Input from manufacturer on the 2nd draft assessment "USTEKINUMAB FOR THE TREATMENT OF ADULT PATIENTS WITH MODERATELY TO SEVERELY ACTIVE ULCERATIVE COLITIS (UC) WHO HAVE HAD AN INADEQUATE RESPONSE WITH, LOST RESPONSE TO, OR WERE INTOLERANT TO EITHER CONVENTIONAL THERAPY, A BIOLOGIC, OR HAVE MEDICAL CONTRAINDICATIONS TO SUCH THERAPIES"

Project ID: PTJA07

eunethta

EUROPEAN NETWORK FOR HEALTH TECHNOLOGY ASSESSMENT

Comments on the 2nd draft rapid assessment on ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (uc) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies

The objective of this reviewer form is to standardise the process of reviewing rapid relative effectiveness assessments.

The 2nd version of the Rapid Assessment on ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (uc) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies was open to review by the manufacturer [Janssen] between **07/10/2019 and 11/10/2019**.

Comments received from:

Market Authorisation Holder
Janssen

All received comments are formally responded in this combined document, to be published on the EUnetHTA website, name of organisation/institution (or individual names of the reviewers/affiliations) disclosed.

Comments on the 2nd draft rapid assessment on ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (uc) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies

Comments from Market Authorisation Holder [Janssen]

Page	Line	Comment	Character of comment[i]	Reply from author
15	8-9	Text from EUnetHTA's assessment: "Active treatment maintenance endpoint data were also imputed for induction non-responders in 8 GEMINI I and induction responders in OCTAVE trials." Comment Janssen: The author states that active treatment maintenance endpoint data were imputed for the induction responders of OCTAVE. This is incorrect as published data from the tofacitinib FDA report (1) were used to inform maintenance outcome data for the active arms. Imputation for this study were only needed for the placebo arms (as described in section 5.3.2.1 of the submission).	2	The MAH was asked to check for factual accuracy of the document. This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
15	10-12	Text from EUnetHTA's assessment: "For the maintenance treatment arms in non-biologic failure patients, the MAH did not find statistically significant dose response relationship for the treatments with different dosage regimes and 11 concluded that pooling of the doses for the same treatment was appropriate." Comment Janssen: "statistical significance" referred to for the dose response relationship but no statistical tests were conducted as explained in the submission section 5.3.2.2.	2	The text was changed into: For NBF patients in the maintenance treatment arms, the MAH did not identified dose response relationship (higher dose/shorter interval between doses did not lead to higher clinical responses) and concluded that pooling of doses for the same treatment was appropriate. No formal dose-response relationship testing was done.
15	49-50	Text from EUnetHTA's assessment: "[...] three studies related to 49 adalimumab (ULTRA 1; ULTRA 2; NCT00853099 Suzuki et al, 2014);[...]" Comment Janssen: Did not include the trials related to vedolizumab (GEMINI and VARSITY) and missed VARSITY as a study for adalimumab – this should be corrected to be "four" instead of "three" studies for adalimumab.	2	Thank you, we made changes needed.

Comments on the 2nd draft rapid assessment on ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (uc) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies

Page	Line	Comment	Character of comment[i]	Reply from author
16	4-5	Studies listed to be included in the induction NMAs in brackets is missing the NCT00787202 trial.	2	Thank you, we made change needed.
20	Paragraph 6 – under indirect evidence	Text from EUnetHTA's assessment: "The evidence drawn from all the networks was based on scarce, indirect data in a setting where heterogeneity or inconsistency of a network could not be statistically assessed, and assessment of statistical heterogeneity between trials informing each comparison was limited and likely underpowered." Comment Janssen: These statements cannot be made to all networks together as there are clear differences highlighted in the discussion section of the submission to show that the evidence networks for 1-year non-biologic failure patients was more robust. This network included a loop for clinical remission and mucosal healing (therefore not relying solely on indirect data for these connections), pooled treatment arms (increasing the power in the analyses) and were based on more homogenous patient populations. The authors also acknowledge in the discussion section of the report (p.145) that the comparisons to anti-TNFs were robust and that minimal or no imputations were needed in these cases. This is contradictory to comment made in the report regarding the uncertainty and low level of evidence of the NMAs. It should be specified that not all treatment comparisons from networks are subject to the same limitations.	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
20	Paragraph 6 – under indirect evidence [Also p. 23 paragraph 2]	Text from EUnetHTA's assessment: "Even in the networks which included a head-to-head comparison, due to data imputation in one of the trails informing a closed loop inconsistency assessment was not feasible." Comment Janssen: It was feasible to assess inconsistency as although maintenance outcomes were imputed for induction non-responders receiving placebo in GEMINI, the data can still be considered as "direct" as opposed to "indirect", which would technically refer to between study treatment comparisons. Therefore, it is incorrect to say that inconsistency could not be assessed within this loop.	2	The text was changed into: "Even in the networks which included a head-to-head comparison, due to data imputation in one of the trails informing a closed loop, inconsistency assessment between different sources of evidence was not performed."

Comments on the 2nd draft rapid assessment on ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (uc) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies

Page	Line	Comment	Character of comment[i]	Reply from author
20	Paragraph 8 – under indirect evidence	Text from EUnetHTA's assessment: "Credible intervals (95% CrI) from most outcomes were quite wide, with a few comparisons showing statistical signifi-cance of usekinumab versus another active treatment." Comment Janssen: The credible intervals have been used to determine the significance of results in the EUnetHTA assessment report. This is incorrect to do as the analysis was performed in a Bayesian and not Frequentist framework, which has a different interpretation. • Statistical significance is a concept from frequentist statistics (i.e. if the probability for observing results as observed under the null-hypothesis (reflecting no difference in treatment effects) is lower than a pre-specified threshold (alpha, traditionally 5%), null hypothesis is withdrawn, and results are considered to be statistically significant). • NMA results were generated in a Bayesian framework, which rather provides a probability distribution (posterior distribution) on most likely values for a particular parameter, provided all available evidence. Neither the EUNETHTA guidelines (2) or the Cochrane Handbook for Systematic Reviews of Interventions (3) mention this approach of frequentist interpretation of Bayesian results. In applying in such context an arbitrary cutoff of P(OR>1)=97.5% (by simply copying the typical frequentist 2 sided test with alpha=0.05) and ignoring any further evidence on comparative efficacy estimates, this may lead to suboptimal decision making, in a context where all compared treatments are available to patients and their treating physicians, and a treatment decision has to be made anyhow. In such context, giving preference to treatments with high probabilities for being better than any of the comparators (based either on SUCRA or pairwise comparisons) by definition is expected to lead on average to better decision making. This point on the interpretation of Bayesian results was discussed at the face to face scoping meeting with EUNETHTA (held in Diemen, Netherlands on 22nd March 2019), where there was agreement on the approach.	1	Thank you, the text was changed into: The credible interval (CrI) for most outcomes was quite wide, indicating high uncertainties in the results.

Comments on the 2nd draft rapid assessment on ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (uc) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies

Page	Line	Comment	Character of comment[i]	Reply from author
20	Paragraph 8 – under indirect evidence	Text from EUnetHTA's assessment: "In example in the case of NMA results for clinical remission in biologic failure patients at induction, the upper limit of the credible interval for OR reaches the value of 95 which seriously undermines confidence in results of such a model." Comment Janssen: This illustrates that NMA results for biofailure in induction are more uncertain compared to other results: overall, such extreme upper credible interval values are rare in the entire set of results. The OR for UST versus placebo (with upper CrI of 95) is entirely informed by the UNIFI-trial? It illustrates uncertainty about the size of the benefit, but does not question at all whether there was benefit versus placebo, with the lower CrI bound being way above 1 as well.	2	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
20	Paragraph 9 – under indirect evidence	Text from EUnetHTA's assessment: "All the networks presented with largerly uncertain treatment effects, so the probability of a treatment being *better* which is known to underpreform in such cases was not used for interpretation. Comment Janssen: This is not true for all networks, more robust evidence has been presented for the non-biologic failure population 1-year NMAs as per section 5.4.3.2. of the submission, and particularly for the comparisons to anti-TNFs. Additionally, the study data from ULTRA II and GEMINI I used in the NMA was consistent with the study data from VARSITY. Also, it cannot be said that the probability would "under perform" as there is uncertainty for the treatment effect to be higher and lower than the median point estimate (i.e. not in one direction only).	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
20	Paragraph 9 – under indirect evidence	Text from EUnetHTA's assessment: "The SUCRA values were likewise derived from low certainty [...]. " Comment Janssen: It is not correct to refer to all of the evidence presented in the NMA as "low certainty". It has been highlighted in the submission that the non-biologic failure NMA of 1-year outcomes is more robust than the bio failure NMA of 1-year outcomes.	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.

Comments on the 2nd draft rapid assessment on ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (uc) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies

Page	Line	Comment	Character of comment[i]	Reply from author
21	Paragraph 2	Text from EUnetHTA's assessment: "For all of the induction NMAs, after a comparison of the DIC values, the fixed effects model was preferred over the randoms effect. However, the observed DIC values of fixed and random models were very similar and did not provide evidence-base for preference of the fixed effect model. Given there is a good intuitive clinical rationale that heterogeneity may be expected but insufficient studies are available to detect it, the random effects model should have been selected. " Comment Janssen: No evidence on between trial variability shown from our assessment of comparability in the baseline populations of trials (appendix 3) or the placebo rates generated after re-calculating to treat-through arms. Also there was insufficient evidence in a number of the induction NMAs to reliably inform a RE model, so the results from fitting these models were considered to be less informative than FE results. Additionally, in case DIC values are similar, it's good (and common) practice (acknowledged by guidelines) to opt for the more parsimonious model (fixed effect).	2	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
21	Paragraph 3	Text from EUnetHTA's assessment: "Pooling the data from the two OCTAVE trials which was done in NMAs for clinical response and mucosal healing in non-biologic failure, and for clinical remission in biologic failure population; in-creased the risk of bias." Comment Janssen: It is incorrect to say that pooling increased the risk of bias as the trial results and populations for the OCTAVE induction 1 and induction 2 trial arms were similar where these were pooled (clinical remission rates are 26.2% and 21.7% for the tofacitinib arm of induction 1 and 2; 15.5% and 7.7% for the placebo arm of induction 1 and 2). Moreover, for all comparators with the exception of infliximab, both doses can be used according to the EU SmPCs.(4)	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
21	Paragraph 4	Text from EUnetHTA's assessment: "However, estimates were very uncertain and it was not possible to state if the effect size was small, moderate, or high." Comment Janssen: This comment refers to uncertainty however the point estimate of the OR still indicates the most likely value and therefore can still be used to determine the effect size.	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.

Comments on the 2nd draft rapid assessment on ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (uc) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies

Page	Line	Comment	Character of comment[i]	Reply from author
21	Paragraph 5	Text from EUnetHTA's assessment: "Based on the SUCRA values, interventions with ustekinumab positioned better than placebo, and adalimumab 160/80 mg for all endpoints, and than golimumab 200/100 mg for clinical response and mucosal healing". Comment Janssen: This has incorrectly interpreted the SUCRA values as P(OR>1) – the statement above is rather based on P(OR>1) >=97.5%.	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
22	Paragraph 3	Text from EUnetHTA's assessment: "For the comparisons of UST 6 mg/kg to other active treatments, CrI estimates were non-informative." Comment Janssen: In these cases, the lower CrI is acknowledged to be greater than 1 but the difference is considered to be non-informative which is not justified. No reference is given to how an effect size is considered to be highly/moderately harmful or beneficial. It is incorrect to state this difference is non-important without further details.	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
23	Paragraph 2	Text from EUnetHTA's assessment: "Consequently, all the networks were based on scarce indirect data." Comment Janssen: This is not true for clinical remission for both populations and mucosal healing for the non-biologic failure population.	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
23	Paragraph 2	Text from EUnetHTA's assessment: "This was done on a data where it is reasonable to assume heterogeneity." Comment Janssen: Baseline characteristics were similar between trials as shown in Appendix 3. Areas of identified heterogeneity were addressed in the base case NMAs or in sensitivity analyses. Moreover, the assumption for heterogeneity is based on two articles, that have insufficiently been discussed in light of the data submitted in this NMA (see later comments on Ghosn 2016(5) and Su 2007(6)).	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.

Comments on the 2nd draft rapid assessment on ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (uc) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies

Page	Line	Comment	Character of comment[i]	Reply from author
23	Paragraph 3	Text from EUnetHTA's assessment: "This resulted in heavy imputation of data." Comment Janssen: Incorrect to suggest that all 1-year NMAs involved heavy imputation. As described in section 5.3.2.1. of the submission, no imputations were needed for ACT-1, ULTRA-2 and VARSITY, the only imputation needed for PURSUIT and UNIFI was for placebo induction non-responders (with very low rates).It is incorrect and misleading to state that heavy imputations were needed without specifying that this was not the case for the comparisons in the non-biofailure population of ustekinumab, golimumab, infliximab, adalimumab and vedolizumab.	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
23	Paragraph 4	Text from EUnetHTA's assessment: "To increase statistical power in this model, MAH pooled doses for treatment arms as no dose re-sponse relationship was observed." Comment Janssen: For ustekinumab, the reason to pool was that observed response rates were higher for the lower dose, which is counterintuitive, and reason to believe observed differences are due to random variation. This was similar for the comparators. The author states that by pooling the results this increases the sample size which affects the results. Instead this is not expected to bias the point estimates, but to impact the uncertainty only.	2	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
24	Paragraph 2	Text from EUnetHTA's assessment: "While the point estimates for these comparisons were high, ranging from OR of 2.4 to 8.7, for some of the said comparisons the lower boundary of CrI for estimates effect sizes were such that there was a possibility that the true effect was not clinically important." Comment Janssen: The credible interval bounds represent the interval where 95% of the simulations lie for the treatment effects. Given the treatment effects expressed as log odds ratios are normally distributed, the lower and upper end of the CrIs are where fewer simulations occur (i.e. less chance of the true treatment effect being the lower or upper intervals). These are used to provide the measure of uncertainty around the treatment effect and not to show where the treatment effect is most likely to occur (the mean and median of the posterior are used for this). Therefore it is incorrect to base statements on the importance of the treatment effect on the lower CrI alone.	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.

Comments on the 2nd draft rapid assessment on ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (uc) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies

Page	Line	Comment	Character of comment[i]	Reply from author
24	Paragraph 3	Text from EUnetHTA's assessment: "count that the evidence on which the SUCRA rankings are based is of very low certainty, the results could be misleading (Table 0.9)." Comment Janssen: This is not true for all networks, non-bio failure network is robust (reasons provided in section 5.4.3.2 of the submission). Cannot make a statement that this applies to all networks considering the difference in evidence, loop in the network and fewer imputations.	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
24	Pargraph 4	Text from EUnetHTA's assessment: "[...]NMA models used [...]." Comment Janssen: Within the paragraph discussing limitations, this text is unclear what the issue is with the NMA model used. Does it refer to the model code, the applied approach (Bayesian vs. frequentist or fixed versus random?) – Please see responses elsewhere in the response document.	2	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
24	Table 0.8	Table 0.8: First column only mentions the treatment, without specifying the dose received at each time (induction-maintenance)	2	Thank you, we made changes needed.
25	Pargraph 4	Text from EUnetHTA's assessment: "[...] the certainity of estimated effect sizes was low as the reported CrIs ranged from OR ~2 to ≥9." Comment Janssen: "low effect size" is not quantified or clinically justified. An OR of ~2 translates to the odds of clinical remission or response for patients receiving ustekinumab being 2 times that for other treatments and therefore represents a substantial and medically relevant increase in likelihood of the event. Within the same paragraph the frequentist interpretation of the significance of the treatment effect is used which is incorrect for a Bayesian analysis (see comments above).	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.

Comments on the 2nd draft rapid assessment on ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (uc) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies

Page	Line	Comment	Character of comment[i]	Reply from author
25	Pargraph 4, last sentance	Text from EUnetHTA's assessment: "For that reason, similar to what has been discussed for the maintenance NMAs in non-biologic failure patients, the fact that all the point estimates for all the comparisons to active treatments, although not significant, point towards increase in odds of achieving clinical response or clinical remission in patients treated with ustekinumab can not be used to support the absolute UST efficacy." Comment Janssen: It has been acknowledged in the submission that the biologic failure population NMAs are limited. However, it would not be correct to say that all of the same limitations apply to the non-bio failure network and to imply here that the NMA results cannot be used to support the understanding of the efficacy for UST, especially in the non-biofailure group versus 3 anti-TNFs and vedolizumab.	2	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
30	Section about safety comparison NMA-Third line	Text from EUnetHTA's assessment: "These demonstrated that overall ustekinumab 6mg/kg was associated with similar likelihoods of overall adverse events, serious adverse events and infections." Comment Janssen: It should be precised that it is overall infections.	2	Thank you, we made change needed.
31	Paragraph 1	Text from EUnetHTA's assessment: "Publication bias was not assessed". Comment Janssen: This is incorrect as the risk of bias assessment results were provided to EUnetHTA has part of the responses to action point 7 post-submission. This assessment showed no risk or unclear risk for all studies with the only exception being ULTRA II maintenance data showing high risk (which was not included in the NMA).	2	Thank you, we made change needed.

Comments on the 2nd draft rapid assessment on ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (uc) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies

Page	Line	Comment	Character of comment[i]	Reply from author
31/104	Paragraph 6/2	Text from EUnetHTA's assessment: "The probability of a treatment being *better* is known to underperform when treatment effect of a intervention is largerly uncertain as an intervention with the most uncer-tainty can have the largest probability." Comment Janssen: The text cited by the author for this statement is from Chamani 2017(7) (cited on page 104). This publication shows that ranking treatments based on the probability of a treatment being the best [P(BEST)] may give misleading conclusions (especially when based on most imprecise estimate). So the statement in the report from EUnetHTA is not correct, and based on the reference it should read: "treatment being best (among a range of treatments) is known to underperform...". Moreover, in the submission, the probability of a treatment being better than another (pairwise) was provided and not P(BEST). Therefore, this statement should not apply. Comparisons on relative efficacy of active of treatments presented in the submission were based both on SUCRA values and pairwise comparisons of efficacy for all comparators versus ustekinumab. We disagree that both approaches, which provide consistent results in the current case, are prone to bias or "are known to underperform". The references cited in the review cannot be used to backup this statement.	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
111	Paragraph 4-Line 4	OR=1.07, 95% CrI = [0.62;1.87]	1	Thank you, we made change needed.
114	Figure 4.11	TOF and not TOC after NTC00787202	3	Thank you, we made change needed.
120	Table 4.20	First column of the table, treatment are not reported with doses different in each part (induction and maintenance)	2	Thank you, we made changes needed.
123	Text below figure 4.18	Seven studies and not six in the following sentence: "The data inputs from the six studies included..."	2	Thank you, we made change needed.
124	Text above Figure 4.20	Seven and not six in the following text: "A total of six studies were included in the analysis."	2	Thank you, we made change needed.
125	Text below figure 4.20	Seven studies and not six in the following sentence: "The data inputs from the six studies included..."	2	Thank you, we made change needed.
185	2.2.2	Same paragraph duplicate "Outputs generated..."	3	Thank you, we deleted duplicated text.
191	Second line and last line	All the reference missing in this page are the following: [Jansen. Ulcerative colitis_Ustekinumab STC Final. 2019-07-01.pdf]	3	Thank you, we made changes needed.

Comments on the 2nd draft rapid assessment on ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (uc) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies

Page	Line	Comment	Character of comment[i]	Reply from author
104	Third line first paragraph	"Generally, credible intervals from most endpoints were quite wide...". Not true for induction results	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
105	Second paragraph – Line 11	OR=1.07, 95% CrI = [0.62;1.87]	1	Thank you, we made changes needed.
105	Section 4.7.2.1.2 second paragraph, third line	SUCRA of 79-86%	1	Thank you, we made changes needed.
150	Biologic failure population section- first paragraph	Text from EUnetHTA's assessment: "The 1-year NMA of effectiveness performed in the biologic failure group show significantly higher odds of reaching response, remission and mucosal healing...". Comment Janssen: The mucosal healing should be deleted from this sentence as the 1-year NMA on biologic failure population was not run for mucosal healing.	2	Thank you, we made change needed.
150	Last paragraph before safety section	Text from EUnetHTA's assessment: "Overall, given that the NMA analysis of mucosal healing in the biologic failure group was not feasible as imputation data needed for placebo were not available in this population, and taking into account various methodological issues assigned to maintainance models in biological failure patients (i.e. (in)di-rectness and paucity of data, data imputation, expected heterogeneity, and NMA models used) the resulst should be considered very low evidence." Comment Janssen: This implies that the non-feasibility of the mucosal healing 1-year NMA in the biologic failure population leads to low evidence for the other results that are presented for the biologic failure population. This is not correct as it is a limitation of one of the analyses only.	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
150	Last paragraph	Suggest adding words, replacing "No relative safety is availlable because various..." by "No relative safety results after 1 year are available.."	3	Thank you, we made change needed.

Comments on the 2nd draft rapid assessment on ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (uc) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies

Page	Line	Comment	Character of comment[i]	Reply from author
24		Text from EUnetHTA's assessment: "Overall, given that various methodological issues related to maintenance models in non-biological failure patients were raised and discussed, such as directness and paucity of data, data imputation, potential impact of bias due to pooling of doses, expected heterogeneity, and NMA models used, the results should be considered very low evidence." Comment Janssen: The methodological issues have been incorrectly described without details, are based on assumptions that are not corroborated by the data, and make no distinction for which comparisons methodical limitations apply, and for which they do not (see follow-up comments below). Proposed text: For certain comparisons, in particular the comparison in the biofailure population and the comparison in the non-biofailure population versus tofacitinib, certain methodological limitations apply such as data imputation, potential heterogeneity, the results should be considered very low evidence.	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
31		Text from EUnetHTA's assessment: "The evidence drawn from all the *networks* (*NMAs*) was based on indirect data - Bayesian network meta-analyses in a setting where heterogeneity or inconsistency of a network could not be statistically assessed due to paucity of data, and assessment of statistical heterogeneity between trials informing each comparison was limited and likely underpowered." Comment Janssen: Referenced articles have been incorrectly cited (see detailed comments on the 2 publications below). Proposed text: The evidence drawn from all the *networks* (*NMAs*) was based on indirect data - Bayesian network meta-analyses.	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.

Comments on the 2nd draft rapid assessment on ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (uc) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies

Page	Line	Comment	Character of comment[i]	Reply from author
31		Text from EUnetHTA's assessment: "Overall, trials that made the body of evidence for NMAs were judged as comparable at baseline in terms of patients' characteristics. However, due to paucity of data, statistical assessment of heterogeneity in distribution of potential effect modifiers across comparisons was not possible." Comment Janssen: This conclusion inadequately lumps together assessments for which imputations were necessary and those for which none, or very limited (placebo non-responders) imputation was necessary. The placebo rates, including the long-term placebo rates, for the trials with anti-TNFs and ustekinumab in the non-biofailure population are available and very similar (included in submission document, p.136, Table 29), the inclusion criteria and baseline characteristics are similar, and thus, the paucity of data does not apply to these comparisons. Neither does it apply to the comparison with vedolizumab, as vedolizumab is linked by two trials to other comparators. Proposed text: However, due to paucity of data, statistical assessments of heterogeneity in distribution of potential effect modifiers across comparisons were more limited for some of the comparisons, i.e. those in the biologic failure population, and those in the non-biologic failure population specific to tofacitinib.	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.

33		Text from EUnetHTA's assessment: "The NMA of 1-year regimens provides ambiguous results of the very low evidence certainty of the treatment effects based on treat-through comparisons. Adding to concerns about meeting the key assumption of a valid meta-analysis, the NMA models in both subgroups were built on scarce, in-direct and heavily imputed data in a setting where heterogeneity is expected but could not be tested due to paucity of data, using the fixed effects models as the only available model option. Namely, based on the differences that were observed among registrational clinical trials of biological therapies in adults with inflammatory bowel diseases, and in meta-analysis of placebo remission rates in UC patients with active disease showing significant heterogeneity of these rates ranging from 0-40% [Ghosh S, Sandborn WJ, Colombel JF, Feagan BG, Panaccione R, Hanauer S, et al. Interpreting Registrational Clinical Trials of Biological Therapies in Adults with Inflammatory Bowel Diseases. Inflammatory bowel diseases. 2016;22(11):2711-23. Epub 2016/10/19. Su C, Lewis JD, Goldberg B, Brensinger C, Lichtenstein GR. A meta-analysis of the placebo rates of remission and response in clinical trials of active ulcerative colitis. Gastroenterology. 2007;132(2):516-26. Epub 2007/01/30.], it is reasonable to assume some level of heterogeneity between the trials included in network meta-analyses." Comment Janssen: a) the assessment is factually wrong on the term 'heavily' imputed. Some of the comparisons required imputation; however, the comparisons in the non-biofailure population of ustekinumab and all three anti-TNFs did not require imputation, or only of the placebo non-responders (a group of patients with very low remission and response rates, limiting the uncertainty). Moreover, the comparison with vedolizumab required imputation for the GEMINI-1 trial, but not for the VARSITY trial. Given that the comparisons are very similar either with or without VARSITY, the statement on low evidence only applies to the comparisons with tofacitinib, and the comparisons in the biofailure population. b) The assessment has failed to factually compare the issues raised in the Su 2007(6) paper and how they were addressed in this NMA a. The Su 2007 paper(6) assessed trials in ulcerative colitis in general, not specific those trials in moderate to severe patients who failed conventional therapy. Obviously, a broader population leads to heterogeneity, which is not relevant to this particular NMA, as described below. The paper listed several sources of heterogeneity, that are not applicable to this NMA b. Different outcome definitions: the outcome definitions are homogeneous in this NMA, with the sole exception that endoscopies were read centrally in the trial with tofacitinib c. Study duration is highly similar in this NMA for the 1-year comparison, which is the most clinically relevant comparison. In addition, the number of study visits was similar in these long-term trials. d. Disease activity was homogeneous across trials in this NMA e. All trials included endoscopic evaluations of efficacy f. Prior medication failure was homogeneous in this population	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.

Page	Line	Comment	Character of comment[i]	Reply from author
		c) Ghosh 2016 (5) (funded by Abbvie) has raised a number of theoretical issues, however without accompanying data. In fact, the only data displayed compare PRECISE1 and 2 data, highlighting that indeed randomised withdrawal trials need to be re-calculated in order to compare them with treat-through trials. The paper focusses on IBD trials in general and fails to highlight that in many randomised withdrawal trials in UC, patients responding to placebo continue on placebo, significantly reducing the differences in design between randomised withdrawal trials and treat-through trials. The authors 'speculate' that patients in current trials are atypical with regard to prognosis. However, the data provided by Janssen in this submission, i.e. the short and long-term placebo rates of the first phase III study (ACT-1), PURSUIT and the last phase III study (UNIFI) do not support this hypothesis. As a result, this assessment "reasonably assumes" heterogeneity based on a paper that is not factually compared with the submitted data. Moreover, the speculation that patients in treat-through trials are 'atypical' because the treat-through design "does not reflect clinical practice [since] Physicians do not persist with drug therapy beyond induction unless the patient experiences a clinically important benefit" does not corroborate with the design of the randomised withdrawal design in UC. All the randomised withdrawal trials in UC allow continued treatment with active drug beyond the induction time period after non-response to induction.		

33		Text from EUnetHTA's assessment: "In the non-biologic failure population in the maintenance phase all the point estimates for all the comparisons, irresectable of their significance, point towards high increase in odds of achieving an endpoint in patients treated with ustekinumab compared to those treated with other active treatments or placebo, all that one could conclude about the effect sizes from width of nonsignificant CrIs (ranging from i.e. 0.6-3.5) is that not much is known about the effect, and that more data are needed. Even in case of significant findings in this population, the interpretation is hindered." Comment Janssen: Translating point estimates (and CrI) for the OR into risk differences highlights the large differences in probabilities and distribution of these. An example for the non-biologic failure 1-year network for clinical remission demonstrates that differences in absolute response rates (derived from the OR) for UST vs. most active comparators exceed differences in absolute response rates for some active treatments versus placebo, (which are all considered as clinically meaningful). Risk difference (95% CrI) for ustekinumab vs. each treatment: Vs. VDZ 300mg-VDZ 300mg pooled: 0.15 (-0.07-0.36) Vs. TOF 10mg-TOF pooled: 0.19 (0.01-0.36) Vs. IFX pooled-IFX pooled: 0.23 (0.05-0.39) Vs. GOL 200/100-GOL pooled: 0.32 (0.16-0.46) Vs. ADA 160/80/40-ADA 40mg EOW: 0.37 (0.20-0.52) Vs. placebo: 0.48 (0.36-0.59) Even when applying an overly strict rule that not the lower bound of the 95% CrI of the risk difference needs to be clinically meaningful, it is clear that the lower bound of the risk difference with most anti-TNFs exceeds the RD between adalimumab and placebo (12.6 % in the ULTRA-2 trial), which was considered to be clinically meaningful. Thus, clearly establishing the differences observed are clinically meaningful against most each one of the anti-TNFs. Since these differences are observed against the anti-TNFs, with trials conducted for the anti-TNFs over a long period of time with different designs, it these results can be established as clinically meaningful and high	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.

Page	Line	Comment	Character of comment[i]	Reply from author
		evidence. Please refer as well to comment below on Bayesian versus frequentist interpretation of results.		
31		Text from EUnetHTA's assessment: "The list of available endpoints for the NMA could be drawn from the table reporting on characteristics of included studies in the submission file, but judging from the data presented for the UNIFI trial, not all the endpoints are listed. Specifically, major secondary endpoint corticosteroid-free clinical remission was not considered for NMA analyses which increases the risk of bias in selection of the reported result of the indirect comparison analysis (as discordant pieces of evidence possibly drawn from a relevant secondary endpoint may not be able to inform decision making if available but excluded from analysis)." Comment Janssen: Just like other endpoints, such as those based on IBDQ, corticosteroid free remission was considered for the NMA. However, data for corticosteroid free remission from comparator trials are not reported by biofailure/nonfailure subgroup, neither are the results reported for the delayed responders in the randomised withdrawal trials, rendering an NMA not feasible. Of note, in the UNIFI study, 97.2% of patients in clinical remission at week 44 (of the re-randomised responders) (139 of 143) were corticosteroid-free at week 44. Proposed text: … Specifically, major secondary endpoint corticosteroid-free clin-ical remission was not considered for NMA analyses as data for the subgroups of the comparators were missing, and results for delayed responders were also not available.	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.

Comments on the 2nd draft rapid assessment on ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (uc) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies

Page	Line	Comment	Character of comment[i]	Reply from author
35		Text from EUnetHTA's assessment: "Non-biologic failure population In the non-biologic failure population, patients having received ustekinumab 6 mg/kg in induction followed by ustekinumab 90mg had significantly higher odds of achieving clinical response, remission, and mucosal healing after a year long treatment regimen than placebo, golimumab and adalimumab. While the point estimates for these comparisons were high, ranging from OR of 2.4 to 8.7, the findings need to be viewed with caution as the certainty of estimated effect sizes was low, as evident from wide CrIs. In addition, clinical response was significantly more likely in patients treated with ustekinumab than those treated with infliximab or tofacitinib. However, it should be stressed that various methodological issues were raised and discussed about maintenance models, which undermines one's confidence in these results. Overall, given that various methodological issues related to maintenance models in non-biological failure patients were raised and discussed, such as directness and paucity of data, data imputation, potential impact of bias due to pooling of doses, expected heterogeneity, and NMA models used, the results should be considered very low evidence." Comment Janssen: The assessment factually misses to make a distinction between those comparisons where some of the issues highlighted may exist, and those comparisons were those issues do not apply. Moreover, considering the level of evidence for comparisons in the non-biologic failure population versus adalimumab, golimumab, infliximab and vedolizumab, to be similarly low to the evidence for comparisons versus tofacitinib and in the biofailure population would entirely ignore this.	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
33		Text from EUnetHTA's assessment: "The data from earlier clinical trials on biologics are not likely to be comparable with more recently conducted RCTs as they include different populations of patients. Since a valid network meta-analysis relies on the assumption that the different sets of studies included in the analysis are on aver-age similar in all important factors that may affect the relative effects, for these NMAs it is very uncertain if the results of NMAs are valid or are due to bias/artefact of analysis." Comment Janssen: This statement is based on a publication (funded by Abbvie (5)) that listed theoretical considerations but without actual data (see above). The data provided by Janssen in this submission, both for short- and long-term placebo, do not demonstrate a meaningful difference between the rates observed in ACT-1, PURSUIT and UNIFI.	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.

Comments on the 2nd draft rapid assessment on ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (uc) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies

Page	Line	Comment	Character of comment[i]	Reply from author
33		Text from EUnetHTA's assessment: "The NMA of 1-year regimens provides ambiguous results of the very low evidence certainty of the treatment effects […]." Comment Janssen: The term 'ambiguous' is not substantiated. Irrespective of the level of evidence (see previous comments), the results are straightforward, given that SUCRAs, ORs with credibility intervals and probability of being better all point in the same direction.	1	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
11	Line 38, 39, 40	Antibiotics are not a commonly recommended treatment for ulcerative colitis. In the ECCO guidelines it states: "antibiotics, only if infection is considered [such as in a first attack of short duration; after recent admission to hospital; or after travel to an area where amoebiasis is endemic], or immediately prior to surgery. Controlled trials of oral or IV metronidazole, tobramycin, ciprofloxacin, or vancomycin in acute UC have shown no consistent benefit in addition to conventional therapy". Proposed amendment (in strikethrough): Medications to treat UC include: aminosalicylates (5-ASAs), corticosteroids, immune modifiers (immunomodulators), biologics, ~~antibiotics,~~ and OTC medications, such as antidiarrheals and pain relievers.(8)	2	Thank you, we made change needed.
39	First paragraph	Same suggested change as above (i.e. for p.11)	2	Thank you, we made change needed.
91	GEMINI 1 Column	Reference for GEMINI 1 is incorrect Proposed amendment: Replace reference with the following for Feagan 2013 (9): Feagan BG, Rutgeerts P, Sands BE, Hanauer S, Colombel JF, Sandborn WJ, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. 2013. N Engl J Med. 2013;369(8):699-710	3	Thank you, we made changes needed.

Page	Line	Comment	Character of comment[i]	Reply from author
139	Sixth paragraph	Comment regarding standardisation of endoscopic readings was misinterpreted. The EMA does mention that "adjudication of endoscopic evidence of activity should be performed, preferably by central reading of the examinations. If decentralised reading of examination is performed, standardization of reading should be demonstrated." This was a guideline created updated in 2018 to advise manufacturers on how to design their trial and how to accurately assess endoscopic evidence. Therefore, Janssen does not have the information to discuss the standardisation between trials that have not been performed by Janssen. In other words, it is not possible to standardise the endoscopic readings post-study. <u>Proposed amendment:</u> Remove this sentence (in strikethrough): A potental source of heterogeneity due to differences in outcome measurements was however observed for the mucosal healing endpoint as the endoscopic score for the efficacy analyses used in the NMA was assessed in 6 out of 8 studies by a local endoscopist (VARSITY trial didn't report who performed the readings, while in OCTAVE readings were performed centraly), ~~and no info on wheather standard-isation of local reading was performed in these studies was presented by MAH. EMA guidelines on UC [EMA Guideline on the development of new medicinal products for the treatment of Ulcerative Colitis. 2019. https://www.ema.eu-ropa.eu/en/documents/scientific-guideline/guideline-development-new-medicinal-products-treatment-ulcerative-colitis-revision-1_en.pdf] recommends that adjudication of endoscopic evidence of activity should be preferably done by central reading of the examinations, and if local reading of examination is performed, standardization of reading should be demonstrated.~~	2	This comment is not related to a factual inaccuracy and is, therefore, outside the scope of a fact check.
200	ULTRA 2 Column	Reference for ULTRA 2 (NCT00408629) is incorrect. <u>Proposed amendment:</u> Replace reference with the following for Sandborn, 2012 (10) Sandborn WJ, Van Assche G, Reinisch W, Colombel JF, D'Haens G, Wolf DC, et al. Adalimumab induces and maintains clinical remission in patients with moderate-to-severe ulcerative colitis. Gastroenterology. 2012;142(2):257-65. e3	2	Thank you, we made change needed.

Comments on the 2nd draft rapid assessment on ustekinumab for the treatment of adult patients with moderately to severely active ulcerative colitis (uc) who have had an inadequate response with, lost response to, or were intolerant to either conventional therapy, a biologic, or have medical contraindications to such therapies

Page	Line	Comment	Character of comment[i]	Reply from author
248	Table A22	Missing treatment group in ACT 1 (Induction) row: <u>Proposed amendment:</u> Add IFX 5mg/kg treatment group Reference: Rutgeerts 2005 (11)	2	Thank you, we made changes needed.
252	Table A28	Additional placebo row under PURSUIT – M. Proposed amendment: Remove additional PBO row (i.e. typo).	2	Thank you, we made change needed.
		The question “Is the information regarding the reimbursement status of your product correct?” from the Fact Check Checklist cannot be answered since no reimbursement information for ustekinumab was included in the assessment document. Reimbursement details regarding ustekinumab were included in Section 3.1 and 3.2 in the original submission document, sent by Janssen.	2	Since the submission dossier is publically available, we did not duplicate this information in the Joint Assessment Report

References

1. FDA. FDA Advisory Committee Meeting Briefing Document (sNDA 203214 Supplement 018). Tofacitinib (CP-690,550) Ulcerative Colitis.; 2018.
2. European Network for Health Technology Assessment (EunetHTA). EUnetHTA Guidelines: COMPARATORS & COMPARISONS: Direct and indirect comparisons. 2015.
3. Higgins JPT TJ, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019) 2019 [Available from: www.training.cochrane.org/handbook.
4. Xeljanz: summary of product characteristics [Available from: https://www.ema.europa.eu/documents/product-information/xeljanz-epar-product-information_en.pdf.
5. Ghosh S, Sandborn WJ, Colombel JF, Feagan BG, Panaccione R, Hanauer S, et al. Interpreting Registrational Clinical Trials of Biological Therapies in Adults with Inflammatory Bowel Diseases. Inflamm Bowel Dis. 2016;22(11):2711-23. doi: 10.1097/MIB.0000000000000909.
6. Su C, Lewis JD, Goldberg B, Brensinger C, Lichtenstein GR. A meta-analysis of the placebo rates of remission and response in clinical trials of active ulcerative colitis. Gastroenterology. 2007;132(2):516-26. doi: 10.1053/j.gastro.2006.12.037. Epub Dec 20.
7. Chaimani A, Salanti G, Leucht S, Geddes JR, Cipriani A. Common pitfalls and mistakes in the set-up, analysis and interpretation of results in network meta-analysis: what clinicians should look for in a published article. Evidence Based Mental Health. 2017;20(3):88-94.
8. European Crohn's and Colitis Organisation (ECCO). e-Guideline Algorithms: Ulcerative Colitis [cited 2018 Nov 26] 2017 [Available from: http://www.e-guide.ecco-ibd.eu/.
9. Feagan BG, Rutgeerts P, Sands BE, Hanauer S, Colombel JF, Sandborn WJ, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. N Engl J Med. 2013;369(8):699-710. doi: 10.1056/NEJMoa1215734.
10. Sandborn WJ, van Assche G, Reinisch W, Colombel JF, D'Haens G, Wolf DC, et al. Adalimumab induces and maintains clinical remission in patients with moderate-to-severe ulcerative colitis. Gastroenterology. 2012;142(2):257-65.e1-3. doi: 10.1053/j.gastro.2011.10.032. Epub Nov 4.
11. Rutgeerts P, Sandborn WJ, Feagan BG, Reinisch W, Olson A, Johanns J, et al. Infliximab for induction and maintenance therapy for ulcerative colitis. N Engl J Med. 2005;353(23):2462-76. doi: 10.1056/NEJMoa050516.

[i] Character of comment
- 'major'=1
- minor'= 2

'linguistic'=3