the study while receiving open label drug (because they may have escaped for the reason of inadequate response) would be considered not to have achieved the efficacy endpoints.

The effect of underestimating the efficacy of adalimumab in the model results in a high ICER for adalimumab compared to surgery and conventional therapy.

AbbVie therefore wishes these points to be taken into consideration and discussed in more detail by the Assessment Group when presenting the results and conclusions from the model. AbbVie also suggests a

scenario with an increased efficacy of adalimumab as the current ICER underestimates the effectiveness of adalimumab.

1.2 The modelled assumption on dose escalation does not employ an escalation rate that is reflective of the population under evaluation and the model assumptions on non-responders does not account for the efficacy of adalimumab dose escalation.

The rate of dose escalation (27.4%) used in the model would not reflect the use of adalimumab according to the licence, which states that some patients who experience a decrease in their response may benefit from a dose increase.[2] The licensed indication wording therefore indicates that only patients achieving a response should be considered for dose escalation. As a result, using dose escalation at the rate applied in the model overestimates the cost of adalimumab.

The rate relied upon in the model is based on all patients in the ULTRA2 trial (including both anti-TNF naïve and experienced patients) and was not derived from week 8 responders. Since the original AbbVie submission, Wolf et al. (2014) has published the rate of dose escalation among week 8 responders in the ULTRA2 clinical trial.[3] In the publication the rate of escalation was 16.3% for week 8 responders and 38.4% for week 8 non-responders. As the model only uses the data from the week 8 responders, it would be appropriate to use a dose escalation rate of 16.3%. Additionally, the model has assumed patients with dose escalation will escalate adalimumab immediately after week 8. This assumption, however, does not reflect clinical practice where the median time to dose escalation is approximately 41 weeks from the start of therapy.[4] In the ULTRA2 study, the median time to weekly dosing in the week 8 responder group was not estimable due to low rate of dose escalation which meant it had not been reached at the end of 52 weeks. In the original and the updated AbbVie submission models, a dose intensity approach was utilised which takes the timing of dose escalation into consideration.

Most importantly however, the definition of treatment efficacy of adalimumab in the ULTRA1 and 2 studies were based on NRI, a method which categorises all patients who dose-escalate on adalimumab as non-responders irrespective of Mayo score improvement. As patients who respond, or are in remission, and who dose escalate were defined as non-responders in the trial protocol, dose escalation is not applicable to the adalimumab cohort in the model. This is because these patients have been indirectly classified as non-responders in the model and therefore would stop treatment in the model.

Additionally, discontinuing patients in the model who were classified as non-responders in the clinical trial is not in line with the clinical trial protocol where patients who were non-responders could continue treatment. The indirect assumption in the model that patients who dose escalate would stop their adalimumab treatment does not take into account that patients could have achieved better efficacy. Based on results from the ULTRA2 trial[3], 9 out of 14 UC patients (64.3%) who achieved response at week 8 and escalated to once-weekly dosing with adalimumab after relapsing on treatment, regained remission or response within the trial period following escalation (based on observed Mayo scores).

In the model, drug costs associated with adalimumab dose escalation are therefore included, while at the same time the model indirectly makes the assumption that patients who dose-escalate are non-responders and thus discontinue adalimumab treatment. If the model is to assume (as it currently does) that dose escalation with adalimumab leads to non-response, then it is not appropriate to consider dose escalation in the model as an option available to adalimumab patients, nor is it appropriate to include the full cost of such dose escalation when evaluating the ICER of adalimumab compared to conventional therapy and surgery.

By making the assumption that patients with dose escalation are non-responders to adalimumab, yet including the full cost of dose escalation, AbbVie believes that the model overestimates the cost, and thereby the ICER, for adalimumab compared to surgery and conventional therapy.

AbbVie therefore suggest that no dose escalation be used in the model, or that the model incorporates the additional benefit that dose escalation with adalimumab offers.

1.3 The treatment sequence strategy and patient pathway assumed in the Assessment Group model does not reflect UK clinical practice

The model assumes that all patients in the anti-TNF arm who discontinue therapy at week 8 due to non-response will receive conventional treatment and will only be able transition to surgery after receiving at least one cycle of conventional treatment.

In a population who has already failed or are intolerant to conventional therapy (as per the decision problem scope), it seems counterintuitive to require that all patients who failed anti-TNFs start with treatment they are intolerant to or treatment that has shown to be less effective than another treatment option, such as surgery. Whilst the model allows patients on anti-TNFs to proceed to surgery after receiving at least one cycle of conventional treatment, the probability of proceeding to surgery in patients on conventional treatment is 0% following the induction phase and <1% following the 26-week maintenance phases. Allowing less than 1% of patients to proceed to surgery whilst on conventional treatment (which has shown to be less effective than surgery), appears counterintuitive and does not seem ethical or reflective of clinical practice. Indeed, the British Society of Gastroenterology (BSG) guidelines for the management of inflammatory bowel disease (IBD) states that there is evidence that delay in surgery as a result of prolonged first or second line medical therapy in patients with severe IBD or who are refractory to treatment, may increase morbidity[5]. Additionally, feedback obtained by AbbVie from gastroenterologists with a special interest in IBD indicated that patients who failed conventional treatment and then failed anti-TNFs would have the option to have surgery rather than go on to conventional therapy. Results from the 2014 IBD Audit provides some support for the BSG guidelines and clinician feedback: 45% of patients who were admitted for elective surgery and 78% of patients admitted for non-elective surgery did so due to failure of medical therapy[6].

As a result of this treatment sequence pathway, patients on the anti-TNF treatments incur the additional cost of conventional treatment, but with less benefit than if they progressed to surgery. This weights the ICER against anti-TNFs, favouring the conventional treatment and surgery cohorts.

In light of the feedback from clinicians and evidence from the BSG guidelines, AbbVie disagrees with the treatment sequence applied and suggests that the AG model be amended to allow at least some, if not all, patients to proceed to surgery following the failure of anti-TNFs.

1.4 The model uses probabilities of surgery that may not be reflective of the patient population under evaluation.

Solberg et al.[7] was selected for inclusion in the model on the basis that the IBSEN study was large and did not specifically relate to patients who had experienced UC flare. Hillson et al., used by AbbVie, was dismissed on the basis that (amongst others) it does not specifically relate to a moderate to severe UC populations.

AbbVie believes that Solberg et al. likely underestimates the probability of surgery given the differences in patient populations between Solberg et al. and the evaluation population. Solberg et al. is a Norwegian study which reports rates of surgery that are lower than other rates identified by the Assessment Group's review of the literature (Table 67 in the Assessment Report). Solberg et al. reports a surgery rate of 9.8% at 10 years, whereas the rate of surgery identified in the Assessment Group's review ranged from 14.5% to 65.0% among other publications with similar or even shorter follow-up times.

AbbVie also believes that Solberg et al. underestimates the rate of surgery in the evaluation population for several reasons. First, Solberg et al. focuses on newly diagnosed UC patients, a patient population that is inconsistent with the population in the present evaluation (patients who have previously experienced inadequate response to conventional non-anti-TNF therapy). Typically, patients with moderate to severe UC are not newly diagnosed and have in fact experienced disease conditions for several years. For example, the ULTRA2 trial includes a patient population that previously experienced inadequate response to conventional therapy. On average, patients in the ULTRA2 trial had disease duration of 8.3 years.[8] Additionally, Solberg et al. includes newly diagnosed UC patients of all severity levels, not just patients with moderate or severe disease activity. Indeed, Figure 1 of Solberg et al. shows that 55% of patients described their disease over the previous 10 years as in remission or mild after initial high activity. Furthermore, the surgery rates reported in Solberg et al. were not stratified by disease severity. It is reasonable to assume that the surgery rate among moderate to severe UC patients will be higher than the average rate observed among all newly diagnosed UC patients. This was the finding in the Hillson et al. study: although the overall one-year prevalence of surgery was small (2.7%), the percentage of severe UC patients requiring colectomy was substantially higher (18.2%)[9]. This is also evident in the results of Solberg et al., as patients with erythrocyte sedimentation rate (ESR) ≥ 30 mm were reported to have a much higher colectomy rate than patients with ESR < 30mm, with an estimated rate of 21% versus 3%. As has been demonstrated in multiple publications, high ESR is known to be correlated with more advanced UC disease severity;[10,11,12] based on the Truelove and Witts scale, moderate UC is characterised by ESR values of 20mm-30mm, and severe UC is characterised by ESR values of > 30mm. As noted from Solberg et al., the majority of patients (366/ 519) had an ESR <30 mm, indicating a population with mild to moderate disease.

Given the above considerations, AbbVie thinks that the surgery rate used in the model is not reflective of UK clinical practice and underestimates the frequency of surgery in a moderate to severe UC population. This results in an ICER that favours the surgical cohort as fewer patients on anti-TNFs and conventional treatment proceed to surgery, which in the Assessment Group model, has shown to offer better outcomes.

As such, AbbVie recommends that consideration should be given to using an alternative source for surgery rates.

1.5 The model does not fully account for the costs and disutilities associated with surgery-related complications and does not consider the disutility of surgery

In its model, there is an assumption that patients may experience both transient and chronic complications related to surgery, and that transient complications can only arise during the same cycle as surgery and not in subsequent cycles. Furthermore, the model uses Woehl et al. to derive post-surgery utilities, assuming that there are no disutilities associated with transient complications. This simplifying assumption does not, however, reflect the consequences and complications of surgical intervention in UK clinical practice.

For instance, patients may incur surgery-related complications many years after colectomy, may develop complications that are recurring, and may even develop multiple complications. In fact, Arai et al.,[13] a study

used in the Assessment Group model, illustrates that many transient complications occur at a mean onset time of more than one year following surgery (Table 2 in the Arai publication). The model, however, only considers surgery-related transient complications within the same cycle (26 weeks) of surgery and does not include the flexibility to incorporate surgery-related complications which occur more than 26 weeks after surgery. Additionally, Arai et al. demonstrates that patients may have multiple complications (Table 3 in the publication); a total of 199 surgically-related complications were documented in Arai et al. among the 156 patients who experience post-surgical complications (Table 2 in the publication). However, the calculation of the complication rate related to colectomy does not take this into consideration. It is also worth noting that the rate of pouchitis reported in Arai et al. is lower than other rates reported in the literature among Western populations. This difference is acknowledged by Arai et al., and is attributed to the exclusion of certain types of pouchitis[13]. In addition, Arai et al. was conducted in a Japanese healthcare setting and among a Japanese population, therefore it is unlikely that these results are generalisable to patients in the UK.

Furthermore, as transient complications may require surgery, hospitalisation, and additional medical treatments, AbbVie believes that it is not reasonable to assume there are no disutilities for transient complications.

It is not clear from the report what the rationale is for only selecting pouchitis as a chronic complication. In Arai et al., for example, staple line ulcers occur in nearly twice as many patients with a similar mean onset time as pouchitis (1.50 years and 1.72 years, respectively)[13], yet staple line ulcers are not included as a chronic complication in the Assessment Group's model. In addition, other important chronic complications associated with high costs and/ or worse quality of life that have been documented in previous studies (e.g. infertility and sexual impotence) were not considered in the model[14,15,16,17,18,19].

As a consequence, AbbVie believes the model underestimates the chronic complications associated with surgery. The BSG guidelines, for example, makes note of the complications of surgery and specifically mentioning pouchitis, pouch failure, nocturnal seepage, urgency and the ability of women to reproduce[20]. To this point, evidence from two meta-analyses[21,22] based on a large number of observational studies provides information on the complications after ileoanal pouch anal anastomosis (IPAA). Of note in these meta-analyses are the rates of pouch failure (4.3% and 8.5%), which has not been taken into consideration in the economic evaluation. Burns et al.[23] offer UK specific information on, amongst others, pouch failure rates in UC patients based on Hospital Episode Statistics data. They found that over 12 years in the UK, pouch failure occurred in 6.4% of cases. More recently, the BSG guidelines state that the incidence of failure was 7.7% based on the UK pouch registry[20]. As pouch failure requires further surgery and long term care, the cost of this complication may be substantial. It is also not unreasonable to expect that it would also affect patients' quality of life.

In a recent systematic review, Leonard et al.[24] investigated complications associated with colectomy in patients with UC. In their meta-analysis, 17.8% of patients experienced small bowel obstruction, amongst others, with 24.5% of these requiring re-operation. Again, the cost of further surgery for this complication may be substantial.

AbbVie also wants to reiterate that surgery can also lead to sexual dysfunction in men and women caused by injuries to nerve plexuses or distortions in the pelvic structure sustained during surgery, or fertility issues. The BSG guidelines state that reproductive ability in women may be reduced by 40-50% following IPAA[20]. As noted in the Swinburn study, approximately 40% of the individuals who had undergone surgery reported experiencing problems related to sexual functioning and 10% had fertility issues. Bach et al. states that 38%

of patients who had undergone IPAA were unable to conceive following 1-year of unprotected intercourse[25] whilst Leonard et al.[24] found there was a six fold increase in the prevalence of infertility post-surgery (7.8% pre-surgery vs. 47.9% post-surgery). The cost of treating sexual dysfunction in affected men and in vitro fertilisation in affected women and female partners of affected men, could add considerably to the cost of surgery.

Lastly, the model assumes there is no disutility for surgery. Surgery for UC patients is a major procedure that requires hospitalisation. In AbbVie's view, it is not reasonable to assume there is no disutility associated with surgery.

Due to the reasons listed above, AbbVie concludes that the post-surgery utility has been overestimated in the economic evaluation and the costs associated with surgery underestimated, making surgery appear to be a more attractive treatment option compared to conventional and anti-TNF therapy.

AbbVie therefore suggests that consideration should be given to incorporating a disutility for transient complications, using more appropriate sources for complications, incorporating all relevant chronic complications and including the cost of treating these additional, transient and chronic complications.

1.6 The cost of the surgery procedure and follow up procedures are underestimated in the model

The economic evaluation makes a number of assumptions regarding the cost of surgery with which AbbVie disagrees. Each of these points is discussed below.

1.6.1 The model only incorporates the cost of one surgery for patients who proceed to surgical intervention. It therefore does not consider the probability that a patient may require more than one surgical procedure. As the Crohn's and Colitus UK website and NHS Choices website point out, IPAA generally requires two operations[26,27]. This is supported by feedback from clinicians that few patients will undergo only one surgical procedure, some will require two and most will require three. Further evidence to this point can be found in Tsai et al. where the cost of IPAA was assumed to be equivalent to two complex and one major procedure, and ileostomy was assumed to be equivalent to two complex procedures[28].

Therefore, only incorporating the cost of one procedure in the model underestimates the cost of surgery, thereby favouring the surgery cohort over the other treatment cohorts and making surgery appear more cost-effective than it is in clinical practice.

AbbVie therefore suggests that the assumptions that only one surgical procedure is performed be reconsidered and that the cost of additional surgical procedures in the model be incorporated.

1.6.2 The economic evaluation also does not incorporate the cost of stoma care. As discussed by CCUK, patients that undergo a proctolectomy with ileostomy will require an external bag that is fitted onto the ileostomy to collect waste. This bag is emptied or changed as necessary and is required for the lifetime of the patient. Patients who undergo IPAA will require a temporary stoma bag until the newly-formed pouch has healed – which generally takes around three months. These stoma bags will need to be changed about twice a week, but this depends on the type of stoma bag as some may need to be changed each time stools are passed. Visits to the stoma nurse will initially be required to familiarise the patient with the stoma bag and educate them about the aftercare. Buchanan et al. in their economic evaluation estimated continued follow-up visits to be every four

months, based on feedback from UK clinicians and nurses[29]. Belts, retention strips, adhesive removers, skin protectives, fillers, deodorants, stoma underwear, support garments and bag covers are all accessories that may be required for use in conjunction with the stoma bag.[26,29,30,31,32]

The cost of these bags and accessories are covered by the NHS[33,34]. As these costs have been omitted from the model, the true cost of surgery is underestimated and favours the ICER for the surgery cohort.

AbbVie recommends that the long term costs of surgery be reconsidered and incorporated into the model.

1.7 Hospitalisation risk does not reflect the relative risk observed in clinical trials of adalimumab

As stated in the AbbVie submission, adalimumab reduced the risk by 30% for all cause hospitalisation compared to placebo and by 50% for all UC-related-hospitalisation compared to placebo.[35] However, in the model a RR of 0.70 was used when comparing hospitalisations between adalimumab and placebo, based on the MSD NMA. As stated by the report, there is lack of clarity on the MSD NMA model, specifically, the methods whereby the results therein were obtained. It is also not clear if this figure relates to all-cause or UC-related hospitalisations. As a result of using the MSD input, potentially using a 30% risk reduction when it should be 50% increases the cost associated with adalimumab, resulting in a higher ICER compared to conventional treatment and surgery.

AbbVie believes it is not appropriate to use the results from the MSD NMA and recommends that the RR from the Feagan publication is used.

1.8 Oral 5-ASA costs and doses do not reflect UK clinical practice for moderate to severe UC patients

Within the model assumptions are made regarding the cost and doses of oral 5-ASA which result in an underestimation of the cost of conventional therapy.

1.8.1 5-ASAs are indicated for the treatment of mild to moderate UC and maintenance of remission. As such, their recommended doses may not reflect the patient population under review, namely moderate to severe UC patients. Specifically, the report states that the dose for mesalazine is 2.4mg/day (the model itself employs the cost of 2.4g/day.) Whilst the difference between the doses in the report and model appears to be a typographical error, AbbVie notes that the maximum licensed dose for Ipocol® is 2.4g whereas the maximum dose for other preparations exceeds 2.4g, and is up to 4.8g[36]. The BSG guidelines state that greater clinical improvement is associated with doses >3g/day and although the rate of remission is similar on 2.4g/day and 4.8/day, faster symptom resolution occurs on 4.8g/day compared to 2.4g/day. Furthermore, the European Crohn's and Colitis Organisation (ECCO)[37] states that doses ≥2.0g/day are more effective than <2.0g/day for remission. Gastroenterologists with whom AbbVie has consulted have indicated that the doses used in the model for mesalazine would not be appropriate for patients with moderate to severe UC. Specifically, for the induction phase a dose of up to 4.8g/day would be used. Similarly, for the maintenance of remission in patients with moderate to severe UC, the higher maintenance dose would be used, i.e. 2.4g/day. Using an inappropriately low dose for 5-ASAs underestimates the cost of the conventional treatment group and results in a high ICER for anti-TNFs compared to the conventional treatment group.

AbbVie therefore suggests that the cost of 5-ASAs should be based on the recommended doses from the BSG and ECCO.

1.8.2 The model also uses the cost of the least expensive conventional treatment product where multiple products were available. AbbVie believes that using this assumption artificially lowers the cost of conventional treatments by underestimating the cost thereof. Using the least expensive drug is a strong assumption to make as it assumes that pharmacies will always stock the least expensive brand. Furthermore, it does not take into account that the extended release preparations may be preferred as it allows once daily dosing compared to divided doses with non-extended release preparations. A more accurate assumption would be to use the average cost of all the drugs.

AbbVie therefore suggest that the cost of 5-ASAs is based on the weighted average cost of all drugs within a class be used.

1.9 The cost of dying is not incorporated into the model.

AbbVie notes that no costs have been allocated to the death state in the AG model.

It would be reasonable to assume that some deaths in UC patients occur in hospital and as such, would incur costs to the NHS. This assumption is supported by results from the 2014 IBD audit[6] which showed that of those who were admitted to hospital and subsequently died, the median length of stay was 26 days (Interquartile range: 18, 34). Similar data is not available for those patients who underwent surgery and subsequently died, but it could reasonably be assumed that there is a cost associated with terminal care due to surgery related death in addition to the cost of the surgical procedure. Not incorporating a cost of death due to surgery underestimates the true costs of the treatment options , and lowers the ICER in favour of surgery.

AbbVie therefore suggests the cost of death be incorporated in the Assessment Group model.

1.10 Factual inaccuracies and clarifications

Page #	Text	Suggested text or action
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.
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.
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
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.

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
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 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).
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.
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.[38]. 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
84	Acute infusion reactions occurred in similar numbers of patients in both treatment arms through week 54 of ACT1 (infliximab 5mg/kg 2/121, 9.9%; placebo 13/121, 10.7%).	Acute infusion reactions occurred in similar numbers of patients in both treatment arms through 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.
219	Table 65 Health state utilities for all presurgical and post-surgical states Woehl *et al*[109]	
225	Of the 10 potentially relevant EQ-5D studies, those reported by Woehl *et a*[109] …	
225	Woehl et al[109] collected EQ-5D utility scores from 180 patients with active UC.	
227	The utility score for post-surgery was based on the mean value reported by Woehl et al[109] …	
245	Table 85 Source of health utilities	The incorrect reference number is assigned to Woehl in the text.

	Woehl et al[160] (chronic pouchitis valued using Arseneau et al110)	The correct reference is 162.
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.**
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%). In the adalimumab arm, 154 (62%) patients completed the study out of 248 who were randomised.

1.11 Miscellaneous

Concluding statements regarding adalimumab's efficacy relative to placebo

The report concludes (Page 253 of the Assessment Report) that "For patients classified as being in remission at the end of the induction phase, all treatments except for adalimumab were associated with beneficial effects relative to placebo." The conclusion that adalimumab is not associated with beneficial effects relative to placebo for maintenance of remission is directly contradicting what was observed in the ULTRA 2 randomised adalimumab clinical trial.[39] In the NMA results none of the anti-TNFs are significantly different than placebo for maintenance of remission from 8-32 weeks. AbbVie considers that given the difference in conclusion reached in this analysis versus those for each of the maintenance trials further exploration of these results and the validity of this conclusion is warranted.

Section 2. Response to the Assessment Group's Critique of the AbbVie Model

2.1 Critique 1: Deviations from the NICE Reference Case and final NICE scope (with respect to omission of other anti-TNFs and surgery as comparators and model time horizon

2.1.1 Time horizon

In its critique of the original AbbVie submission model, the Assessment Group notes that AbbVie has constrained the model to a 10-year time period. Following suggestions from the Assessment Group, the submission model has been extended to include a lifetime time horizon.

2.1.2 Choice of comparator - Surgery

The report also brings into question the exclusion of surgical treatment as a comparator to adalimumab in the original submission model. Given the early onset of UC (patients are typically diagnosed between the ages of 15 and 25), the potential for complications associated with surgical intervention, and the life-altering implications of UC-related surgery, surgery is still widely viewed as salvage therapy and many patients are not comfortable with undergoing such a procedure.[40] For these reasons, surgery was not included as a comparator to adalimumab in the initial AbbVie submission model.

To address this critique, a separate model has been developed which compares treatment with adalimumab versus surgery. In this model, two alternative treatment options are compared; one option is patients would directly receive surgery, and the other option is patients would initiate on anti-TNFs first followed by surgery treatment upon failure on anti-TNFs. In this mode, moderate to severe UC patients who relapse on treatment with adalimumab will switch directly to surgery rather than conventional non-anti-TNF therapy.

2.1.3 Choice of comparators – infliximab and golimumab

The Assessment Group suggests that infliximab and golimumab should be included as comparators to adalimumab in the AbbVie submission model.

Whilst AbbVie disagrees with the approach to base the evidence of anti-TNFs and conventional treatment on a NMA, we accept that this is the most pragmatic solution to determine the efficacy between treatment options relevant to the decision problem.

We do however want to reiterate that the reliability of the estimates obtained in a NMA is dependent upon similarity between the trial design, setting and the endpoints being compared. As stated in the submission, there are substantial differences in study design and alternative scoring methods across trials for evaluating treatment efficacy which could prevent a meaningful comparison across these anti-TNFs using an NMA methodology. Furthermore, we want to emphasise that a NMA utilising the proportion of patients achieving remission or response from each trial would inherently assume that these endpoints have been defined in exactly the same manner across trials. AbbVie also believes that the NMA results used in the Assessment Group model will negatively impact on the efficacy estimates for adalimumab for the following three primary reasons:

2.1.3.1 A major difference across the clinical trials of adalimumab, golimumab, and infliximab was how Mayo scores were determined; as detailed in the submission document, the worst sub-scores of

stool frequency and rectal bleeding from the past 3 days were used to calculate the Mayo scores in the adalimumab trials, while average scores were used in the infliximab and golimumab trials. On page 211 of the Assessment Report, it is suggested that "fluctuations in Mayo score evaluations over the course of 3 days are likely to be minor." Given the stringency of the Mayo score definition of remission, whereby a one-point change in overall score can mean the difference between achievement of remission or not, AbbVie believes this statement underestimates the magnitude of impact Mayo scoring methodology can have on clinical trial endpoints. Due to the waxing and waning nature of UC symptoms, use of worst-rank methodology, which was pre-specified in the adalimumab trials, will tend to underestimate differences between patients receiving an effective drug and those receiving placebo. Indeed, data from the M10-477 clinical trial, in which daily patient diaries were available to allow different post-hoc methods of Mayo score calculations show that an 'average scoring method' led to higher rates of remission, more so for the adalimumab treated patients than placebo treated patients, compared to the 'worst sub-score method' of stool frequency and rectal bleeding over a 3-day period, and the average scoring method led to substantially larger treatment effect sizes than the worst-scoring method. Remission rates calculated using the worst sub-score method for placebo, adalimumab (80/40mg), and adalimumab (160/80mg) were 11.5%, 13.8%, and 10.0% respectively, whereas remission rates calculated using the average scoring method for placebo, adalimumab (80/40mg), and adalimumab (160/80mg) were 13.5%, 18.4%, and 14.4% respectively.[41] The difference in patients achieving remission across these two methods was 2.0% for the placebo arm, 4.6% for adalimumab 80/40mg, and 4.4% for adalimumab 160/80 mg. Therefore, evaluating treatment efficacy using Mayo scores calculated from the worst sub-score method may lead to a decreased treatment efficacy for adalimumab compared to a scoring system based on average Mayo scores. Additionally, the decrease in efficacy is more pronounced among patients treated with adalimumab compared to placebo, which results in lower relative effect comparing adalimumab to placebo using the worst scoring method than when using the average scoring method. The odds ratios compared to placebo were 1.23 for adalimumab (80/40mg) and 0.86 for adalimumab (160/80mg) using the worst scoring method. The odds ratios of remission compared to placebo were 1.44 for adalimumab (80/40mg) and 1.08 for adalimumab (160/80mg) using the average scoring method, which are both higher than the estimated odds ratios when using the worst scoring method. As a result, this would cause a decreased relative effect of adalimumab versus other anti-TNFs in an indirect comparison using placebo as the common comparator. Unfortunately, due to lack of access to individual patient diary records in the ULTRA2 trial these scores cannot to be re-analysed to calculate a Mayo score using average scoring method.

2.1.3.2 The model employs a definition of non-response which bases the efficacy of adalimumab on NRI, a method which categorises all patients who dose-escalate on adalimumab as non-responders. As discussed in Section 1, this would lead to an underestimate of adalimumab efficacy. The lack of consideration for the additional efficacy associated with dose escalation is not a concern for other anti-TNFs within the evaluation.

2.1.3.3 Heterogeneity in adalimumab, infliximab and golimumab trial designs could generate biased results in an indirect comparison of these anti-TNFs. As acknowledged in the report, the golimumab trial has a different design for assessing maintenance than the adalimumab and infliximab trials. In the golimumab maintenance trial, initial responders to golimumab were re-randomised to either placebo or golimumab for the assessment of maintenance. Thus, the placebo efficacy during the maintenance period was measured among patients who were responders to anti-TNFs. In the adalimumab and infliximab trials, patients are randomised only at baseline and followed for the entire study period. Thus, in the adalimumab and infliximab trials, the placebo efficacy is evaluated among placebo responders, while the golimumab trial is among initial golimumab responders who may or may not respond to placebo (i.e. patients in the adalimumab trial remained on placebo, a treatment they had already responded to; patients in the golimumab trial switched from golimumab to placebo and were being removed from a treatment to which they had already responded). The NMA has estimated the placebo efficacy during the maintenance period in the adalimumab and infliximab trials using placebo responders from the induction period. However, patients responding to golimumab (in the golimumab trial) and patients responding to placebo (in the adalimumab and infliximab trials) could be inherently different, which could prevent a reliable indirect comparison across anti-TNFs during the maintenance period using placebo as the common comparator.

Given the combined limitations mentioned above, we feel that the efficacy of adalimumab versus infliximab and golimumab cannot be fully considered using an NMA framework.

2.2 Critique 2: Questionable choice of cycle length necessitating the use of other external evidence on transition probabilities which should not be required

In its critique, the Assessment Group states that the use of a 2-week cycle in the original AbbVie submission model could lead to less accurate results compared with a longer cycle length. A 2-week cycle was used in the submission model to reflect the every-other-week dosing schedule of adalimumab. We believe that the use of a 2-week cycle length allows for additional flexibility in the model and a more accurate representation of patients' dosing and discontinuation patterns in UK clinical practice.

In addition, the Assessment Group suggests in its critique that using a longer cycle length of 6 weeks would eliminate the need for external data to inform treatment efficacy for adalimumab's induction period. However, the use of external data has little impact on the model results when determining treatment efficacy during the induction period.

To evaluate the impact of such external data, a sensitivity analysis has been implemented with the assumption that patients do not change disease states during the first 8 weeks of treatment with adalimumab, and instead can only change disease states after week 8. This sensitivity analysis is using the same assumption as in the Assessment Group's model, where patients remain in the same disease state for the first 8 weeks of treatment. The results from this sensitivity analysis revealed that the ICER is not sensitive to this change, and the ICER of adalimumab compared to conventional therapy increased by £89, and the ICER of adalimumab compared to surgery increased by by £88 to £23,115 per QALY gained, and the ICER of adalimumab compared to surgery increases by £239 to £15,989 per QALY gained after removing the use of external data in the sensitivity analysis.

2.3 Critique 3: Concerns regarding the selection and use of evidence to inform HRQoL parameters

2.3.1 Mapping utilities from the study to the EQ5D

The report states that data on the SF-36 Health Survey were collected in the ULTRA2 trial and should have been used in the original submission model to derive SF-6D HRQoL parameters. However, the NICE Guide to the Methods of Technology Appraisal emphasises that EQ-5D is the preferred instrument for QALY impact assessments in adult populations[42]. AbbVie would like to highlight that mapping from SF-6D to EQ-5D, as stated in the submission document (Appendix I Section 2.9 Utilities of the submission document), would likely overestimate HRQoL for patients with more severe disease states. Furthermore, the available trial-based SF-6D results were only available for pre-surgery states and were not available for post-surgery states.

2.3.2 Choice of utilities for health states

In the original AbbVie submission, pre-surgery utilities were derived from Swinburn et al.[43] The report notes that Swinburn et al. has only been published in abstract and poster form, although additional detail was provided to the Assessment Group in the submission report. Swinburn et al. was chosen as the source for health related quality of life (HRQoL) estimates in the AbbVie submission model for several reasons: 1) HRQoL estimates are reported for patients in remission, mild, and moderate-severe disease states, the same set of pre-surgery disease states used in the AbbVie original submission model, 3) Swinburn et al. contained a large sample size, 4) Swinburn et al. reports details on the characteristics of patients included in the study, and 5) Swinburn et al. studied a geographically diverse patient population in the UK.

In terms of post-surgery utility, the AbbVie model incorporates the utility of post-surgery without complication and post-surgery with transient complication reported in Tsai et al.[44] For utilities after surgery without complication, AbbVie acknowledges that Tsai et al. contains limited information on how utility levels were derived. However, the post-surgery utility reported in Tsai et al. (0.61) is similar to the post-surgery utility reported in Swinburn (0.59)

2.3.3 Utility values for chronic complications

Finally, the critique asserts that the derivation of utilities for chronic complications in the original submission document was unclear, and that different health instruments were used to inform utility values for chronic complications. The Assessment Group has also acknowledged that the utility value for post-surgical chronic complications does not have a material impact on the ICER. To clarify the calculation of these estimates, and to address any concerns on the use of different health instruments, the utility value for post-chronic complications has been updated in the updated submission model and additional details have been provided below regarding these utility calculations.

The utility value for the chronic complications state was calculated from utility values of chronic pouchitis, infertility, and sexual dysfunction, weighted based on the respective rates. The utility for chronic pouchitis was 0.40 as reported in Arseneau et al 2006[45]. To estimate the utility for chronic pouchitis, the Assessment Group has taken the difference in utility values between ileostomy and chronic pouchitis reported in the same reference, and applied the estimated disutility to the utility of post-surgery state used in the Assessment Group model. Since it is not clear in Arseneau et al. whether ileostomy refer to a state of surgery without complication, and since the only difference between this state and chronic pouchitis is the

presence of chronic pouchitis, the disutility approach is not used in the AbbVie models; rather the direct utility reported was used in the AbbVie submission models (0.40).

The utilities for infertility and sexual dysfunction were calculated by subtracting the disutility values for each complication from the utility value of 0.61 for the post-surgery without complication state used in the AbbVie original and updated submission model. The disutility of infertility was 0.18 as reported in Hu 2009[46]. The disutility of sexual dysfunction was estimated at 0.07, based on the difference in utility value between states with no sexual impairment and with sexual impairment (i.e. 0.96-0.89; Smith et al. 2002)[47].

The respective rate for each chronic complications was 10.25% for chronic pouchitis (Abdelrazep et al. 2008)[48], 7.22% for infertility (Targownik et al. 2012; Johnson et al. 2004)[4950], and 2.74% for sexual dysfunction (Krausz et al. 2005)[51].

The overall utility for chronic complication state was calculated as the weighted average of all three complications, estimated at 0.43.

AbbVie acknowledges that different health instruments are used to derive utilities of chronic complications (time trade off method for chronic pouchitis and sexual dysfunction, and Health Utility Index for infertility). However, these are based on the best available evidence.

2.4 Critique 4: Questionable source of surgery rate

The assessment report states that Hillson et al. is a US-based study with a shorter 1-year study period, and that the population of interest in Hillson et al. did not relate to moderate to severe UC patients.

We acknowledge that Hillson et al. is a US-based study which might not reflect the exact surgical rate among a UC patient population in the UK[9]. However, we feel that the additional critiques raised regarding the time horizon and study population in Hillson et al. do not apply. Hillson et al. contains a proxy measure which classifies patients into moderate or severe disease categories, categories which are utilised in the AbbVie submission model to estimate the probability of undergoing surgery among moderate to severe UC patients. Additionally, Hillson et al. is a cross-sectional evaluation, and the estimated results reflect average probabilities of surgery among moderate to severe patients. The patient population should include those who were in moderate to severe disease states for a long period, and patients who were in moderate to severe disease states for a short period, thus the rate reported already took into consideration of surgery might occur over a long period after patients were in moderate to severe disease states.

Section 3. AbbVie Model Updates and Results in Response to Critiques

Two separate models have been developed to update the original AbbVie submission model in response to the Assessment Group's critiques:

1. An updated submission model comparing adalimumab to conventional medical therapy. In this model, the majority of patients who relapse while on treatment with adalimumab will switch to conventional therapy and a small proportion of patients will switch to surgery based on the surgery rate observed in the clinical papers discussed in Section 2.

2. An updated submission model comparing surgery to treatment with adalimumab. In this model, all patients who relapse while on treatment with adalimumab after the first 8 weeks (induction period) will switch directly to surgery. Results from this model will be used to inform whether patients should undergo surgery directly or receive treatment with anti-TNFs prior to surgery.

Both models have been extended to consider a lifetime time horizon, and general population mortality has been added to the model framework due to the extension to lifetime horizon. The same life table for natural mortality used by the Assessment Group has been incorporated into AbbVie's updated submission models.

In the base case for the updated submission models, the efficacy of adalimumab and conventional therapy is derived from the anti-TNF-α naïve subgroups in the ULTRA2 and ULTRA3 (the open label extension study of ULTRA 1 and 2) trials; the anti-TNF-α naïve subgroup is used as the base case population to be consistent with the base case population in the Assessment Group model. Deterministic sensitivity analyses are provided to use efficacy results from the ITT population of adalimumab trials, including both anti-TNF-α naïve and experienced patients. ICER estimates based on the ITT population are similar to those based on the anti-TNF-α naïve subgroup for the comparison of adalimumab to conventional therapy (£22,809 and £23,027 per QALY gained, respectively) and for the comparison of adalimumab compared to surgery (£15,434 and £15,750 per QALY gained, respectively). Unless otherwise specified, the ICER values presented in this document are based on the anti-TNF naïve subgroup.

The rate of surgery has also been revised in the updated submission models to reflect a sample-weighted average of the rates of surgery in Actis et al., Molnar et al., Mocciaro et al, and Gustavsson et al. These publications are from the list of literature summarized in Table 67 of the Assessment Report

Additionally, the discontinuation rule for adalimumab after year 2 has been changed in the updated submission model comparing adalimumab to conventional therapy. In the original submission model, the discontinuation rate from the ULTRA3 trial was used; in the updated submission model, adalimumab-treated patients who relapse to a moderate-to-severe UC disease state after year 2 will discontinue treatment (the same rule for discontinuation utilized by the Assessment Group). The transition probabilities and dose intensity of adalimumab after year 2 have been re-calculated based on this new discontinuation rule (i.e., patients who relapse to a moderate-to-severe UC disease state will not contribute to the transition probabilities and dose intensity estimation after their relapse in disease severity).

In the new models, the base case ICER was £23,027/QALY comparing adalimumab to conventional non-anti-TNF therapy and £15,750/QALY comparing adalimumab to surgery at life time among moderate and severe UC patients who failed conventional non-anti-TNF therapy, respectively.

Several deterministic sensitivities have also been incorporated into the updated submission models in response to critiques from the Assessment Group:

1. To evaluate the impact of using external data to inform transition probabilities during the induction period (discussed in Section 2.2) a sensitivity analysis has been implemented excluding the use of external data. The results from this sensitivity analysis reveal that the ICER is not sensitive to the use of this external data, generating an ICER of £23,115 per QALY gained comparing adalimumab to conventional therapy, and an ICER of £15,989 comparing treatment with adalimumab followed by surgery to direct surgery. .

2. A sensitivity analysis has been added to the AbbVie submission models which uses the same set of pre-surgery and post-surgery utilities relied upon by the Assessment Group model (Woehl et al.) Results from this sensitivity show that incorporating these post-surgery utilities does have an impact on the ICER for adalimumab compared to conventional therapy and surgery, however, it does not change the economic conclusions that should be drawn based on these ICERs. The ICER of adalimumab compared to conventional therapy increased to £25,470 per QALY gained, and the ICER of adalimumab compared to surgery increases to £21,719 per QALY gained.

3. A sensitivity analysis has been added to the AbbVie submission models where the utility of post-surgery chronic complications has been changed to reflect the value used in the Assessment Group model: a value of 0.54 (post-surgery, without complications utility of 0.71 minus the disutility due to chronic pouchitis of 0.17), has been incorporated in the updated AbbVie models. Results from this sensitivity analysis show that this modification has a minimal effect on the ICER per QALY gained (an increase of £625 to £23,652/QALY comparing adalimumab to conventional therapy and an increase of £621 to £16,371/QALY for adalimumab compared to surgery).

4. To evaluate the impact of using a different rate of surgery in the updated AbbVie submission model, an alternative surgery rate was calculated based on the publications summarised in Table 67 of the Assessment Report. Table 67 included the following 6 studies: Actis et al.,[52] Gower-Rousseau et al,[53] Molnar et al,[54] Mocciaro et al,[4] Gustavsson et al.,[55] and Solberg et al.[7] Among them, Gower-Rousseau et al. was excluded because it focused on paediatric UC patients rather than adult UC patients. Solberg et al. was excluded based on the reasons described above. Using the remaining 4 studies (excluding the steroid responder arm in Molnar since it deviates from the evaluation population), a surgery rate was calculated based on an average weighted by study sample size, and this alternative average rate of surgery was incorporated into the updated model. To evaluate the impact of this input on ICER, a sensitivity analysis is conducted by including Solberg et al. as one of the data sources in the surgery rate estimate in the updated AbbVie submission model comparing adalimumab to conventional treatment and to surgery. After including Solberg et al. as a data source in the surgery rate estimate, the ICER increases by £2,004 to £25,030/ QALY. This sensitivity analysis is not evaluated in the updated submission model comparing adalimumab to surgery since in that model all adalimumab-treated patients who relapse while on treatment with adalimumab will switch directly to surgery after the induction period of 8 weeks.

The base case ICERS are robust in deterministic sensitivity analyses, with ICER ranging from £18,218 to £25,594 per QALY gained comparing adalimumab to conventional therapy, and from £13,172 to £21,719 per QALY gained comparing adalimumab to surgery. The probability that adalimumab is cost-effective at the

willingness to pay threshold of £30,000 per QALY gained is 73% compared to conventional therapy, and 94% compared to surgery.

References

[1] Feagan, BG et al. Adalimumab therapy is associated with reduced risk of hospitalization in patients with ulcerative colitis. Gastroenterology 2014; 146(1):110-118

[2] AbbVie. Summary of Product Characteristics. Available from: https://www.medicines.org.uk/emc/medicine/21201 [Accessed: 31 July 2014]

[3] Wolf D et al. Escalation to weekly dosing recaptures response in adalimumab-treated patients with moderately to severely active ulcerative colitis. Aliment Pharmacol Ther. 2014. EPub ahead of print

[4] Mocciaro et al. Cyclosporine or infliximab as rescue therapy in severe refractory ulcerative colitis: early and long-term data from a retrospective observational study. Journal of Crohn's & Colitis 2012; 6(6):681-686

[5] Mowat C et al. Guidelines for the management of inflammatory bowel disease in adults. Gut 2011; 60(5): 571-607

[6] Royal College of Physicians. National clinical audit of inpatient care for adults with ulcerative colitis: UK inflammatory bowel disease (IBD) audit. June 2014. Available from: https://www.rcplondon.ac.uk/sites/default/files/adult_ibd_inpatient_care_audit_web.pdf [Accessed: 24 July 2014]

[7] Solberg IC et al. Clinical course during the first 10 years of ulcerative colitis: results from a population-based inception cohort (IBSEN study). Scan J Gastroenterol 2009; 44: 431-440

[8] Sandborn WJ et al. Adalimumab induces and maintains clinical remission in patients with moderate-to-severe ulcerative colitis. Gastroenterology. 2012;142:257-265

[9] Hillson, E et al. Health care expenditures in ulcerative colitis: the perspective of a self-insured employer. J Occup Environ Med 2008; 50(8):969-977

[10] Turner D et al. C-reactive protein (CROP), erythrocyte sedimentation rate (ESR) or both? A systematic evaluation in pediatric ulcerative colitis. Journal of Crohn's and Colitis. 2011; 5:423-429

[11] Osada T et al. Correlations among total colonoscopic findings, clinical symptoms, and laboratory markers in ulcerative colitis. Gastroenterology. 2008;23 (Suppl 2):S262-S267

[12] Manz M et al. Treatment algorithm for moderate to severe ulcerative colitis. Swiss Medical Weekly. 2011; 141:w13235

[13] Arai, K et al. Incidence and outcome of complications following restorative proctocolectomy. Am J Surg 2005; 190(1):39-42

[14] Johnson P et al. Female infertility after ileal pouch-anal anastomosis for ulcerative colitis. Dis Colon Rectum. 2004;47:1119-1126

[15] Gorgun E et al. Fertility is reduced after restorative proctocolectomy with ileal pouch anal anastomosis: a study of 300 patients. Surgery. 2004;136:795-803

[16] Lepisto A et al. Female fertility and childbirth after ileal pouch-anal anastomosis for ulcerative colitis. British Journal of Surgery. 2007;94:478-482

[17] Berndtsson I, Oresland T, Hulten L. Sexuality in patients with ulcerative colitis before and after restorative proctocolectomy: a prospective study. Scand J Gastroenterol. 2004;4:374-379

[18] Farouk K et al. Functional outcomes after ileal pouch-anal anastomosis for chronic ulcerative colitis. Annals of Surgery. 2000;231(6):919-926

[19] Lindsey I et al. Impotence after mesorectal and close rectal dissection for inflammatory bowel disease. Dis Colon Rectum. 2001; 44(6): 831-835

[20] Mowat C et al. Guidelines for the management of inflammatory bowel disease in adults. Gut 2011; 60(5): 571-607

[21] Hueting WE et al. Results and complications after ileal pouch anal anastomosis: A meta-analysis of 43 observational studies comprising 9,317 patients. Dig Surg 2005; 22: 69-79

[22] de Zeeuw S et al. Update of complications and functional outcome of the ileo-pouch anal anastomosis: overview of evidence and meta-analysis of 96 observational studies. Int J Colorectal Dis 2012; 27: 843-853

[23] Burns EM et al. Volume analysis of outcome following restorative proctolectomy. Br J Surg 2011; 98: 408-417

[24] Leonard et al. Complications associated with coelcteomy in patients with ulcerative colits. Poster presented at Digestive Disease Week, May 19-22; 2012

[25] Bach SP et al. Revolution and Evolution: 30 Years of Ileoanal Pouch Surgery. Inflamm Bowel Dis 2006; 12: 131-145

[26] Crohn's and Colitis UK. Information Sheet - Surgery for ulcerative colitis (UC). Available from: http://www.crohnsandcolitis.org.uk/Resources/CrohnsAndColitisUK/Documents/Publications/Info-Sheets/Surgery%20for%20Ulcerative%20Colitis.pdf [Accessed: 17 July 2014]

[27] NHS Choices. Ileostomy – how it's performed. Available from: http://www.nhs.uk/Conditions/Ileostomy/Pages/How-is-it-performed.aspx [Accessed: 28 July 2014]

[28] Tsai HH et al. A model of the long-term cost effectiveness of scheduled maintenance treatment with infliximab for moderate-to-severe ulcerative colitis. Aliment Pharmacol Ther 2008; 28: 1230-1239

[29] Buchanan J et al. Managing the long-term care of inflammatory bowel disease in patients: The cost of European health providers. J Crohn's Colitis 2011: 5: 301-316

[30] NHS Northamptonshire. NHS Northamptonshire Comprehensive Stoma Care Guidelines – Information for general practitioners. Available from: http://www.neneccg.nhs.uk/resources/uploads/files/Stoma%20Care%20Guidelines.pdf [Accessed: 24 July 2014]

[31] Hull and East Riding Prescribing Committee. Guidelines for the prescribing of stoma care appliances in primary care. Available from: http://www.hey.nhs.uk/herpc/guidelines/StomaCareGuidance.pdf [Accessed: 24 July 2014]

[32] Ipswich and East Suffolk Clinical Commissioning Group. Cost effective prescribing of ostomy products. Available from: http://www.ipswichandeastsuffolkccg.nhs.uk/Portals/1/Content/Members%20Area/Clinical%20Area/Medicine%20management/Guidelines%20and%20Policies/ostomy%20and%20continence%20products%20final.pdf [Accessed 24 July 2014]

[33] NHS Business Services Authority. New arrangements under Part IX of the Drug Tariff for the provision of stoma and urology appliances. Available from: http://www.nhsbsa.nhs.uk/PrescriptionServices/2818.aspx [Accessed 24 July 2014]

[34] NHS Electronic Drug Tariff. Part IXC – Stoma appliance. Available from: http://www.ppa.org.uk/edt/January_2014/mindex.htm [Accessed on 24 July 2014]

[35] Feagan, BG et al. Adalimumab therapy is associated with reduced risk of hospitalization in patients with ulcerative colitis. Gastroenterology 2014; 146(1):110-118

[36] BMJ Group and Pharmaceutical Press. British National Formulary. 2014. July edition. Available from: https://www-medicinescomplete-com.proxy1.athensams.net/mc/bnf/current/index.htm [Accessed: 21 July 2014]

[37] Dignass A et al. Second European evidence-based consensus on the diagnosis and management of ulcerative colitis Part 2: Current management. J Crohn's Colitis 2012; 6: 991-1030

[38] Colombel J-F et al. Adalimumab Maintains Remission for up to 4 Years in Patients with Ulcerative Colitis. 2014; Poster presented at the 9th ECCO Congress in Copenhagen, February 20-22, 2014 (P571)

[39] Sandborn WJ et al. 2012. Adalimumab induces and maintains clinical remission in patients with moderate-to-severe ulcerative colitis. Gastroenterology 2012; 142:257–265

[40] Bewtra M et al. Patient preferences for surgical versus medical therapy for ulcerative colitis. Inflamm Bowel Dis. 2014 Jan; 20(1):103-14.

[41] Suzuki Y et al. Efficacy and safety of adalimumab in Japanese patients with moderately to severely active ulcerative colitis. J Gastroenterol. 2014;49:283-294

[42] National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. Available from https://www.nice.org.uk/article/PMG9/chapter/5-The-reference-case [Accessed 30 July 2014]

[43] Swinburn et al. PTU-127 The impact of surgery on health related quality of life in ulcerative colitis. Gut. 2012; 61:A237

[44] Tsai HH et al. A model of the long-term cost effectiveness of scheduled maintenance treatment with infliximab for moderate-to severe ulcerative colitis. Alimentary Pharmacology & Therapeutics 2008; 28(10):1230-1239.

[45] Arseneau KO et al. Do patient preferences influence decisions on treatment for patients with steroid-refractory ulcerative colitis? Clin Gastroenterol Hepatol. 2006;4(9):1135-1142

[46] Hu D et al. Cost-effectiveness analysis of alternative first-trimester pregnancy termination strategies in Mexico City. BJOG : an international journal of obstetrics and gynaecology 2009;116(6):768-779.

[47] Smith DS et al. Patient preferences for outcomes associated with surgical management of prostate cancer. J Urol. 2002;167(5):2117-2122

[48] Abdelrazeq et al. Predictors for acute and chronic pouchitis following restorative proctocolectomy for ulcerative colitis. Colorectal Dis 2008; 10(8):805-813

[49] Targownik LE et al. The epidemiology of colectomy in ulcerative colitis: results from a population-based cohort. Am J Gastroenterol. 2012;107(8):1228-1235

[50] Johnson P et al. Female infertility after ileal pouch-anal anastomosis for ulcerative colitis. Dis Colon Rec2004;47(7):1119-1126

[51] Krausz MM, Duek SD. Restorative proctocolectomy with ileal pouch-anal anastomosis for ulcerative colitis and familial adenomatous polyposis: twenty years follow-up in 174 patients. Isr Med Assoc J. 2005;7(1):23-27

[52] Actis, G.C et al. Colectomy rate in steroid-refractory colitis initially responsive to cyclosporin: a long-term retrospective cohort study. BMC Gastroenterology 2007; 7

[53] Gower-Rousseau, C et al. The natural history of pediatric ulcerative colitis: a population-based cohort study. Am J Gastroenterol 2009; 104(8):2080-2088

[54] Molnar, T et al. Response to first intravenous steroid therapy determines the subsequent risk of colectomy in ulcerative colitis patients. J Gastrointestin Liver Dis 2011; 20(4):359-363

[55] Gustavsson, A et al. Clinical trial: colectomy after rescue therapy in ulcerative colitis - 3-year follow up of the Swedish-Danish controlled infliximab study. Aliment Pharmacol Therap 2010; 32(8):984-989