Targeted Immune Modulators for Ulcerative Colitis

Response to Public Comments on Draft Evidence Report

September 11, 2020

Table of Contents

#	Comment	Response/Integration
Patient Organizations		
Crohn's & Colitis Foundation		
1.	The Foundation is concerned that the long-term cost effectiveness model is too disparate from the real-world patient experience to be a reliable source of cost-effectiveness information. We recommend ICER include verbiage around this uncertainty in the draft evidence report, and that ICER make recommendations for how health policy decision-makers should utilize the uncertain conclusions. Further, ICER could partner with organizations like the Foundation to identify information gaps and launch efforts to support the cultivation and use of more real-world data."	We recognize that UC patients switch TIM therapy many times over the course of disease. While the cost effectiveness of treatment sequences is an important question, we sought to answer a wholly different one—the cost effectiveness of each TIM after failure by conventional or initial TIM therapy. This required striking a balance between adding subsequent lines of treatment and maintaining a model structure with sufficient focus on the initial therapy. Furthermore, we tested these assumptions in sensitivity and scenario analyses. We have added language around the uncertainty associated with these decisions to Section 5.4 of the report.
2.	ICER's assumptions for biologic efficacy and use do not align well with real-world efficacy data. Per the model assumptions, the median patient is on the specific targeted immune modulator (TIM) for less than one year and the average patient-duration on the specific TIM is less than 1.5 years. OPUS, a European registry, finds a median infliximab use duration of 19.8 months. Furthermore, ICER's ongoing maintenance transition probabilities are "memoryless" – where the patient has the same probability of response failure each subsequent cycle. In the real world, treatment success often persists, with remission history begetting more continued response.	In the model, persistence to TIMs is driven by the probability of achieving and maintaining response, which is driven by results of the ICER NMA. We assume that a proportion of patients will discontinue the TIM based on clinical trial data. These assumptions may differ from individual patient and physician decisions in clinical practice but reflect the relevant clinical trials and guidelines. Of note, 18 months (1.5 years) is relatively similar to the 19.8 months observed with infliximab. While the "memoryless" feature is a known limitation of Markov models, this particular Markov model is based on the probability of losing response each cycle given that a patient has already achieved response or remission. Due to data limitations, we were unable to generate a set of maintenance transition probabilities separately for those entering maintenance with clinical response without remission versus remission.

3.	ICER's model does not consider infliximab as a treatment for biologically-experienced patients. Infliximab was the first biologic approved for ulcerative colitis (2005), and its FDA indication does not distinguish between biologically-experienced and naïve patients. However, infliximab is commonly used for biologically-experienced patients. ICER, in its primary model, should model infliximab and the newer infliximab bioequivalents as drugs for biologically-experienced patients. In the ICER sensitivity analysis, using infliximab after vedolizumab was a cost-effective strategy. For those with comorbidities or moderate disease severity, this may be a highly cost-effective sequence of therapies. This finding further questions the key assumptions of the primary model and reinforces the lack of external validity of the current ICER report and key findings.	We agree, based on stakeholder and clinician interviews, that infliximab is sometimes used in clinical practice for patients who have been failed by prior biologics; however, were unable to include infliximab in the biologic-experienced population due to lack of RCT data. Response in a biologic-naïve population is known to differ from response in a biologic-experienced population. Thus, including infliximab in the biologic-experienced population based on biologic-naïve data would introduce bias in favor of infliximab or treatment sequences that include infliximab due to the population studied. A scenario analysis was undertaken to explore the sequence of vedolizumab followed by infliximab using the biologic-naïve population efficacy data for infliximab. The cost per QALY for vedolizumab was reduced under this scenario compared to the base case but remained well above the $150,000 per QALY threshold for both the biologic-naïve population and biologic-experienced population.
4.	Further, ICER assumptions for biologic efficacy and use do not align well with real-world biologic use. While ICER's assumptions for biologic discontinuation are too high, ICER underestimates that portion of discontinuations that are for reasons other than failure to respond to the drug. Non-response is one of many reasons for discontinuation; other reasons include drug costs (patient cost sharing), health plan approvals and hassles, concerns over long-term safety, side effects, and pregnancy.	We have attempted to model a complex set of patient-level decisions regarding TIM persistence using the data available to us. Within the model, TIM discontinuation is based on lack of response or for reasons other than efficacy from the clinical trials. Discontinuation rates from clinical trials are often lower than in the real world due to reasons stated. Overall, the duration of persistence in our model is similar to that reported in the literature. Please see: • Khan S, et al. Real-world evidence on adherence, persistence, switching and dose escalation with biologics in adult inflammatory bowel disease in the United States: A systematic review. • Chen C, et al. Real-world pattern of biologic use in patients with inflammatory bowel disease: treatment persistence, switching, and importance of concurrent immunosuppressive therapy.

5.	Finally, over a lifetime, patients may have more than two rounds of biologic drug use, which is not incorporated into the model. Patients, per the model, will spend less than two years of treatment on biologics and so nearly all of their 30+ year life expectancy on conventional treatment or with a colectomy. The concept that patients would undergo treatment with two TIMs and then be cycled on long-term steroids or immunomodulators is really not consistent with current clinical practice. The Foundation has found that patients may try several other TIMs along their disease journey as they seek to avoid surgery at all costs, and their health care providers seek to avoid long-term steroids.	We agree that over a lifetime, patients may try additional treatments beyond those included in our model. However, our intent is only to evaluate the relative cost effectiveness of the initial TIM of interest. The addition of downstream treatment options beyond those included in the model is unlikely to substantially impact the relative cost effectiveness of the initial TIM. This is reflected in the scenarios evaluating shorter time horizons, the results of which did not change any conclusions regarding cost effectiveness.
6.	Similar to the patient experience assumptions, ICER's cost estimates are inconsistent and not aligned with the costs to payers in the real world, which limits the reliability of ICER's findings. ICER should cite this uncertainty in its report and, given these uncertainties, recommend how the report can be leveraged by health policy decision-makers. Further, the Foundation encourages efforts to better understand the patient perspectives on UC health state utility values as ICER's assumptions do not align with a commonly held priority of patients to avoid surgery.	Regarding costs, please see our response to Comment 8. We agree there is an unmet need to better understand utility values in UC and have reviewed our utility value assumptions in order to select best estimates from published literature within the model.
7.	ICER's drug costs are manufacturer net revenues and not the average cost to payers, which impacts the relative cost effectiveness of some TIMs compared to others in the draft report. ICER's costs are estimates of the average manufacturer revenue for the drugs, net of discounts, rebates, concessions to wholesalers and distributors, and patient assistance programs. Payers instead pay a price that is marked-up by wholesalers, PBMs, and providers (pharmacies, physician offices, and hospital outpatient departments). Physician-administered drugs (paid for as a medical benefit) generally have higher markups than self-injected and oral drugs (paid for as a drug benefit), particularly if the drug is administered in a hospital outpatient setting. Including mark-ups in the cited drug costs will increase infliximab's cost relative to adalimumab, golimumab, tofacitinib, and ustekinumab and decrease its cost effectiveness advantage.	We thank you for the comment and agree that pricing transparency would improve our ability to accurately compare costs across TIMs. In response to public comments, the revised model is based on Medicare ASP Payment Allowance Limits for Medicare Part B Drugs for all IV-infused products (infliximab, infliximab-abda, infliximab-dyyb, vedolizumab, and the ustekinumab IV loading dose), and remain based on net pricing data for self-injected or oral medications that are typically sourced through traditional specialty drug channels (adalimumab, golimumab, and tofacitinib).

8.	ICER's non-drug healthcare cost assumptions do not represent U.S. averages, have several gaps, and are poorly documented. Based on the forward of the draft evidence report, we expect to see healthcare cost assumptions that represent average costs for patients with commercial, Medicare, Medicaid, and other insurance. Instead we find a mix of costs. For example, some costs are pulled from Medicare but omit what commercial payers pay. Further, costs frequently do not take into account associated costs of key procedures and hospitalizations/outpatient care. These inconsistencies need to be addressed to strengthen the findings of the draft evidence report. Given space constraints, we are not including specific examples in this paper, and we can share more details upon request.	Non-drug health care costs are comprised of health state costs, colectomy costs, adverse event costs, and cost of administration. For the most part, non-drug costs are based on a private payer perspective. For example, annual cost for each health state (Cohen 2015) is based on privately insured employed patients and includes all direct costs except the drug (inpatient, emergency department, and outpatient). Key procedures, hospitalizations, and outpatient care are non-drug costs that are captured within these estimates. Cost of administration is based on the Medicare Physician Fee schedule without adjustment to a typical commercial payment, but this is unlikely to have a substantial impact on the conclusions of the cost-effectiveness analysis.
9.	Additionally, per ICER's QALY assumptions and our calculations, colectomy by far dominates biologics for maximizing QALYs, implying that UC patients who are candidates for biologics should instead want an immediate colectomy. Since this is clearly not the case, we question the overall veracity of ICER's utility values and treatment efficacy assumptions. While biologics provide a potential short-term boost in QALYs compared to conventional treatment, colectomy (per ICER's assumptions) produces a long-term boost for everyone except the 15% of the population who ICER assumes develop chronic pouchitis. We estimate that the net (of pouchitis) lifetime QALY boost from a colectomy, per ICER's assumptions, is a multiple of the 0.20 gain from biologics. Yet, in the real world, patients who are candidates for biologics are not on the phone asking their payers for approval for colectomies. In fact, ICER estimates that only 18.9% of patients with conventionally-treated UC will get a colectomy over a period of 10 years. More work needs to be done to assess patient perspectives to inform accurate health-state utility values.	Colectomy is dominant over TIMs in our model and in prior evaluations of biologics in UC. For example, see the prior NICE evaluation and resulting publication: • Archer R, et al. Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy (including a review of TA140 and TA262): clinical effectiveness systematic review and economic model. • Tappenden P, et al. A model-based economic evaluation of biologic and non-biologic options for the treatment of adults with moderately-to-severely active ulcerative colitis after the failure of conventional therapy. Despite this finding, we acknowledge that colectomy is an option of last resort for most patients and for this reason, our model positions colectomy as an option for a small proportion of patients each cycle who do not receive benefit with medical treatment rather than as a treatment option to directly compare to TIMs. Additional work to characterize the utility of UC patients post-colectomy would be a valuable contribution to scientific literature but is outside the scope of this review.

10.	ICER should reframe the lifetime-horizon data point. A casual reader of the cost-effectiveness summary and tables might reasonably, yet incorrectly, conclude that ICER modeled the costs of a patient receiving a single biologic drug for the remainder for their life vs. the cost of the patient receiving conventional treatment. Section 5.4, Summary and Content, starts with "We estimated the cost effectiveness of TIMs over a lifetime horizon…" and the data tables name single biologics. As noted above, per ICER's model, the median patient is on the specified (first) biologic for less than one year and the mean patient is on for less than 1.5 years. If the patient fails the induction phase of the first biologic, the patient receives a second biologic, which the patient will also most likely be on for a very short time. The modeled patient will therefore have biologic treatment for less than 10% of their expected 30+ year life expectancy and the biologic costs reported by ICER for the specified biologic include the costs for both the first and, when applicable, second biologic drug.	We thank you for bringing this to our attention. We have revised the wording in the report accordingly to prevent misinterpretation.
11.	ICER should share its pathways from cited sources to the assumptions used in the modeling. Most of the assumptions are not straight "pick-ups" from the cited paper(s) into the model, and it is clear that some degree of transformative judgment and blending of papers was required. The transformations are as critical to assumption soundness as the underlying citation, yet they are described in only the most general terms, if at all. NMA workpapers would fill-in the missing information. Without access to these models, we and other stakeholders cannot fully understand the dynamics of the modeling or check for potential errors.	As with many cost-effectiveness models, a fair amount of data transformation is required to make published estimates fit within the model framework and cycle length. The report intends to strike a balance between transparency and redundancy of common transformations.

12.	ICER's review of payer drug-approval policies ignores/understates payer-approval complexities and the lack of alignment with clinical guidelines. The Coverage Policies and Clinical Guidelines section presents default commercial payer coverage policies for 17 of the largest U.S. national and regional commercial medical payers. While ICER finds an inconsistent jumble of policies that often do not align with clinical guidelines, the reality is worse. For example, UnitedHealthcare has separate policies for the use of infliximab for commercial insurance and Medicare Advantage and the language within both policies makes it clear that numerous exceptions exist. Payer policy is further complicated when a patient has prescription drug benefits via a PBM whose benefits are not integrated with the patient's medical benefits. Divergent health plan and PBM policies are particularly stressful to UC patients as some UC drugs (those that require physician administration) are typically paid for as a medical benefit and other UC drugs are paid for under the prescription drug benefit. PBMs set their own benefit policies and their policies vary by health plan, self-insured employer, and state.	We thank you for this comment. While some of the more nuanced inconsistencies mentioned here are described in Section 2 (Patient Perspectives), we have added to the text in this Section to address the additional issues raised.
13.	Finally, we will short-change our patients if we do not note that patients and their physicians often have to navigate multiple policies as patients move between health plans and health plans modify their payment policies. We further note that it is estimated that 98% of top health plan formularies are not aligned with the American Gastroenterological Association's (AGA) UC clinical guidelines, further complicating access to appropriate and needed care.	We have amended Section 2.2 (Impact on Patients) to integrate these concerns.

14.	ICER should recommend against the use of insurance-mandated step therapy among the UC patient population, particularly for moderate to severe patients, in its final policy recommendations paper. For decades, the Foundation has invested in medical research and has played a part in the discovery of every recent therapy brought to market so that UC patients have many options to treat the unique manifestation of their condition. Despite these advancements in research, high pharmaceutical costs and insurance-mandated medical management limit patient access to the full suite of available treatments. The draft evidence report notes there are currently limited head to head studies and sequencing information, therefore, providers and patients must make the best decision based on clinical guidelines, the provider's experience, and the patient's known characteristics, treatment preferences, and medical history. It is impossible for an insurance-mandated step therapy protocol to take all of these factors into account. Further, the consequences of undertreatment are quick and devastating for UC patients – many patients who do not receive adequate treatment in time experience irreversible damage to their large intestine and can degrade so far as to need a colectomy. It is often thought – though difficult to prove – that the colectomy could have been avoided if the patient had access to the right treatment at the right time. Further, ICER notes the high cost per QALY of every assessed TIM. Requiring patients to step through inappropriate treatments reduces the value of the payer's-preferred TIM that is unlikely to work. Therefore, the Crohn's & Colitis Foundation strongly urges ICER to recommend against the use of step therapy among UC patients.	It is expected that step therapy will be a featured element of discussion at the meeting's Policy Roundtable.
15.	ICER should encourage pharmaceutical companies and insurance plans to work together to improve fair patient access at affordable prices. Even if many of the TIM prices referenced in Table 5.17 Cost of Induction and Maintenance were slashed in half, many UC patients would still be unable to afford their out-of-pocket treatment costs. Insurance coverage is a must for patients to access the therapies they need to thrive. While the Foundation recognizes the purpose of the UC Value Assessment is to assess whether the TIM drug prices meet the usual standards for cost effectiveness, ICER must acknowledge the critical lifeline these treatments present for patients, and urge manufacturers and payers to come together to ensure that patients can afford their treatments and maintain consistent, regular therapy without undue formulary burdens.	While Section 2 does provide cited information on out-of-pocket costs, we have expanded this information to include additional figures and citations.

16.	As stated above, while ICER plays an important lead role in advancing the use of value assessments in the American healthcare system, ICER should recognize the limitations of its comparisons and valuations, and encourage further studies leveraging RWD. Better study inputs, such as carefully gathered RWD, would improve the validity and usefulness of the value assessments. The Foundation will continue to develop and promote IBD Plexus as a platform for RWD that can support better pricing and formulary decision making.	We agree that robust collection of real-world data would be useful in assessments of the long-term impact of UC therapies; while available real-world evidence is summarized in the report, these studies were often conducted in highly selective populations and typically focused only on one or two TIMs.
Partnership to Improve Patient Care		
1.	UC is a highly heterogeneous condition. The presentation of symptoms and disease course can vary substantially among patients. In some, the disease course may reflect periods of active disease and remission, while in others the symptoms are persistent despite increased use of medical therapy. In addition, there are noted but currently poorly understood differences in how racial and ethnic minorities experience UC. The Crohn's & Colitis Foundation highlighted this heterogeneity in their initial letter to ICER, noting that treatment needs may vary greatly based on the specific patient's presentation of the disease: "UC is heterogeneous and the needs of each patient unique…Because each patient is unique and UC is a chronic and generally progressive disease, optimal care for the UC patient requires timely access to the full suite of treatments currently available." Though ICER acknowledges the reality that UC patients are heterogeneous in their report, they neglect to represent this in their base case for cost-effectiveness and continue to base these judgements off of an "average" patient. Reliance on averages in cost-effectiveness analyses has shown to be illogical and unscientific. It also results in very real harm for many patients whose experiences do not sit conveniently close to the averages portrayed in these sampling-based summaries of widely varying sets of outcomes. As a result, new therapies that are likely to have significant impacts on the lives of patients with life-altering levels of discomfort and pain will not be made available – or will have their access restricted – simply because other patients deemed to fall into the same disease category experience far less of such pain and discomfort, or experience it infrequently rather than constantly, diluting the effects of the former group.	ICER always seeks to address concerns around heterogeneity of disease impact and treatment effects in its review of available evidence. The challenge is that evidence of heterogeneity is often missing from available clinical trials and other comparative studies due to limited presentation of subgroup data. In this case, we stratified data from the outset according to where subgroup data was robust—namely, according to prior biologic use and induction versus maintenance treatment. In the cost-effectiveness analysis, we also conducted extensive sensitivity analyses that varied the quality-of-life impact of the active, response, and remission states of health in UC, and found that, while results varied substantially, they generally did not cross common cost-effectiveness thresholds.

2.	PIPC would like to reiterate the point it has made to ICER in past comment letters that the use of the Quality-Adjusted Life Year (QALY) is inappropriate in assessing treatments for chronic illnesses. For many UC patients, incremental improvements in health without having to undergo surgery are significantly beneficial to their quality of life, even if they never achieve "perfect health." ICER's use of the QALY in this report is particularly concerning because the utility weights used vary considerably from other published estimates.	ICER follows common academic and health technology assessment standards by using the cost per QALY gained, but also presents cost per life year gained and cost per evLYG. The QALY is the gold standard for measuring how well a medical treatment improves and lengthens patients' lives and has served as a fundamental component of cost-effectiveness analyses in the US and around the world for more than 30 years. A recent legal analysis found that the QALY does not disadvantage patients who have a disability or a chronic condition that is not curable: • http://icerreview.org/wpcontent/uploads/2020/03/ICER-Analysesand-Payer-Use-of-Cost-effectivenessResults-Based-on-the-QALY-and-evLYGAre-Consistent-With-ADA-Protections-forIndividuals-With-Disabilities.pdf.
3.	ICER uses utility weights for active UC, clinical response without remission, and clinical remission of 0.69, 0.78 and 0.87 respectively. Other published estimates of the utility weight of time spent in active UC not only vary considerably from the figure used in the base case for ICER's report, but they also tend to have multiple figures describing various levels of severity of active disease. For example, Woehl et al and Tsai et al estimate mild active UC at 0.72, but for moderate and severe disease the utility is estimated at 0.42. Similar estimates have been used in numerous UC treatment models in the last decade. This is also the estimate of utility for severe disease that was suggested by the Evidence Review Group of NICE in recent submissions for amongst others vedolizumab.	Utility values in the model are based on active moderate-to-severe disease, mild disease (clinical remission without response), and remission. Woehl 2007 is a commonly cited, non-peer-reviewed abstract with considerable limitations, which was also used in the analysis by Tsai et al. Several more recent peer-reviewed studies have led to higher estimates for active moderate-to-severe disease than those reported by Woehl et al. After further review and consideration, we revised the source for utility values and made a change to ensure that they represent an active moderate-to-severe population. We are now using Gibson et al. in the base case and include an alternative estimate for active UC based on an average of baseline EQ-5D scores across the InspireADA, GEMINI 1, and GO-COLITIS trials (0.658) in a scenario analysis.

4.	ICER's choice to use a single "active disease" utility weight that represents an average across all patients with the disease, rather than one which represents the population with moderate to severe disease for which these drugs are indicated, is concerning. As mentioned previously, UC is a disease of significant heterogeneity in terms of severity. It is inevitable that the utility weights for patients suffering moderate or severe disease will be lower, as their suffering is more extreme, and the use of an artificially higher utility weight for active disease will underestimate the value of any effective treatment as ICER's model uses the utility weight as the most impactful input variable.	The draft report relied on a meta-analysis to inform utility values for health states, including active disease. In response to comments, we reviewed the studies included in the meta-analysis and identified that some mild active disease was included in the "active UC" estimate. As a result, a supplemental targeted literature review was conducted to identify a single best source for utility weights, which specifically reflect active moderate-to-severe disease and our base-case utility values were modified accordingly. See our previous response to Comment 3.
5.	Patient advocacy groups have voiced concerns to ICER about the narrow scope of symptoms collected in clinical trials, which omit outcomes important to patients, like pain, fatigue, and depression. Published studies confirm the need to incorporate this data. One recent study of nearly 300 Crohn's and UC patients found that 40% of respondents met criteria for chronic pain and nearly 20% reported opioid use, and much of this pain was not directly explained by rate of incidence of disease activity. Despite this consistent message from patient groups and researchers, it does not appear that these outcomes were incorporated into ICER's economic model.	We agree that our model does not capture all outcomes that may matter to patients and we encourage future clinical trials to capture these outcomes in a systematic way. The health-related quality of life impact of these outcomes would be indirectly captured within the health state utility values used in the model. Additional data generation to show an impact of TIMs on patient-important outcomes, such as pain, would enhance our ability to model these outcomes.
6.	ICER defines only three health states for the disease: active UC, response without remission, and response with remission. Even if we assume that a generic patient-reported outcomes (PRO) tool, like the EQ-5D or SF-36, effectively captures all the components of utility in UC including pain, fatigue, anxiety and depression, the fact that many of these will not be directly correlated with disease activity, and that they will vary considerably with severity of active disease, means that ICER's simplistic representation of the disease will have a negative impact across a considerably heterogeneous patient population who are known to show wide variance in terms of treatment effect. That heterogeneous group is currently represented by just one health state when in active disease, which has the detrimental outcome of failing to capture health gains properly for patients with more severe UC.	The health states used in the model are commonly used in clinical trials, guidelines, and previous economic models. They do not represent every combination of disease symptoms that patients may experience but represent key disease health states experienced by patients. We also maintain that model captures the average health gains for the moderate-to-severe population. We do acknowledge a lack of data about patient heterogeneity sufficient to use in our model. Findings from generic and disease-specific instruments, such as the EQ-5D, SF-36, and IBDQ, are not reported consistently across trials, and when they are, summary scores (such as the MCS and PCS of the SF-36) are reported rather than the specific domain scores for pain, fatigue, and other symptoms of interest.

7.	Vedolizumab is a gut-selective biologic agent that is recommended as a potential first-line treatment for both induction and maintenance in UC but is often not covered by insurance until failure of other treatment options. In some cases, patients must try and fail at least two TNF inhibitors before vedolizumab is considered medically necessary and ultimately covered by payers, even despite clinical evidence that many patients do not respond to this biologic class. Patient advocacy organizations highlighted to ICER that one of patients' primary concerns around access to treatment is step-therapy.	We agree that assessing the value of sequencing of the TIMs is an important topic with implications for patients, but it is outside the scope of this review (please see our response to the Crohn's & Colitis Foundation's first comment).
8.	ICER's assessments are frequently considered by insurers as they develop arduous step therapy guidelines. Step therapy is primarily used by payers as a utilization management tool to help contain costs, and there is no evidence to suggest the use of step therapy improves health outcomes. Given that ICER's stated goal is to determine true clinical and cost-effectiveness of treatments, and to optimize value in the United States healthcare system, this would have been an opportunity to evaluate the impact of step therapy on clinical outcomes in UC patients. It is a worthy question to determine whether there is any long term 'value' in the payer community's reliance on step therapy, which has been shown to result in very real harm to patients. This could have been easily modeled through a scenario analysis to assess the value of treatment pathways with and without step-therapy.	Please see our response to Comment 7 above as well as our response to the Crohn's & Colitis Foundation's first comment. As mentioned previously, step therapy will be discussed at the Policy Roundtable.
9.	There is evidence that in the case of ulcerative colitis, the vast majority of insurance medical polices around prescribing for UC are incompatible with current American Gastroenterological Association (AGA) clinical pathway recommendations. ICER's UC assessment represents a missed opportunity to develop a comprehensive modeling exercise comparing step therapy to a system where patients are prescribed the most effective treatment indicated for them based on their physician's expert diagnosis, disease progression, individual patient characteristics, and relevant clinical society guidelines. ICER's decision not to capitalize on this opportunity contradicts its previously stated goal to determine the true value of treatments and is another missed attempt to better account for outcomes that matter most to patients.	We have amended Section 2.2 (Impact on Patients) to integrate these concerns.

Manufacturers		
AbbVie Inc.		
1.	AbbVie would like to reiterate while a network meta-analysis can be an effective way to compare the efficacy of one drug relative to another, the high level of variability among UC clinical trials (including different patient populations, inclusion/exclusion criteria, study designs, and clinical endpoints) may yield inaccurate or invalid conclusions. The impact these variations may have on the true comparative clinical effectiveness are amplified when considering the small differences observed between the individual TIMs (QALY ranges from 15.97-16.04, a .07 maximum difference).	We recognize the challenges associated with conducting a large NMA with multiple agents. Based in part on recommendations from manufacturers, we stratified our NMA approach by prior biologic use and timing of treatment (induction vs. maintenance), adjusted inputs to address differences in trial designs (treat-through vs. re-randomized) and conducted numerous sensitivity analyses to examine population heterogeneity. Our findings are generally consistent across this spectrum as well as with other published NMAs using a similar approach.
2.	The network meta-analysis (NMA) that ICER used to assess comparative efficacy of targeted immune modulators (TIMs) includes trials with meaningfully different clinical definitions of outcomes. For example, adalimumab trials used the worst rank method to measure Mayo scores, which ICER acknowledged may have underestimated effect sizes relative to the method of using average of the scores which most other trials did.	For this particular measure, our only alternative was to use the trial data as reported. As mentioned, we note this as a limitation of our analysis. One of the adalimumab trials (Suzuki et al., 2014) also reported average Mayo scores; when we substituted these figures in our primary NMA we saw no material change in findings.
3.	The trials included in the ICER's NMA reported varying levels of placebo response and remission rates. This may indicate differences in the underlying patient populations across trials and across treatment arms which may lead to potential differences in treatment responses and disease prognosis. For example, authors in Motoya 2019 noted treatment group imbalances that may have resulted in the unusually high placebo response and remission rates. This likely contributed to the lack of statistical difference in clinical response and remission rates between vedolizumab and placebo in the biologic-experienced population. It is worth to keep in mind that the response and remission rates of placebo compared to adalimumab in ULTRA 2 are also notably high compared to the lower placebo response and remission rates in other trials, such as OCTAVE 1 & 2, UNIFI and GEMINI 1. This suggests trial heterogeneity which is not fully adjusted for in the ICER's NMA model.	Our initial efforts suggested that including a placebo adjustment did not improve model fit, and so results were not adjusted. We revisited this and, based on the observed variation in placebo response during maintenance that is noted here, we applied a placebo adjustment for the biologic-naïve maintenance population. Such an adjustment was not feasible for the biologic-experienced maintenance population due to the much smaller size of the network, however, so we could not apply a placebo adjustment in this case.

4.	Furthermore, we would like to note that in assessing the model fit of the ICER's clinical effectiveness NMA, it is important to evaluate possible inconsistency between the direct and indirect trial evidence available. When comparing the indirect and direct estimates as reported in ICER's clinical effectiveness results, there is some evidence of inconsistency that warrants further assessment. We recommend that ICER also take this evidence into consideration during model assessment and publish the results of their inconsistency tests if possible.	We apologize for the oversight. Findings from inconsistency tests are now included in the report. We used node splitting tests for inconsistency and found no statistical differences between direct and indirect evidence in any of the four networks. Results are summarized in the Appendix.
5.	While the ICER's CEA model is structured with eight-week cycles, it uses efficacy outcomes from trials with an induction phase ranging from 6-14 weeks without adjustment for difference in the induction duration. This could introduce potential bias for trials with shorter induction duration, as patients are allowed less time to respond to treatment compared to trials with a longer induction period. We recommend that ICER consider including a time-adjustment parameter within the base-case NMA to predict more accurate 8-week outcomes to be applied within the economic model. This can be incorporated by including a coefficient for time which specifies the average difference between 8 weeks and trial induction time points on the probit scale for clinical response and remission.	Thank you for your comment. Although only two of our included trials had an induction timepoint longer than six to eight weeks, we nevertheless included a time coefficient per your suggestion and found no material or statistical differences in our results.
6.	The ICER's model pools treatment dosages across trials for efficacy assessment. It includes treatment doses which are not approved by the U.S. Food and Drug Administration (FDA), such as infliximab 10 mg/kg, golimumab 400 mg, and vedolizumab maintenance q4w. Respectfully, we suggest that ICER reconsider this approach. Since these doses are not approved by the FDA, their use in clinical practice would be rare, if used at all, likely due to safety concerns. In addition, we see numerical differences in efficacy estimates between the unapproved (higher) doses and the approved doses. For example, golimumab 200/100 mg and golimumab 400/200 mg had clinical response rates of 51.0% and 54.9%, respectively, in its phase 3 trial (Section 4.3, Table 4.2). Therefore, including these unapproved higher doses in the data synthesis may be arbitrary in resulting in a better efficacy profile while not best representing the clinical profile of these treatments in the United States. We recommend that ICER only include treatments that have been approved by the FDA in data synthesis to enhance the validity of the clinical efficacy results.	As noted in our assessment, we excluded doses lower than those approved by the FDA but included higher doses because of the well-documented use of these higher doses in routine practice. We present NMA results with doses unpooled as a sensitivity analysis in the Appendix, with results generally consistent with our primary findings.

7.	Additionally, we noticed some discrepancies in ICER's approach of data synthesis to derive efficacy inputs versus safety inputs. For example, while the efficacy estimates pooled all available doses of tofacitinib and infliximab, the safety estimates only included data for one selected dose of these treatments, despite the availability of safety data for other doses considered in the efficacy evaluation. We suggest that ICER apply a consistent pooling approach to the safety outcomes (i.e., adverse event rates) as the efficacy outcomes, and to leave out FDA unapproved doses in data synthesis for both outcomes.	Because we did not conduct an NMA on safety data, adverse event findings are presented for each dose arm in available trials that met our entry criteria (i.e., FDA approved or higher dose). Note that the table presented in the main body of the report is for the longer maintenance period only (induction results are available in the Appendix), so only relevant doses are included. We did inadvertently exclude vedolizumab 300 mg every four weeks from this table, which has now been added.
8.	The current model structure is not consistent with routine clinical practice in the U.S. as patients who do not respond to the second TIM discontinue to conventional treatment. The assumption of 2nd-line TIM use as a "market basket" of other treatment options except the initial TIM may not sufficiently capture downstream costs for this patient population, as it is common in clinical practice for patients to cycle through more than two TIMs before discontinuing to conventional treatment. While we acknowledge the lack of empiric data to estimate more than two lines of TIM use, we suggest that ICER consider a longer duration of the 2nd-line TIM use to better reflect the additional cycles of TIMs patients would use in real-world practice. Modeling multi-line TIM use for a sufficiently long period may reduce the differences in long-term costs between treatment sequences.	We appreciate the comment and suggestion for how to address the uncertainty in treatment sequences over a lifetime time horizon. As an alternative approach to balancing differences in subsequent treatments across agents, we have revised our approach to assume an equal market basket of subsequent TIMs across all agents and also allow the conventional treatment comparator arm to cycle through subsequent treatment.

9.	In the current CEA model, the source that ICER uses to inform the post-colectomy utility value in its base case measures only the post-colectomy health state. Without a broader spectrum of the UC health states considered together, a single utility value from one study may not accurately reflect the relative differences in utilities when considering pre-colectomy and post-colectomy health states. Therefore, it is difficult to assess whether the health state utilities for the pre-surgical and post-surgical health states in ICER's model are entirely comparable. For example, the ICER's model assumes that the utility of post-colectomy remission is higher than that of clinical response without remission (Section 5.2, Table 5.12). This assumption was not aligned with existing evidence which evaluated utility for both pre-surgical and post-surgical health states in the same model, and may overestimate the benefit of colectomy. Several health technology assessment (HTA) submissions used post-colectomy utility values from studies including both pre-surgical and post-surgical health state valuations, which showed that the utility value of response to treatment is higher than post-colectomy remission. For example, utility values reported in two sources cited in HTA submissions demonstrated 5.26-26.25% reduction in utility from clinical response without remission to post-colectomy, which would result in a post-colectomy value between 0.58 and 0.74 using ICER's current clinical response utility value. We recommend that ICER use a more reasonable utility value for post-colectomy in the base case.	We thank you for the thoughtful comments and have taken the opportunity to carefully review all utility values within the model. There is considerable variation in the literature for post-colectomy utility values, with estimates generally falling above moderate-to-severe active disease because of improvement in UC but below medically induced remission because of long-term consequences. It is important to note that both utility values for post colectomy cited in the 2015 vedolizumab HTA evaluation were considered by the committee as equally uncertain and not based on peer reviewed publications. While no perfect single source exists, most estimates are near or slightly below mild disease (clinical response without remission). Considering the base-case estimate (0.79) and that 15.5% of patients experienced chronic pouchitis with a utility estimate of 0.40, the weighted average post-colectomy utility score is 0.73, within the suggested range.
10.	The ICER's CEA model does not consider long-term complications of colectomy other than chronic pouchitis. Other long-term complications such as infertility and male sexual dysfunction can occur in patients who receive colectomy, which can bear significant impacts on patients' quality of life. We recommend that ICER include these complications within the model by incorporating these events into the rate of long-term complications after colectomy. The HTA submission for adalimumab estimated the transition probability for chronic complication among patients who receive surgery as 19.19%, which incorporated estimates for fertility and male impotence in addition to chronic pouchitis. Including the above-mentioned adverse events will allow ICER's model to more accurately capture the impact associated with the colectomy procedure for the patient population of interest.	We acknowledge that colectomy is associated with complications that are not each explicitly modeled individually. The utility value associated with the post-colectomy health state is captured from a survey of patients a mean of 3.7 years after first surgery, many of whom experienced complications such as infertility and sexual dysfunction. The base-case model takes this overall utility score and then applies a further utility decrement for chronic pouchitis and infection. This approach may actually "double count" the reduced qualify of life associated with complications of colectomy but was taken to avoid overestimation of the post-colectomy health state.

11.	The ICER's CEA model only considers serious infections to capture treatment-related adverse events, which may not reflect the larger impact of adverse events experienced by patients receiving treatments. Other adverse events noted by the FDA for various TIMs, including thrombotic events, cardiovascular mortality, and lymphoma, are also of relevance to this patient population. We recommend that ICER consider the above-mentioned adverse events within the modeled costs to best reflect resource use, patient disutility and mortality in the real-world patient population.	Adverse events, such thrombotic events, cardiovascular mortality, and lymphoma, are important and were considered for inclusion in the model during our initial scoping. Because these events are rare and with uncertain differences in rates across TIMs, inclusion of these adverse events is unlikely to have a substantial impact on the conclusions of the model.
12.	Other costs of UC management, including hospitalization and outpatient visits, were included by health state. We recommend ICER also to consider differences in health resource utilization incurred by patients receiving infusion treatments vs. subcutaneous treatments given the reported differences between the two. A study comparing medical costs between patients receiving adalimumab and infliximab for Crohn's disease reported significant differences in health care costs between patient groups. Average disease-related medical service costs, excluding drug costs, for the 6 months after treatment initiation were lower for patients receiving adalimumab (adalimumab: $5,199 vs. infliximab: $9,059, $P < 0.0001$). The authors attribute much of this difference to "a decreased cost of outpatient office visits" for patients receiving adalimumab, and "infusion-related expenses for infliximab." Additionally, ICER should also consider differences in non-drug costs between patients receiving TIMs and patients receiving conventional treatment, as there is evidence that health resource utilization and annual medical service costs are higher among patients receiving conventional treatments compared to patients receiving anti-TNF therapies. We suggest that ICER apply non-drug costs of UC management by treatment type to account for these differences in health resource utilization.	We thank you for your comment and for providing specific supportive evidence for this suggestion. The study cited pertains to a population of patients with Crohn's disease; it is unknown how different a study of UC patients would be. In addition to disease-related costs, all-cause total costs were similarly higher for patients receiving infliximab versus adalimumab, making it difficult to confidently attribute the difference in costs to be directly related to Crohn's disease infusion treatments. With regards to conventional treatment leading to higher non-drug cost than biologics, our model captures this through higher costs incurred because of greater time spent in active UC and response without remission health states. Without knowing the unit health care resource utilization driving the differences in direct costs, it is not possible to distinguish whether costs are higher for conventional treatment due to reduced disease control (which would be captured in our health state costs) or other factors.

13.	ICER applies an assumption that "patients who discontinue conventional treatment after two TIMs will follow the biologic-naïve transition probabilities for the biologic-naïve population," while patients who begin in the biologic-experienced population and discontinue conventional treatment after two TIMs will follow the biologic-experienced transition probabilities. The assumptions applied in the biologic-naïve population could bias results in favor of treatments that do not have data among biologic-experienced population, specifically infliximab. Moreover, it is more realistic in clinical practice to consider patients after two TIMs as biologic-experienced. Therefore, we suggest that patients who discontinue conventional treatment in the TIM-treated arm follow the transitional probabilities for the biologic-experienced population.	In the revised model, we have modified the conventional treatment arm such that after failure by initial conventional treatment patients will initiate subsequent treatment (represented by an equally distributed market basket of TIMs with efficacy in the biologic-experienced population). Upon failure by subsequent treatment, patients will then return to conventional treatment and follow the transition probabilities for biologic-experienced patients on conventional treatment for the remainder of the model time horizon. Using this approach, the outcomes for both subsequent treatment and conventional treatment after subsequent treatment are balanced across comparator arms.
14.	ICER's model assigns a second TIM for patients with a different mechanism of action from the initial TIM in their market-basket approach. However, clinical practice suggests that it is possible for patients to be treated with a 2nd-line TIM with the same mechanism of action. In fact, treatment guidelines published by the American College of Gastroenterology recommend that patients "with moderately to severely active UC who had an initial response but subsequently lost efficacy to one anti-TNF therapy" may transition to an alternative anti-TNF therapy. Thus, we recommend that ICER allow the market basket to include all TIMs except the initial TIM.	We have revised the model so that subsequent treatment is represented by an equally distributed market basket of TIMs with efficacy in the biologic-experienced population.
15.	The CEA model currently considers a modified societal perspective as a scenario analysis. This includes indirect health care costs by the health state, but does not consider costs associated with different routes of administration. Given that there may be disruption to productivity and daily life for patients receiving IV therapies due to work loss because of infusion appointments and/or associated travel costs, we suggest that the model consider the impact on indirect costs related to mode of administration. Furthermore, we recommend that ICER include additional indirect costs incurred by UC patients who undergo colectomy, as increased indirect costs among patients who receive UC-related surgery has been reported in the literature.	We agree that these are both important considerations and have revised the model to include the indirect cost of an office visit for administration of IV-infused TIMs and lost productivity following colectomy procedure.

16.	We suggest revising language in the ICER's draft report which states that some TIMs "carry a black box warning in their FDA labels for an increased risk of lymphomas and other malignancies, based on clinical trials and real-world evidence for these TIMs when studied for other indications" (Section 4.3, "Harms") to also acknowledge that the warning is primarily noted for adolescent and young-adult patients.	Thank you for this comment. We will clarify that the warning is focused on younger patients.
17.	We recommend that ICER publish additional interim clinical outcomes that are generated by the model to improve the transparency of the report. In particular, it would be helpful to report the distribution of the patients in each health state over time for each evaluated treatment separately.	An additional table has been added to the Appendix to present health state distribution over time for each comparator in both the biologic-naïve and biologic-experienced populations.
Amgen Inc.		
1.	The high utility value assumed for the active UC health state risks underestimating the quality of life burden of uncontrolled disease. ICER should consider other utility estimates for active the UC health state that have been used in previous HTA assessments in UC. The draft evidence report presents the significant impact the utility estimate for patients with active UC has on the model results in the deterministic sensitivity analysis (and is by far the most impactful parameter). The point estimate and upper bound assumed in the analysis risk overestimation of the quality of life of patients with moderate to severe UC who are not responding to treatment or are untreated. In particular, the upper bound is close to the general population utility value used by ICER to calculate evLYG (0.851), and the base case utility value for patients with moderate to severe UC in clinical remission is higher than this general population value.	We agree there is an unmet need to better understand utility values in UC and have reviewed our utility value assumptions in order to select best estimates from published literature within the model. Woehl 2007 is a commonly cited, non-peer-reviewed abstract with considerable limitations, which has been historically used in HTA some submissions. In response to comments, we conducted additional supplemental targeted literature reviews of utility estimates and carefully reviewed and revised the source for utility values. However, more recent peer-reviewed studies as well as the baseline EQ-5D of the InspireADA (0.6), GEMINI 1 (0.675), and GO-COLITIS (0.7) trials of moderate-to-severe UC patients initiating treatment with a TIM have consistently presented higher estimates for active moderate-to-severe disease than those reported by earlier studies.
2.	The draft evidence reports cited an alternative source for utility estimates, previously used by NICE. This publication provides estimates for moderate to severe UC with values for active UC ranging between 0.41 and 0.66, all below the mean value adopted by ICER in this draft evidence report. The ICER draft evidence report notes that Archer 2016 has been used to inform previous cost-effectiveness evaluations in UC, which the primary source selected for the base case analysis has not. The choice of the values (both mean and assumed distribution) used in the model should be based on the robustness of the evidence, and the selection should be justified. Additional consideration should be paid to the severity of disease for the patient population in the assessment, and potential differences in the utility of patients with active UC following the failure of prior	Please see our response to the prior comment.

	biologic treatment, compared to those who are biologic-naive.	
3.	Despite the high estimates of indirect costs, the modified societal base case is reported only as a scenario analysis. ICER should consider presenting the modified societal perspective as a co-base case. As stated in the 2020-2023 ICER Value Assessment Framework, when "the societal costs of care for any disease are large relative to the direct health care costs, and that the impact of treatment on these costs is substantial (i.e., there are substantial differences in the cost-effectiveness findings between the two perspectives), the societal perspective is included as a co-base case, presented directly alongside the health care sector perspective analysis." (ICER, 2020) Based on the results of the draft model and the scenario analysis included in the draft report, the above conditions appear to be met. The estimated indirect costs are almost as great as the estimated direct costs, so ICER should consider the inclusion of the modified societal perspective as a co-base case.	This review, which was announced in September 2019, is being conducted under the 2018-2019 ICER Value Assessment Framework and therefore, the modified societal perspective is reported as a scenario analysis. We do point out in the text describing this scenario that the cost per QALY was reduced substantially for all TIMs, but that this results in only infliximab and infliximab biosimilars falling below commonly used cost-effectiveness thresholds.
4.	The definition of the second line TIM market basket varies for each first line TIM and is not included in the conventional treatment (CT) arm. This limits the validity of the comparisons made. ICER should revise the comparisons made to ensure they are appropriate and interpret these accurately. HTA assessments in UC. The draft evidence report states that the second line TIM market basket is assumed to be made up of the treatment options other than the treatment received as first line therapy. This means that the second line treatment is different in each treatment arm, limiting the comparisons that can be made between them. Furthermore, the second line TIM is not included in the CT comparator arm, and so these comparisons are also invalidated by discrepancies in treatments received.	We have revised the model so that subsequent treatment is represented by an equally distributed market basket of all TIMs with efficacy in the biologic-experienced population. Subsequent treatment is also now part of the conventional arm treatment pathway. These changes were made to enhance the comparability across treatment arms.

5.	By defining the comparisons as they are stated in the draft evidence report, ICER is, in effect, modeling sequences of TIMs, but interpreting the results as if for a single first line treatment. Furthermore, the cost-effectiveness of any first line TIM is limited by the inclusion of the least cost-effective treatment in the second line basket. This penalizes the biologic treatments being assessed and introduces bias in favor of CT. To include multiple treatment lines in the analysis, ICER should specify specific sequences of biologics and state the assumed sequences in the evidence report so that the sequence of treatments is clear when interpreting the cost-effectiveness results.	The scope of this review is not intended to evaluate the cost effectiveness of specific sequences. Instead, we have equalized the market basket of subsequent treatment across all TIMs and conventional treatment to reduce any bias introduced by later line treatments.
6.	The per-cycle rate of colectomy and the cumulative colectomy rate are incompatible, meaning that the "cap" is reached before the end of the model. ICER should calibrate the rate of colectomy to match the cumulative rate in the comparator arm. When discontinuing a TIM, and at any time in the active UC health state on CT, patients have a probability of undergoing colectomy, incurring the costs and outcomes of the surgery procedure. The draft model contains a cap on colectomy incidence of 25.4%, and once this is reached, the probability of any future surgeries is set to 0. ICER should comment on the clinical validity of this assumption and provide greater detail on how the rate and cap were derived and applied. ICER should consider calibrating the per-cycle probability of surgery, so that it is consistent with the cumulative colectomy incidence at 20 years as stated in the draft evidence report.	As suggested, we have revised the model to include the point estimate of colectomy per cycle to be based on the 20-year colectomy rate and removed the arbitrary cap on colectomies.
Genentech, Inc.		
1.	The use of conventional therapy (CT), defined as corticosteroids for induction followed by azathioprine or mercaptopurine, as a baseline comparator is not appropriate in moderate-to-severe UC. American Gastroenterological Association (AGA) and American College of Gastroenterology (ACG) guidelines cite a stronger grade of evidence for the use of targeted immune modulators (TIM) in the induction and remission of moderate-to-severe UC population over CT. Instead, we recommend following a guidelines-based approach supplemented by real-world practice, of which infliximab and adalimumab have most prevalent usage. Adopting this approach will ensure that the base case represents the appropriate relative value of treatment in the moderate-severe UC population.	As the first evaluation of this class of medications in UC, we chose to compare all TIMs to conventional treatment to assess the relative cost effectiveness of the entire class. Because adalimumab and infliximab are entrenched in clinical practice, we also provided comparisons of TIMs to infliximab and adalimumab in a biologic-naïve and biologic-experienced population, respectively

2.	We recommend ICER update the definition of CT to denote the placebo arm of RCTs rather than the current definition of "corticosteroids for induction followed by maintenance with 50:50 split of azathioprine and mercaptopurine." The current CT definition is flawed, as there is significant heterogeneity within the placebo arms of each RCT, and thus may not accurately reflect ICER's definition of CT. For example, the range of corticosteroid use in the baseline placebo group from UC clinical trials ranges from 38.9% in GEMINI I to 65.3% in PURSUIT-SC. The use of immunomodulators such as thiopurines in placebo is reported as 12.1% in GEMINI I to 43.8% in ACT I. Some trials, such as the OCTAVE I, II, and SUSTAIN cohorts, do not report immunomodulator use. As such, we encourage ICER to further acknowledge the uncertainty of using CT as a comparator due to its high heterogeneity. This will ensure the report provides the relevant context that is necessary for correct interpretation of any conclusions or comparisons made concerning the usage of CT in UC.	Thank you for the comment. The description as noted in quotes was purely for costing purposes in the model. Effectiveness estimates were driven by the placebo arms of the trials. We have clarified the description of conventional therapy to acknowledge its variability in both the overall evidence discussion and controversies and uncertainties sections.
3.	The justification for a surgery cap references a small study sample of 369 patients within one county in the United States. As such, we caution on its generalizability to a national real-world population. Furthermore, it is unclear from the reference whether these patients are moderate-to-severe UC patients or inclusive of all severity states, which may underestimate the true cumulative rate of surgery. By capping surgery rates, this potentially biases the outcomes for weaker comparators as they may have a higher cost and disutility burden of surgery that is removed from the model due to the artificial cap. Therefore, to improve model transparency and show the impact of this artificial construct, we recommend a scenario analysis removing the cap limiting colectomy within the model.	We have revised the model to remove the artificial cap on colectomy in the base case.
4.	Per the International Society for Pharmacoeconomics and Outcomes (ISPOR) Good Research Practices Task force, it is recommended that parameter estimates follow evidence-based medicine principles, e.g., "seek to incorporate all evidence, rather than selectively picking a single source." Current adverse event rates for TIMs within the model are selected from a single source using an absolute incidence, which could bias the disutility and costs of certain TIMs due to the heterogeneity of each RCT. A more robust approach would be to utilize relative risks and adjusted probabilities from ICER's NMA. For example, a recent NMA by Singh et al. and Jairath et al. found significant differences in serious adverse events and infection rates of TIMs.	The meta-analysis by Singh provides a good source of estimates for rate of serious infection across trials based on an NMA. In this analysis, the authors conclude that no agent was significantly different than placebo in the rate of serious adverse events. The rate of serious infection was too low to be deemed appropriate to conduct a meta-analysis. Although some statistical differences among agents were found, the absolute difference in risk is very low and does not impact the conclusions of the cost-effectiveness evaluation.

Janssen Pharmaceuticals		
1.	The NMA should appropriately account for differences between re-randomized maintenance placebo rates across trials. These placebo rates are statistically different across re-randomized maintenance trials. The differences observed can be explained by the varying rates of carry-over effects patients experience when they receive active treatment in induction trials, and are subsequently re-randomized to receive placebo in maintenance trials. (Figure 1: "Placebo" rates across re-randomized maintenance arms in the maintenance phase) Therefore, placebo is not a common comparator as reported across re-randomized maintenance trials, and cannot appropriately be used to conduct an NMA without further adjustment.	As noted previously, our initial efforts suggested that including a placebo adjustment did not improve model fit, and so results were not adjusted. We revisited this and, based on the observed variation in placebo response during maintenance that is noted here, we applied a placebo adjustment for the biologic-naïve maintenance population. Such an adjustment was not feasible for the biologic-experienced maintenance population due to the much smaller size of the network, however, so we could not apply a placebo adjustment in this case.
2.	Notably, placebo efficacy rates in the Stelara UNIFI re-randomized maintenance trial are significantly higher than the placebo efficacy rates reported in other re-randomized maintenance trials. (Figure 1) This is due to the prolonged efficacy of Stelara following a single intravenous (IV) dose during the induction trial. All patients entering the maintenance trial received Stelara during the induction trial, the effect of which carries over into the maintenance trial, and continues to benefit patients who are re-randomized to placebo in the maintenance trial. Evidence of this carryover effect at a molecular level was presented by Li et al. This difference in placebo efficacy across re-randomized maintenance trials cannot be attributed to heterogeneity across patient populations in the trials since clinical response for placebo arms at end of induction were similar across trials.	We note that the "carry-over" hypothesis for ustekinumab has been presented in other HTA settings. The NICE evidence review group, for example, was unconvinced that this differed according to TIM, and cited results of other clinical reviews suggesting that placebo effects were likely to be multifactorial and manifested primarily in differences in population characteristics and other prognostic factors across trials. Please see: • https://www.nice.org.uk/guidance/ta633/evidence/appraisal-consultation-committee-papers-pdf-8774360605, pp. 399-400.
3.	The methods ICER has employed for this analysis result in: 1. Substantial underestimation of efficacy for Stelara in ICER's maintenance NMA 2. Overestimation of efficacy for conventional treatment, which particularly impacts clinical response in maintenance trials 3. Overestimation of efficacy for products that have a more rapid decrease in efficacy following induction (i.e., treatments which lack a carry- over effect). This particularly impacts clinical response in maintenance trials.	The results of our placebo-adjusted NMA in the maintenance biologic-naïve population have somewhat shifted the rankings among TIMs and resulted in nominally higher risk ratios versus placebo but have not changed the nature of between-TIM comparisons, which remain largely nonsignificant.

4.	Placebo arms across the UC maintenance trials are not a common comparator. Examining these trials independent of induction trials and converting treat-through maintenance trials to randomized withdrawal trials results in flawed NMA results. ICER's analysis violates the homogeneity assumption, rendering it invalid, and making it contrary to guidelines issued by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) for how to appropriately conduct indirect treatment comparisons. Welty et al. NMA: An NMA with Janssen authors recently published by Welty et al, utilized alternative methodology that attempted to correct for the variation in placebo efficacy rates across maintenance trial.	Strict adherence to the homogeneity assumption would result in the conduct of zero NMAs. Instead, responsible adjustment where warranted as well as conduct of multiple sensitivity analyses can provide context around findings. There is published precedent for the use of our approach converting the minority of treat-through studies to re-randomized. Please see: • https://bmjopengastro.bmj.com/content/6/1/e000302.info • https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0165435 We also note criticisms of the approach described in Welty et al. (conversion of re-randomized to treat-through designs), including significant data imputation requirements and mention by one HTA body that they were "unable to verify whether the methodology had been correctly and reasonably applied." Please see: • https://www.nice.org.uk/guidance/ta633/evidence/appraisal-consultation-committee-papers-pdf-8774360605, p. 410.
5.	The SSR Health net cost assumptions used in ICER's base-case cost-effectiveness model does not account for rebates that differ between Stelara doses used for UC (90 mg syringe) and those used in other indications (45 mg syringe). Additionally, SSR Health net prices lag behind those relevant in the current market.	SSR Health calculates both WAC and net price changes at the strength/form level and aggregates up to the product level. In the absence of data from manufacturers about their net prices, SSR Health is the best available source for net prices of non-infused products.
6.	Limitations of QALYs in assessing product value in UC: Janssen fundamentally disagrees with the use of quality adjusted life year (QALY) alone to determine the value of a medication for patients and the healthcare system, as QALY rates the value of human life relative to a subjective standard of perfect health. QALYs and cost/QALY results are highly sensitive to utility assumptions which are themselves subjective measures of health and may not adequately capture variability in UC disease severity. This is evident from the results showing low QALYs gained relative to conventional treatment and costs/QALY vs. conventional therapy that are higher than the $150K/QALY threshold for all advanced therapies.	ICER follows common academic and HTA standards by using the cost per QALY gained, but also presents cost per life year gained and cost per evLYG. The QALY is the gold standard for measuring how well a medical treatment improves and lengthens patients' lives and has served as a fundamental component of cost-effectiveness analyses in the US and around the world for more than 30 years.

Pfizer Inc.		
1.	We appreciate ICER's choice to compare the TIMs to Adalimumab in biologic-experienced patients in their base case analysis. However, ICER's threshold analysis/conclusions on cost effectiveness is based mostly on the analyses of TIMs compared to conventional treatment (corticosteroids, mercaptopurine, azathioprine) alone. We recommend the selection of conventional treatment in the bio-experienced population as a comparator. Cost-effectiveness of new interventions should be made from comparison to standard of care based on clinical guidelines. It is debatable whether conventional treatment alone is considered standard of care for patients that already failed a biologic. Instead, we believe that long-established TIMs, like Adalimumab, represent standard of care for these patients.	As the first evaluation of this class of medications in UC, we chose to compare all TIMs to conventional treatment to assess the relative cost effectiveness of the entire class. Because adalimumab and infliximab are entrenched in clinical practice, we also provided comparisons of TIMs to infliximab and adalimumab in a biologic-naïve and biologic-experienced population, respectively
2.	To substantiate our request, we want to highlight recommendations from the latest American College of Gastroenterology (ACG) guidelines. In the key concept statement section ACG specifically identifies Vedolizumab, and Tofacitinib (Ustekinumab is not included in the guidelines because it was not yet approved for moderate to severe UC) as recommended treatment options for induction and maintenance of remission in patients who fail a TNF blocker (pasted below). "In patients with moderately to severely active UC who have previously failed anti-TNF therapy, we recommend vedolizumab for induction of remission (strong recommendation, moderate quality of evidence)" "In patients with moderately to severely active UC who have previously failed anti-TNF therapy, we recommend tofacitinib for induction of remission (strong recommendation, moderate quality of evidence)."	We are unclear what the request is. Comparisons are made for each TIM to both conventional treatment and adalimumab in the biologic-experienced population to address the very concerns raised by this comment. We have, however, included additional text in our summary of ACG guidelines to address the concepts raised for vedolizumab and tofacitinib here.
3.	ACG makes no mention of steroids and or 5-ASA in moderate to severe UC patients who have failed a TNF blocker. It is important to note that the guidelines do not state that azathioprine or 6-monotherapy for the maintenance of remission are appropriate therapeutic options for patients who have failed TNF blockers. In addition, systemic steroids are not recommended for maintenance of remission either. Also, budesonide MMX has not been studied for maintenance of remission of previously moderately to severely active UC. Finally, treatment with 5-ASA therapy has been shown to be efficacious and safe as monotherapy for induction of moderately but not severely active UC, therefore, it may not be appropriate to consider 5-ASA therapy in moderate to severe UC patients.	As above, we have revised our description of the ACG guidelines.

4.	We strongly recommend ICER to compare the TIMs to Adalimumab only in biologic experienced patients and estimate economically justifiable price of TIMs versus Adalimumab (and not conventional therapy). At the minimum, we recommend that ICER include in their "limitations" discussion about the uncertainty of whether standard of care should be represented by commonly used TIM (Adalimumab) or conventional treatment given the US guidelines recommended use of TIMs in bio-experienced population.	We have clarified our approach and described the limitations further in Section 5.4. Given that this is the first ICER review in UC, there is a need to understand the value of all products, informed by the comparisons made in pivotal clinical trials (i.e., to conventional therapy); comparisons to common TIMs are also included to provide additional context.
5.	We support ICER's decision to use SSR Health pricing for pharmacy benefit products (orals, self-injectables), since there is no publicly available source for net price. However, for products dispensed through medical benefit (infusions), e.g. Infliximab, Infliximab-dyyb, Infliximab-abda, CMS published ASP values are the most transparent source of net pricing, since these include discounts to providers and reflects how payers reimburse as well. ASP values for Medical benefit products are updated every quarter and publicly available.	We have revised the model so that pricing for IV-infused products are based on Medicare ASP Payment Allowance Limits for Medicare Part B Drugs.
6.	In this case, we strongly recommend using the actual price data from CMS rather than estimating price using SSR data. ASP is a market-based price that reflects the weighted average of all manufacturer sales prices and includes all rebates and discounts that are privately negotiated between manufacturers and purchasers (with the exception of Medicaid and certain federal discounts and rebates). This methodology mirrors reimbursement for physician administered drugs in the commercial market. A manufacturer's ASP must be calculated by the manufacturer every calendar quarter and submitted to CMS within 30 days of the close of the quarter. Since ASP is accurate enough to be the official determinant physician's actual reimbursement, it should also be the basis for ICER's assessment of cost-effectiveness. Below are the ASP and SSR based cost tables that we propose to be used to replace the SSR values in the cost effectiveness analysis. We have provided costs for all infliximab molecules for transparency and completeness.	Please see our above response.
7.	We observed that mortality resulting from UC without colorectal cancer (CRC) was higher than UC with CRC. This can be observed from the mortality estimates in the "Inputs sheet" of the model where the estimates in cells P357 to P457 (average mortality rates for UC without CRC) are higher than those in cells M357 to M457 (average mortality rates for UC with CRC). We believe mortality resulting from UC with CRC should be higher than mortality resulting from UC without CRC.	We thank you for the comment. Column M should be specifically labeled as colorectal cancer related mortality in colorectal cancer rather than mortality rates for UC with colorectal cancer.

8.	We do not believe the formula for discounting costs and benefits implemented in the model is correct. The correct formula (and widely used) is the following: $$\frac{(Undiscounted\ costs\ or\ benefits)}{\left(1+(Discount\ Rate)\right)^{Time}}$$	We thank you for the comment. Although the equation used in the model is expressed slightly differently, we have confirmed that both lead to the same numeric results.
9.	The model structure assumes that patients that discontinue their assigned TIM are switched to a second TIM, represented by a market basket of other treatment options. While multiple biologics are used following treatment failure and the inclusion of a second TIM would be reflective of clinical practice, the inclusion of subsequent treatments introduces additional uncertainty into interpreting model results. Inclusion of a second TIM convolutes the cost-effectiveness ratio of the initial TIM because the benefit and costs are attributed in part to the second TIM. The ICER analysis as it stands therefore deviates from the intended health-economic question about the cost-effectiveness of the initial TIM, and instead estimates the cost-effectiveness of the treatment sequence.	As stated in a prior response, the scope of this review is not intended to evaluate the cost effectiveness of specific sequences. Instead, we have equalized the market basket of subsequent treatment across all TIMs and conventional treatment to reduce any bias introduced by later line treatments.
10.	The basket of second TIMs modelled also possesses some limitations due to limitations in data, which further introduces uncertainty in the cost-effectiveness ratios of the initial TIM. For example, in the draft evidence report on page 95, it states that, "No RCT data was identified for infliximab or infliximab biosimilars in a biologic experienced population, so these therapies were excluded from the market basket of second TIMs…" These assumptions strongly influence the cost-effectiveness ratio. If there is uncertainty in the 2L TIM basket, then this translates to uncertainty in the final cost-effectiveness ratio. You can see this effect in Table 5.40 of the report. Replacing the 2LTIM basket with just infliximab lowered the cost-effectiveness ratio by 18%.	Subsequent treatment does have an impact on the cost effectiveness of the initial TIM. In order to maintain focus on the initial TIM and reduce unintended bias from the market basket of subsequent treatment, we have revised the model so that each TIM and conventional treatment have the same subsequent treatment market basket. Note that the scenario involving infliximab as the second-line treatment also employs efficacy data from a biologic-naïve population, which may reflect more favorable outcomes than if data were available for infliximab in a biologic-experienced population, making this overall strategy appear more cost effective.
11.	We also note that because of some of these limitations, the results of the base case analysis are inconsistent with the results of the EJP, which excludes the second TIM in its calculation. We recommend removing the second TIM from the base case analysis or at the very least running a scenario analysis that excludes the second TIM. Alternately, we also recommend addressing the uncertainty created by the inclusion of the second TIM in the limitations section.	In the shift to allow the same subsequent treatment basket across all comparators, including conventional treatment, we have effectively mitigated this issue. EJPs will be calculated for the base case without removing the option for second TIM.

12.	For the biologic-experienced population, the scenario analysis of dose escalation showed the cost per QALY gained for tofacitinib increased (Table 5.32 of ICER's report). This is surprising because the cost of 10 mg and 5 mg are the same (flat pricing) and we would expect the overall cost to remain the same.	While tofacitinib has flat pricing, some patients who initiate tofacitinib will discontinue and move to subsequent treatment, which contains a market basket of TIMs that do have additional costs for dose escalation. Thus, the cost of the initial TIM (tofacitinib) remains the same but the overall cost of the treatment strategy in increased.
Takeda Pharmaceuticals		
1.	Given that the health state utilities (HSU) associated with the active disease state have by far the greatest impact on the final cost per QALY in the draft ICER report, it is important to ensure that these estimates are valid and reliable for meaningful interpretation of results.	Based on commends received, we have reviewed the literature and revised our choice of utility values.
2.	ICER's univariate sensitivity analyses upper bound range for active disease HSU (0.814) results in a higher HSU value than the base case HSU value for the response health state (0.727), undermining the value of obtaining clinical successes with treatment.	Upper and lower bounds for utilities are taken from 95% confidence intervals reported for these point estimates. If the range is wide and the input parameter is impactful, the one-way sensitivity analysis will appropriately characterize the uncertainty of this parameter on the resulting incremental cost-effectiveness ratio for each TIM.
3.	Malinowski and Kawalec (M&K) should not be used as the base case source for HSU values, as it lacks robust validity in several ways, making it inappropriate for use: • M&K do not explicitly state how they aggregated studies with different health states. • M&K included 'active' as a health state but 'active' disease as defined in the included publications do not align with the definition of the population to be modeled by ICER, "moderate-to-severe UC being treated with TIMs." • M&K did not include the most cited sources for HSU values in their meta-analysis (MA), some of which were available to them at the time of their analysis. • Face validity is limited. Moderate-severely active disease HSU (0.7977) is greater than the HSU for moderate (0.6969) and mild (0.7834) disease. Active disease HSU (0.6992) also has a lower utility preference than moderate-severe (0.7977) and severe (0.7059). Normally, an approximate linear relationship between severity and utilities is to be expected. Higher quality of life for those with more severe disease highlights the danger of combining studies with differing methodologies. • M&K state their MA revealed a significant heterogeneity among studies and warn that their results should be interpreted with caution. Cochrane's Q test and I illustrate this significant heterogeneity.	As stated in prior responses, we have reviewed the literature and revised our choice of utility values to best reflect the moderate-to-severe UC population.

4.	The limitations of the M&K study make it inappropriate as the base case utility values. We therefore recommend that ICER conduct an MA to generate HSU using inclusion criteria that would better reflect the population being modeled.	See prior response.
5.	If conducting a targeted MA is not an option, it may also be appropriate to use HSU values from Woehl et al. or Arseneau et al. for the base case and use M&K (and/or other updated sources) as a scenario analysis. This reflects a more appropriate population and would decrease uncertainty around interpretation of the results.	See prior response.
6.	While the therapeutic area is still developing, therapeutic goals have evolved from relief of IBD-related symptoms to the "treat-to-target" approach of endoscopic improvement. Long-term studies are particularly effective in identifying durability of treatment. A network meta-analysis (NMA) conducted in 2017 explored endoscopic improvement data separately over the short-term, using eight short-term studies, and after longer follow-up using five long-term studies.	Thank you for the comment. We evaluated endoscopic improvement in available trials as a secondary outcome of interest and report these results descriptively for all trial phases and using an NMA for the induction phase. As stated in the report, differences in trial design and limited available data precluded maintenance NMAs on this outcome. Finally, given that its measurement is also included as a criterion for response and remission, we are unsure of the added value of attempting a separate long-term NMA.
7.	Because of the lack of optimal incorporation of long-term mucosal healing data, overall cost-savings associated with mucosal healing's clinical and economic benefit, such as colectomy-free survival and reduced colectomy, are undermined in the ICER model. The ICER evaluation anchors analysis on RCTs and thus does not adequately capture long-term real-world outcomes.	The ICER model is anchored on RCTs in order to create a robust comparative effectiveness analysis with minimal bias. We agree that mucosal healing may represent an important outcome that is partially captured within our remission health state (as mucosal healing is part of the remission criteria). Additional research is needed to isolate the specific impact of mucosal healing above our assumptions for remission outcomes.
8.	A similar NMA augmented by the VARSITY trial's long-term endoscopic healing outcome scores should be performed to ensure that the ICER analysis accurately reflects long-term endoscopic/mucosal healing. Alternatively, both the VICTORY consortium data as well as the EVOLVE study should be considered for inclusion in ICER's evaluation to supplement the current deficiency in long-term clinical projections for vedolizumab.	Please see Comment 6 above. In addition, we found no induction data from VARSITY that was stratified by prior biologic experience. Regarding the VICTORY and EVOLVE observational studies, the former did not meet our sample size criteria for comparative study (>500 UC enrollees). Data from the EVOLVE study were identified in our updated search and have now been included in the revised report.

9.	In the base case analysis, adalimumab is the sole TNFα therapy included for 2nd line following vedolizumab, ustekinumab, or tofacitinib, despite its lack of indication in TNFα failure population, poorer efficacy and higher cost than some alternatives. This is contradictory to real-world data that indicates patients often switch to other TNFαs, as evidenced by some plans which require a double step of two non-TNFα prior to switching to biologic therapies, leading to malapportioned 2nd line market baskets. Likely the exclusion of infliximab from 2nd line will introduce bias and produce inflated cost per QALY estimates for all other products.	As stated in prior responses, subsequent treatment does have an impact on cost effectiveness of the initial TIM. In order to maintain focus on the initial TIM and reduce unintended bias from the market basket of subsequent treatment, we have revised the model so that each TIM and conventional treatment have the same subsequent treatment market basket.
10.	ICER leverages the NMA of available multinational randomized clinical trials, as well as clinical trials with local/regional populations. Including regional trials in the base case model analysis could introduce bias in the meaningful interpretation of data for the US population.	We excluded studies conducted solely in Asian populations in a sensitivity analysis, with results generally concordant with our base-case findings; in addition, such exclusions would have reduced the network size and statistical power further. We also consulted with clinical experts, who indicated that there was no clear rationale for such exclusion in UC studies.
11.	Jiang 2015, Motoya 2017, Suzuki 2014, Kobayashi 2016 and NCT01551290 trials introduce variance in population outcomes through underlying patient demographics. Diagnosis criteria, treatment patterns, background treatment and therapeutic access are different from the US healthcare system. In addition, this could be through predisposition to comorbid conditions, nutritional intake, or control on AE reporting which is illustrated by the AE of tuberculosis featured in Suzuki 2014 and the lack of any serious infections in any of their associated trials. The ICER report specifically states that imbalances in key disease characteristics in the Motoya et al. RCT population bias results in placebo favorability, especially in the biologic-experienced group, thus potentially diminishing the relative efficacy of vedolizumab. With major study limitations and heterogeneity in study design, including this study in the NMA risks invalidating its interpretation.	Please see the comment above. In addition, diagnosis, treatment, and therapy access differ from the US in almost any setting, including Europe. We believe our conduct of a sensitivity analysis excluding these studies is an appropriate response to any possible systematic differences.
12.	We suggest that the base case analysis in both the NMA and the cost-effectiveness model focus on high-quality and appropriately generalizable evidence, including well-designed trials in populations that align with ICER's target population, and real-world studies in the US healthcare system. Clinical trials with unbalanced arms or with population characteristics that limit appropriate comparison should be reserved for sensitivity analysis.	Please see comments above.

13.	When evaluating the induction clinical outcomes of vedolizumab, week 6 data from the GEMINI trial should be replaced with week 14 data in the base case induction NMA. In comparison to week 6 data that is currently used, week 14 data more appropriately matched the clinically relevant assessment period for vedolizumab, and still falls within the timepoint range (6-14 weeks) established for the base case induction period.	GEMINI 1 was designed to measure differences at an induction timepoint of six weeks. Our general approach across RCTs was to focus on induction timepoints that were declared as primary or key secondary, and where total Mayo score was calculable (i.e., endoscopy was performed).
14.	Using week 14 GEMINI data in place of week 6 data will align most appropriately with vedolizumab induction usage while remaining consistent with ICER's stated approach.	Please see comment above.
15.	Although ICER compared each therapy to "conventional" non-biologic therapy, it may no longer be considered clinically appropriate for long-term use in patients with moderate to severe UC. Long-term treatment over a year or more with conventional therapy such as steroids are known to cause serious adverse events that require additional medical attention and costs that are not accounted for in the ICER model. Follow the 2019 ACG14 and the 2020 AGA guidelines recommendations of using advanced therapy for the management of moderate to severe UC to reflect reference comparator.	Our synthesis of the evidence and NMA relied on the comparators chosen for the clinical trials and observational studies of interest; for the former, conventional therapy (placebo) was the comparator in all but one trial.
16.	Net prices estimated from SSR database cannot be validated, are often inaccurate and do not reflect real-world costs of drugs in UC paid by various plans. Cost-effectiveness analysis based on such strong assumptions possess enormous uncertainties and conclusions from CEA models cannot be generalized and can be mis-leading. ICER should be transparent that this is a major limitation when making conclusions of comparative cost-effectiveness in the report and be extremely cautious in making policy recommendations based on such analysis.	Please see prior responses. We also note that the drug price was extensively tested in sensitivity analyses for all TIMs and was relatively insensitive to results in comparison to utility and efficacy estimates.
17.	Vedolizumab indication is stated incorrectly (Section 3.1, page 16). Vedolizumab is indicated for adult patients with moderately to severely active UC, similar to ustekinumab. FDA sent Takeda a Complete Response Letter (CRL) in response to Takeda's Biologics License Agreement (BLA) supporting the use of subcutaneous (SC) vedolizumab as maintenance treatment in patients with moderately to severely active UC. Takeda is assessing the details of the CRL, gathering information needed to resolve the FDA's questions, and will work closely with the FDA on a path to approval.	We apologize for these errors; they are corrected in the revised report.

18.	Annual drug costs reported in Table 1.1 do not match those reported in Table 5.16. Golimumab recommended dosage differs between Table 1.1 and Table 5.14 (correct in Table 5.14). Ustekinumab recommended dosage differs between Table 1.1 and Table 5.14 (correct in Table 1.1). Table 4.20 – Tofacitinib and ustekinumab estimates for AEs leading to discontinuation need to be reviewed and revised.	A typo in the recommended dosage of ustekinumab in Table 5.14 has been corrected. Misalignment between Table 1.1 and Table 5.16 has also been corrected.
Economists		
Paul Langley		
1.	When a QALY is constructed time spent in a disease state is multiplied by a utility score on a range 0 = death to 1 = perfect health. Would ICER agree that this requires the utility score to have ratio properties? Ratio property means that the measurement scale must have a true zero. This means that the EQ-5D-3L should have a true zero. In fact, EQ-5D-3L utilities can take negative values (with a range -0.59 to 1). Would ICER agree that this means the EQ-5D-3L is not a ratio scale? In respect of Q2, if ICER believes that the EQ-5D-3L instrument has, despite negative values, a true zero, could ICER provide a proof?	We (and most health economists) have the understanding that the EQ-5D (and other multi-attribute utility instruments) do have ratio properties. The EQ-5D value sets are based on time trade-off assessments (which are interval level), with preference weights assigned to different attributes. We fail to see why this should be considered as an ordinal (ranked) scale. ICER believes that the dead state represents a natural zero point on a scale of health-related quality of life. Negative utility values on the EQ-5D scale represent states considered worse than dead.
2.	It appears to be commonly assumed that the EQ-5D-3L (in common with other generic instruments) meets the axioms for invariance of comparisons. That is, it has interval scoring properties. Would ICER agree?	The EQ-5D multi-attribute utility function is designed so that a utility difference of 0.05 is considered equivalent regardless of the starting point.
3.	ICER cannot provide a proof that the EQ-5D-3L has interval properties, how does ICER justify the creation of QALYs as responses to therapy?	Please see above responses.
4.	Over the past 20 years commentaries from measurement theory specialists have made the case that instruments such as the EQ-5D-3L, in fact the majority of patient reported outcomes instruments, have lacked ratio (and interval) properties. Is ICER aware of this literature and would ICER care to comment?	Please see above responses.
5.	If the EQ-5D-3L (or other generic utility scale) cannot be shown to have ratio (and interval) properties why does ICER persist in creating lifetime cost-per-QALY claims? As the utility scale is ordinal then the QALY is an impossible construct? Would ICER agree?	Please see above responses.
6.	If, given that the QALY is mathematically impossible, would ICER inform its audience that it recognizes this but insists that there is still merit in constructing imaginary value assessments on imaginary QALYS to create imaginary claims information.	As stated above, we do not accept the premise of this question.

Other		
Baysient		
1.	Baysient recommends that the Final Evidence Report identify trending changes to the standard of care that have the potential to alter the analytical conclusions of the report within a 24-month timeframe. These trends include the impact of: • Increasing the application of therapeutic drug monitoring (TDM) combined with CDS as the standard of care. Filling "knowledge gaps" (i.e. the relationship between administered dose and trough concentration) in current clinical practice guidelines. • Providing sufficient evidence to recommend the routine use of TDM with TNF-alpha inhibitors. • Applying better-clinical-outcome-at-lower-cost alternative payment models to individualized infliximab therapy incorporating TDM combined with CDS. • Changing the US infliximab label making it consistent with monographs that include the use of TDM with infliximab. • Providing health care providers with scientifically appropriate and statistically sound dosing information derived from the laboratory assays provided by infliximab manufacturers or other licensed laboratories.	Thank you for these comments. However, our review is focused on current evidence of differences in effectiveness, cost, and cost effectiveness among the TIMs rather than changes in dosing for a single TIM. We also cannot speculate on information on dosing that is still being processed. Should relevant data become available during an update to this topic we will certainly consider it.