INSTITUTE FOR CLINICAL AND ECONOMIC REVIEW

JAK Inhibitors and Monoclonal Antibodies for the Treatment of Atopic Dermatitis
Response to Public Comments on Draft Evidence Report

July 9, 2021

Table of Contents

#	Comment	Response/Integration
Manufacturers		
AbbVie		
1.	The net health benefit of upadacitinib in AD is not "promising but inconclusive" compared to topical therapies or “insufficient” versus dupilumab as ICER concluded. Upadacitinib, in fact, is statistically significantly more efficacious on a range of measures than either placebo or dupilumab based on double-blind, randomized controlled trials and numerous network meta analyses. The statistically significant and clinically meaningful superiority of upadacitinib versus placebo was shown in multiple randomized placebo-controlled trials (Measure Up 1, Measure Up 2, AD Up).2,3 Similarly, the superiority of upadacitinib versus dupilumab was demonstrated in a large Phase III well-designed head-to-head clinical trial (Heads Up). In addition, the long-term safety of upadacitinib has been reported through study of up to 4.5 years of data across multiple clinical trials and indications.	The net health benefit considers not just efficacy but also risks. This comment primarily focuses on the benefits but does not take into account the potential risks or harms. It is when considering both that we have arrived at the current recommendations. These are also reflected in existing black box warnings for oral JAK inhibitors as well as ongoing concerns expressed by the FDA as part of their consideration of oral JAK inhibitors including upadacitinib for this new indication.
2.	The conclusion that dupilumab dominates upadacitinib in the cost-effectiveness analysis lacks validity. One therapy dominates another in a cost-utility model when it produces more quality-adjusted life years (QALYs). Clinical efficacy is the main driver of QALYs gained, and as stated above, upadacitinib was shown to be superior to dupilumab in the head-to-head (Heads Up) clinical trial. In addition, ICER’s own network meta-analysis (NMAs) show that when considering all evidence together upadacitinib 30mg is the most efficacious therapy in eight out of nine comparisons where ICER included upadacitinib.	The benefit of drugs that treat atopic dermatitis includes the initial treatment efficacy, informed by the NMA, and continued benefit informed by the discontinuation rates. In the draft report, although patients receiving upadacitinib, had strong initial response, the discontinuation rate was higher than that for other drugs including dupilimab and therefore the total QALYs was higher for dupilimab versus upadacitinib. This, combined with the higher cost, leads to a “dominated” conclusion for upadacitinib. However, in the updated results of the multinomial NMA, the relative benefit of upadacitinib versus dupilumab is higher and the discontinuation rate that will be used will be slightly lower leading to a revised conclusion.
3.	Finally, the resulting price at which upadacitinib is found to be cost-effective is biased by the various salient clinical and methodological limitations described herein this response letter.	Thank you for your comment. We have addressed the various points you have made in this document.
4.	The 1% discount rate applied to upadacitinib is erroneous and a major driver of high total cost of upadacitinib treatment in ICER’s assessment. The SSR Health data source should not be used to estimate discounts to WAC prices, as clearly demonstrated by the evidence shared in the later section of this response.	We reviewed the data we received from SSR Health information and confirmed that we did not use the quarters of data where the net pricing in SSR Health was above that of the WAC pricing in

		estimating the net price and corresponding discount rates.
5.	The discontinuation rate for upadacitinib in the ICER's assessment are obtained from a Japanese rheumatoid arthritis study that is not representative of the atopic dermatitis population. AbbVie has recently provided confidential data to ICER from the Heads Up trial that show no difference in discontinuation rates for upadacitinib vs. dupilumab. Consistent with evidence in atopic dermatitis, the same discontinuation rate applied to dupilumab should also be applied to upadacitinib.	Thank you for providing the additional discontinuation data. However, the data provided is different than what was requested and that which was provided by other manufacturers. These data only reflect the initial 16-week trial period whereas the parameter is meant to reflect discontinuation in the weeks and years after the initial 16-week period on therapy. In the absence of these key data provided on the discontinuation rate for responders after 16 weeks, we will conservatively assume a rate equal to the highest rate within the class.
6.	ICER's assessment did not capture any economic savings of achieving EASI 90 vs 75 vs 50 scores in the cost-utility analysis. We have provided compelling evidence to ICER that is corroborated by previously published literature to show substantial cost savings associated with higher EASI scores that should be included in the final assessment.	Thank you for providing these data. Although these data have not yet been peer reviewed, we have reviewed the methodology provided by AbbVie and have incorporated the data into the model.
7.	Methodological errors in ICER's NMA, including the omission of key clinical trial data, affected resulting transition probabilities generated for use in ICER's cost-effectiveness model. These incorrect transition probabilities for upadacitinib underestimate the true clinical benefit to patients relative to all other comparators in the model. This error primarily affects the EASI- 50 and EASI-90 transition probabilities for upadacitinib.	Thank you for your feedback. We have revised and updated the NMA for the revised version of the report.
8.	ICER's base case cost-utility analysis does not fully capture the patient value of AD treatments (e.g., improvements in sleep, itch). It also does not capture any benefits of work productivity improvements or of EASI score improvements of less than 50. All these exclusions suggest a substantial portion of the value of AD treatments such as upadacitinib is not reflected in the economically justifiable price calculations or base case cost-utility analysis.	Work productivity is included in the societal perspective in accordance with the ICER value framework and measures of other clinical benefits (including improvements in sleep, itch, anxiety/depression) are provided in the main report for therapies providing PRO data by EASI score. If these data are available for upadacitinib (to capture the patient-level impact), we invite you to provide said data to ICER to be considered in the analyses. Additionally, available evidence supports the assumption that gains in other areas (sleep, itch, etc.) were correlated with EASI score, and that EASI score was therefore an acceptable

		measure of patient improvement for the base case cost-utility analysis.
9.	Topical emollients are not standard of care for moderate to severe AD patients but rather supportive care at best, as assumed in the cost-utility and cost-consequence analyses. The terminology of SoC should be changed to supportive care for moderate to severe AD patients.	Thank you for your comment. We think that our referring to topical emollients as standard of care is clear, guideline recommended, and understandable to readers.
10.	NMA methods are not described sufficiently, and important Phase 3 trial data provided to ICER by AbbVie (e.g., Heads Up, the head-to-head Phase 3 trial of upadacitinib vs. dupilumab) are omitted from the analysis.	Thank you for your comments. We have added additional information detailing the NMA methodology in the revised version of the report.
11.	We are also deeply disappointed with ICER making public our confidentiality provided data to support the AD assessment. Specifically, on p. 312 of the Draft Evidence Report, ICER did not redact the productivity data reported in Table E4.1 that we provided them from the Measure Up 1 and Measure Up 2. This is a violation of ICER's guidelines that state, "[a]cademic-in-confidence data will be redacted from all external and public ICER documents until the earlier of: (a) publication or presentation of such data by the data owner or study investigators; (b) 18 months following the date of the public ICER meeting; (c) for reports that are not subject to a public meeting, 18 months following report publication."4 This carelessness adds to the challenges manufacturers such as AbbVie face when working with ICER to help improve the overall quality of the assessments being undertaken across the various therapeutic areas.	We disagree that this is a violation of our policy. Academic-in-confidence data are redacted to avoid interfering with publication of results. These are calculated values and not direct inputs from the data AbbVie has provided. We do not believe that our publishing these calculated values will interfere with publication of the underlying data. ICER has updated its in-confidence data policy to make this even clearer.
Eli Lilly		
12.	Study Inclusion and Dosing Information: The BREEZE-AD1, BREEZE-AD2, and BREEZE-AD7 clinical trials studying baricitinib in patients with atopic dermatitis include only patients outside of North America and are not representative of a US patient population with moderate to severe atopic dermatitis.[1-3] BREEZE-AD5 is a North American study that best represents the US population.[4] Lilly applauds ICER for highlighting that the 4 mg dose of baricitinib will not be available in the US. Lilly submitted data on the lowest efficacious dose of baricitinib in atopic dermatitis to the FDA at 2 mg.[1-4] Of equal importance, baricitinib 1 mg was studied in clinical trials per regulatory guidance, and this dose will be intended for patients with renal impairment who are unable to take the baricitinib 2 mg dose should baricitinib be approved for the treatment atopic dermatitis. This would be consistent with the current labeling for Olumiant in Rheumatoid Arthritis.[5]	We included all key pivotal efficacy trials of baricitinib in our analyses, including Breeze-AD1, 2 and 7. Comparing Breeze-AD1 and 2 with Breeze-AD5, we did not observe important differences in study design, patient characteristics, or outcomes assessed to suggest that we should not include certain trials in our analyses. We have added additional language in the revised report in Section 3.2 on Uncertainty and Controversies to specify that the 1 mg dose maybe intended for patients with renal impairment using dosing information that is consistent with approved labeling for use in rheumatoid arthritis.

	Lilly Recommendations: • ICER should provide detail on the geographic locations of clinical trials in their reports to allow readers to understand and interpret the patient populations assessed in each clinical trial. Specifically, inclusion of this detail in Table 3.1 or in Table D3.2 is preferred. • ICER should evaluate only FDA-approved doses for the interventions identified within the final assessment. • ICER should state that the 1 mg dose of baricitinib will be intended for patients with renal impairment who are unable to take the baricitinib 2 mg dose consistent with the labeling for Olumiant in Rheumatoid Arthritis in Section 3.2.	There are many characteristics of clinical trials that can be described in a systematic review. We have chosen the ones we felt would be most important to those wishing to understand the underlying trials. Of course the authors of the underlying trials can choose to highlight whatever characteristics they feel are most important.
13.	Outcomes: BREEZE-AD1, BREEZE-AD2, BREEZE-AD5 all investigated patient reported outcome (PRO) measures that are important symptoms of Atopic Dermatitis and important aspects of the impact of Atopic Dermatitis on patients. Key PRO measures in the trials included but were not limited to the following: itch severity (Itch Numeric Rating Scale (NRS)), skin pain for example discomfort or soreness (Skin Pain NRS), night-time awakenings due to itch (Atopic Dermatitis Sleep Scale item 2 (ADSS-2)), quality of life (DLQI, WPAI), anxiety and depression (HADS-Anxiety, HADS-Depression). Data can be found in publications, clinicaltrials.gov, as well as in data submitted to ICER during the data request period earlier this year.[2-4,6-8] For quick reference, Lilly has provided a summary of the publicly available relevant data in Appendix 1 to this Public Comment. Lilly Recommendations: • ICER should recognize additional outcomes data important to patients in the discussion of baricitinib clinical effectiveness in Section 3.2 including itch (Itch NRS), night-time awakenings due to itch (ADSS-2), skin pain (Skin Pain NRS), work productivity (WPAI), and anxiety and depression (HADS-Anxiety, HADS-Depression). [2-4,6-8] ICER should reach out to Lilly if they have difficulty identifying this information in the submissions provided and referenced in this document in Appendix 1. • ICER should at a minimum include the impact of baricitinib 2 mg on all PRO measures (e.g., itch, night-time awakenings due to itch, skin pain) as a part of the Potential Other Benefits section of the	We agree that patient reported outcomes represent important outcomes for atopic dermatitis. We have reviewed these outcomes for patients enrolled in Breeze-AD1, 2 and 5. We are unable to incorporate the PRO measures provided for baricitinib in the economic model, however, as they are not disaggregated by EASI score.

	assessment as these endpoints are important to patients and help to inform prescribing behavior.	
14.	Subgroup Analyses and Heterogeneity: On page 28 in the Disease Severity section, ICER states that baricitinib has qualitatively better outcomes in patients with severe disease compared to those with moderate disease. Lilly's data submissions support baricitinib efficacy in both moderate and severe patients, however, based on analyses of both IGA3 vs. IGA4, and body surface area involvement, the efficacy is qualitatively better in patients with moderate disease. Body surface area (BSA) is a tool utilized in dermatologic disease states to quickly and easily assess the extent of disease in clinical practice.[9] In light of the clinical utility of BSA, it would be valuable to evaluate the baricitinib 2 mg data within this subgroup. The mean affected BSA at baseline in our studies ranged from ~40% to ~50%.[2-4] Post-hoc analyses showed that ~90% of the EASI75 responders, and ~95% of patients achieving a score of 0 or 1 (clear or almost clear) with the validated Investigator Global Assessment for AD (vIGA-AD™) scale, had a baseline BSA between 10-50%.[10-12] Patients who responded to baricitinib 2 mg showed a clinically meaningful improvement in skin inflammation (50% improvement from baseline in affected BSA) and itch (at least a 3-point or greater improvement in the itch NRS) by week 4 and 8, allowing for early medical decision on whether patients should continue on baricitinib 2 mg therapy or not.[10,11] Lilly Recommendations: • ICER should revise their statement about baricitinib efficacy to state that while baricitinib is effective in both moderate and severe patients with atopic dermatitis, it has qualitatively better outcomes in patients with moderate disease compared with severe disease on page 28 of the report. • Due to the clinical utility of measuring BSA involvement in dermatology practice, ICER should include the baricitinib 2 mg impact on patients with BSA involvement of 10-50% within their Potential Other Benefits or Contextual Considerations section.	Thank you for your feedback. We have included disease severity in our discussion of relevant subgroups as indicated. For space reasons in the main report, this section focuses on differences among patients with moderate and severe disease defined using study eligibility criteria. Moreover, data for baricitinib stratified by baseline severity is reported in the supplemental report as academic in confidence (AIC) and not shown.
15.	**Network Meta-Analysis (NMA):** Lilly encourages ICER to honor its commitment to model transparency by providing additional detail on the NMA model parameters in their next release of the Evidence Report for Atopic Dermatitis. Specifically, Lilly would like to understand NMA model parameters such as the details of priors put on the estimates, including the between study standard deviation (SD). The NICE technical supporting	Thank you for your comments. We have added additional information detailing the NMA methodology in the revised version of the report. For the reasons cited in the report, we believe that the model parameters selected best reflect the trial data and fit. We agree that our primary analyses include monotherapy placebo-

	documents (TSD) that are referenced within the NMA section recommend that the baseline is fitted independently.[13] It is not clear in the methods section of the report if these models fit the baseline independently or simultaneously. In addition, Lilly would like to understand the between study SD, the deviance information criterion (DIC) and residual deviance for each model. Since the models are adjusted for baseline risk, Lilly would like to understand the regression coefficient for baseline risk. By adjusting for baseline risk, the model favors treatments with higher placebo response and penalizes treatments where the placebo response is low. Placebo response rates are multifactorial, and while baseline risk adjustment can be used to account for some heterogeneity in trial design, it may not account for this effect sufficiently and can potentially introduce bias. It is therefore important that the report demonstrate the reasons for adjusting for baseline risk. The NICE TSD recommends looking at several different criteria to determine if adjusting for baseline risk is necessary.[13] Lilly encourages ICER to include the following information in the Evidence Report to justify the use of the model adjusting for baseline risk: 1. Establishing whether the regression coefficient was significant by showing that the 95% credible interval (CrI) excludes 0. 2. Establishing whether the between-study standard deviation parameter (and its 95% CrI) was reduced in magnitude when adjusting for baseline risk 3. Establishing whether the DIC and the posterior residual deviance are improved when comparing with the unadjusted model 4. Plotting of the relative risk by placebo response Lilly believes that a multinomial model is more appropriate for fitting EASI response scores as the scores are categorical. Rather than fitting three separate binary models, a probit model is more appropriate. Should ICER choose to convert to this type of model, Lilly asks ICER to include details of this type of model in their methods section in the primary report or in the appendices. Finally, ICER's base-case NMA appropriately includes monotherapy clinical trials with placebo only as a common comparator. However, in a model sensitivity analysis, ICER conducted an NMA including the combination studies with the monotherapy studies in the section "Combined	controlled trials. We also provide analyses examining combination trials that also permitted the use of topical therapies, something that is commonly done in clinical practice.

	Placebo-controlled monotherapy and combination Trials in Adults (short-term)". There is not a common comparator linking these studies making it inappropriate to pool these studies. Further, in ICER's scoping document and research protocol, it lists the interventions of interest as monotherapies. The BREEZE-AD7 clinical trial, as well as Guttman-Yassky phase 2 clinical trial, are trials of baricitinib in combination with topical corticosteroids compared to placebo plus topical corticosteroids.[1,14] Therefore, the trials of the interventions in combination with topical agents compared to placebo in combination with topical agents are out of scope for this assessment and for the NMA. Lilly Recommendations: 1. ICER should provide additional detail on model parameters including the details of whether the baseline is fitted independently or simultaneously, the priors put on estimates, the between study SD, the DIC, and the residual deviance for each model. 2. ICER should include detail on the rationale and parameters to justify the use of a model that adjusts for baseline risk. In addition, ICER should highlight the heterogeneity of key trial criteria that justify the use of this type of model in the Contextual Considerations section of the assessment. 3. ICER should include details of the multinomial model structure in the methods section of the assessment. 4. ICER should keep the base-case NMA using placebo-controlled trials only, and not include the sensitivity analysis NMA with placebo plus topical agents as a comparator to keep consistent with the scope and because there is not a common comparator.	
16.	**Comparative Value Analysis:** Utility Values: Lilly is aligned with ICER's approach to use a pooled utility estimate across therapies to give the most robust understanding of the utility of achieving a clinical response for defined health states based on EASI scores. In Table 4.4, the BREEZE-AD clinical trials for baricitinib are not included in the pooled utility estimate. Because the BREEZE-AD5 utility response rates represent a North American population of atopic dermatitis patients, these utility values would be the most representative of the US patient utility for achieving a clinical response in atopic dermatitis and could be applied across all interventions included in the assessment or pooled with the estimates from other	We included all key pivotal efficacy trials of baricitinib in our analyses, including Breeze-AD1 and 2. Comparing Breeze-AD1 and 2 with Breeze-AD5, we did not observe important differences in study design, patient characteristics, or outcomes to suggest that we should not include certain trials in our analyses. We used a pooled utility estimate approach using the data available by EASI subgroup; disaggregated utility was not provided from Lilly and was therefore not included in the pooled estimates. We did

	intervention trials. Further, additional clarity is needed on the estimates used in this assessment. It is not clear whether separate utility values are used for the moderate vs. severe health states. Lilly Recommendations: • If possible, ICER should provide the weighted averages of the utility estimates (means, standard deviations) without divulging the product or trial specific utilities that were submitted in confidence. • ICER should clarify whether different utility values are used for moderate vs. severe health states. • ICER should include the BREEZE-AD clinical trials, or specifically the BREEZE-AD5 clinical trial given its representation of US patients, in their determination of pooled utility estimates. • ICER should be as transparent as possible in the inputs and assumptions included in this assessment.	not use different utility values for moderate and severe health states, a point that has been clarified in the report.
17.	**Access and Reimbursement Considerations:** We continue to urge ICER to consider clinical, economic, and patient access implications of rebates used to negotiate formulary access in the autoimmune therapeutic class, including with respect to AD. Rebates are rarely equal for all available treatment options and negotiations can create barriers to more cost-effective therapies due to exclusions and step edits. In the autoimmune market this dynamic is known as the "rebate wall," which is an issue that has received significant attention from Congress, the FTC, and ICER itself.[15-22] Further, we encourage ICER to consider the impact of rebate walls as it examines the implications of tiering, step therapy requirements and prior authorization criteria, in its forthcoming "Barriers to Fair Access Assessment" as rebate walls can drive utilization management techniques and formulary decisions.[22] Lilly Recommendation: • ICER should encourage discussion of the implications of using rebates to negotiate formulary access during the forthcoming Roundtable Discussion and should acknowledge these potential implications in the Final Evidence Report for Atopic Dermatitis.	Thank you, we hope to discuss this further during the policy roundtable at the public meeting.
Incyte		
18.	**Recommendation for consistent nomenclature of ruxolitinib:** An oral formulation of ruxolitinib is available in the United States, however the oral formulation is not indicated, nor being evaluated, for use in patients with atopic dermatitis.	Thank you for your suggestions. We have revised reference to ruxolitinib in the revised report as suggested, that we are studying ruxolitinib cream.

	Therefore, we recommend that ***ruxolitinib cream*** is the preferred term, replacing *ruxolitinib* throughout the document. **Draft Evidence Report Text:** Page 11, Paragraph 2: *"While ruxolitinib also appeared to be more effective than a medium potency topical corticosteroid..."* **Suggested revision:** "While ruxolitinib cream also appeared to be more effective than a medium potency topical corticosteroid..."	
19.	**Recommendation to change placebo to vehicle cream:** Phase 3 clinical studies evaluated ruxolitinib cream against vehicle cream. We recommend a global change throughout the document to replace "placebo" with "vehicle cream," for accuracy and consistency.	We have updated the language used in the revised report by replacing "placebo" with "vehicle (placebo)."
20.	**Recommendation to specify safety concerns related to oral JAK inhibitors** When discussing important safety considerations of systemic JAK inhibitor therapies, we recommend the report specify *oral* JAK inhibitors consistently throughout the Evidence Report. We have identified 3 places where the change needs to be made. **Draft Evidence Report Text:** Page 10, Paragraph 3: *"Safety is an important consideration with biologic therapies and, as above there have been particular concerns about the safety of JAK inhibitors when used for other conditions"* Page 10, Paragraph 5: *"Taking into consideration the above information on short-term benefits seen in the trials but concerns about long-term safety, especially for JAK inhibitors.."* Page 32, Paragraph 3: *"In summary, for adults and adolescents with moderate-to-severe atopic dermatitis inadequately controlled with topical or systemic therapies, or for whom topical or systemic therapies are not tolerated or are medically inadvisable, we identified benefits from short-term trials of these four agents but concerns about long-term safety, especially for the JAK inhibitors"* **Suggested Revision:** Include the word "oral" preceding "JAK inhibitors" to read "oral JAK inhibitors" in all 3 abovementioned statements.	We have clarified the terminology included in the report based on your suggestions.
21.	**Recommendation to revise statements based on current evidence** **A. Recommend stating consistently that long-term data were not published at the time of this report** Evidence related to long-term data of ruxolitinib are currently under review at an upcoming Dermatology conference and as such remains embargoed. We therefore recommend ICER make the following edits for consistency:	We have updated our statement regarding long-term data for ruxolitinib cream based upon additional publications since the draft report was published. We believe the revised language accurately reflects the available information about topical ruxolitinib.

	▯ On page 10, "*There is currently inadequate information on long-term safety of topical ruxolitinib"* **Suggested Revision:** Long-term safety data for topical ruxolitinib were unavailable at the time of this analysis. On page 33, and 34, last sentence: *"No long-term data was identified"* **Suggested Revision:** Long-term data were unavailable at the time of this report. On page 37, *"Side effects of ruxolitinib cream were similar to or better than placebo, though long-term safety remains uncertain."* **Suggested Revision:** Side effects of ruxolitinib cream were similar to or better than vehicle cream. Long-term safety outcomes were unavailable at the time of this report.	
22.	Safety concerns due to systemic absorption of ruxolitinib cream Recent publication by Gong X et al, have concluded that plasma ruxolitinib concentrations after treatment with topical ruxolitinib cream in patients in the 3 clinical trials are not expected to lead to systemic plasma concentrations associated with adverse effects commonly associated with oral JAK inhibitors.1 ▯ On page 10, the ICER draft evidence report states: *"As a topical JAK inhibitor therapy, safety concerns are likely not as great as with oral JAK inhibitors, but there still is systemic absorption of the topical agent."* **Suggested Revision:** "Pharmacokinetic study was conducted using data from the phase 3 and phase 2 trials of patients with ruxolitinib cream 0.15%, 0.5%, 1.5% once daily and 0.75% and 1.5% twice daily. Plasma ruxolitinib concentrations after treatment with topical ruxolitinib cream in patients with up to 20% BSA affected by AD are not expected to lead to systemic plasma concentrations that may be associated with adverse effects commonly associated with oral JAK inhibitors." On page 35, first sentence under Uncertainty and Controversies: *"Although ruxolitinib cream is a topical JAK inhibitor and concern for side effects may be lower, systemic absorption still occurs and..."* **Suggested Revision:** "Ruxolitinib cream, a JAK inhibitor, was specifically designed and formulated for topical application to minimize systemic absorption. Pharmacokinetic data for ruxolitinib cream suggest that adverse events associated	Thank you for providing this additional information. Nevertheless, we believe that our statements about safety concerns remain applicable. Given the FDA's recent action, we have revised this sentence to highlight some of the FDA's concerns (https://www.businesswire.com/news/home/20210611005030/en/Incyte-Announces-U.S.-FDA-Has-Extended-the-New-Drug-Application-Review-Period-for-Ruxolitinib-Cream-for-the-Treatment-of-Atopic-Dermatitis).

	with systemic absorption commonly associated with oral JAK inhibitors is not expected."(Gong X et al)	
23.	**Statements related to sub-group analyses** Incyte disagrees with ICER's subjective conclusions made when assessing evidence from subgroup analyses. Subgroup analyses were conducted post-hoc and not pre-specified or powered to show comparative evidence among them. We therefore recommend the following changes: ▯ Page 35, Disease Severity: *"Subgroup analyses based on disease severity at baseline suggest qualitative better outcomes in patients with moderate disease compared to those with mild disease (see Evidence Tables D3.63-66)."* **Suggested Revision:** "Proportion of patients achieving IGA-treatment success in the sub-groups of mild and moderate disease severity were consistent with the overall study (see Evidence Tables D3.63) ▯ Page 35, Last Sentence: *"The effectiveness of ruxolitinib in patients with darker skin complexions may be somewhat less, supporting the need for trials in broader populations."* **Suggested Revision:** "Ruxolitinib cream has demonstrated effectiveness in darker skin population, a population that is often under evaluated in clinical trial studies."	We believe that the current statements are accurate and provide useful information to readers in terms of identifying individuals with varying severity of atopic dermatitis who may benefit from ruxolitinib compared to other treatment options. We have added to the section on sub-groups to include data stratified by race that has been published as part of an abstract presentation. Our statement in the uncertainty/controversies section regarding treatment of patients with atopic dermatitis and darker skin complexions reflects this added information and is intended to alert readers to the need for more information for this patient group.
24.	**Comparative clinical assessment rating of ruxolitinib cream** Incyte respectfully disagrees with ICER's comparative net health benefit rating of C++ (comparable or better) based on the published evidence of ruxolitinib cream. We consider the evidence of ruxolitinib cream compared to topical emollients to be superior and recommend a rating of A based on the rationale below: ▯ **Comparator and Treatment History**: Patients in the 0.75% and 1.5% active arms in the Phase 3 clinical trials were compared to patients randomized to vehicle cream, which is a bland emollient. Other emollients such as Eucerin® cream were allowed during the double-blind period. Moreover, approximately 90% of all patients enrolled in the trials had a history of previous AD medication use, which included topical corticosteroids, calcineurin inhibitors or systemic therapy. Ruxolitinib cream demonstrated a high level of efficacy and was well tolerated in patients with AD regardless of previous use of topical or systemic therapy.2 **Strength of Evidence:** Compared to vehicle, ruxolitinib cream 0.75% and 1.5% met key primary (proportion of participant achieving IGA-TS) and secondary endpoints (proportion of participants achieving EASI75, >4-point	In reviewing available evidence for ruxolitinib cream compared to vehicle (placebo), despite the FDA's recent decision to postpone consideration for an additional three-month review period, we have decided not to lower the evidence ratings in the revised report to "promising but inconclusive" because we continue to feel that net harm is unlikely.

	improvement in Itch NRS, clinically meaningful improvement in PROMIS Short Form-Sleep Disturbance, and Sleep-related impairment) at week 8. Additionally, a clear separation for both active treatment groups from the vehicle cream treatment group was evident at the very first post-baseline assessment (week 2). Antipruritic effect of ruxolitinib cream 0.75% and 1.5% cream assessed using Itch NRS score was evident as early as 12 hours after the first application. Furthermore, ruxolitinib cream has shown significant improvements in other well accepted and important efficacy and patient reported outcomes measures such as SOCRAD, DLQI,/CDLQI, POEM and WPAI.	
25.		
LEO Pharma		
26.	<u>Base Case Model Time Horizon</u> Atopic dermatitis ("AD") is a lifetime condition for some patients. In our clinical trials, patients suffered from moderate-to-severe AD for a median duration of 27 years prior to entry into those trials. The current 5-year base case model time horizon does not adequately capture the nature of the disease, nor the long-term value and potential risks of novel treatments for patients with AD. As such, we strongly recommend that ICER consider a 70-year lifetime horizon for the base case as was done in the 2017 AD review, rather than as a scenario analysis.	Thank you for this suggestion. Given that our cost-effectiveness research questions are related to addressing the value of interventions within our scope (and not subsequent lines of treatment) and because the interventions have no known relationship to changes in disease progression or to mortality, the cost-effectiveness findings at the end of five years should be closely aligned with that of a lifetime time horizon but with the added benefit of retaining clinical face validity. In addition, as one of our scenario analyses in the revised report, we have examined using a lifetime horizon.
27.	<u>Investigational Tralokinumab's Q4W Dosing After 16 Weeks</u> We would also like to note that basing a 5-year model time horizon period solely on 16-week data does not consider the Q4W dosing option for investigational tralokinumab after 16 weeks. Q4W dosing was available to patients who achieved EASI 75 and/or clear or almost skin after 16 weeks of treatment in all three pivotal trials (ECZTRA 1, ECZTRA 2, ECZTRA 3). ICER has acknowledged within its report that dosing and utilization will impact model outcomes, and that inclusion of the option for tralokinumab every four weeks would lower treatment costs. We feel strongly that ICER should conduct a scenario analysis reflecting the Q4W dosing option. Additionally, voting question 13 cannot be adequately assessed if there is not an analysis of the Q4W dosing option included in the report.	We have included a scenario analysis to reflect Q4W dosing in the revised report.
28.	<u>Long-term safety data</u>	We thank you for bringing this new abstract to our attention. As noted,

	ICER noted the need for long-term safety and efficacy data in the evaluation as noted on page 9 of the report: "*Safety is an important consideration with biologic therapies and, as above there have been particular concerns about the safety of JAK inhibitors when used for other conditions. Additionally, though, tralokinumab is a novel inhibitor of IL-13 and we have limited long-term safety data*." It is critical to note that tralokinumab is a fully human monoclonal antibody with a different mechanism of action from the JAK inhibitors. In the report, ICER states the following about JAK inhibitors: "Though abrocitinib, baricitinib, tralokinumab, and upadacitinib appeared to have few serious harms reported from the trials of atopic dermatitis, oral JAK inhibitors approved for other indications, including baricitinib and upadacitinib, have label warnings about potentially causing serious infections, blood vessel disorders, cancer and death, and serious harms are more common at the higher doses studied. Whether certain oral JAK inhibitors or their use in patients with atopic dermatitis is associated with fewer long-term harms remains uncertain." Despite acknowledging that "no similar risks have been reported for tralokinumab," (pg. 32) the Draft Evidence Report subsequently classifies tralokinumab alongside the JAK inhibitors as having a "small (but nonzero) likelihood of a negative net health benefit" (pg.32). This equivalence of safety concerns is not merited by quantitative analysis and contradicts qualitative statements made elsewhere in the report. Regarding longer term data, LEO would like to make ICER aware of key late-breaking clinical data that addresses this need for data on the clinical effectiveness and safety of investigational tralokinumab presented at the 2021 American Academy of Dermatology (AAD) virtual annual meeting. ECZTEND is a 5-year, open label extension trial including subjects from 9 parent trials evaluating the safety and efficacy of tralokinumab. The interim analysis (n=1174 total) included data from 1-year (n=612) and 2-year (n=345) cohorts from 4 parent trials (ECZTRA 1-3 and 5). This ECZTEND interim analysis demonstrated that the long-term use of tralokinumab 300 mg Q2W was well tolerated and the overall safety profile was consistent with the parent trials, with no new safety signals observed.	ongoing experience with tralokinumab continues to accrue. We acknowledge the long-term safety of current biologics and we look forward to seeing updated data for new biologics as well.
29.	<u>Network Meta-Analysis Considerations</u> Trial design differences in AD clinical trials make it challenging to compare trials via typical indirect comparison methodologies. Key differences amongst trials may pose challenges when seeking to compare outcomes, particularly	Thank you for your comments. We have added additional information detailing the NMA methodology in the revised version of the report. For the reasons cited in the report, we believe that the model

	when these differences may impact active treatment and placebo differently. As such, indirect treatment comparison using network meta-analyses conducted from trials with differing methodologies should be interpreted with caution. Given the information shared by ICER, LEO has been unable to fully evaluate the methods used in the NMA.	parameters selected best reflect the trial data and fit.
Pfizer		
30.	**1. Elevation of abrocitinib evidence rating when compared to dupilumab and emollients** On page 32 of the DER, ICER reports an evidence rating of "insufficient" (I) when comparing abrocitinib to dupilumab and an evidence rating of "promising but inconclusive" (P/I) when comparing abrocitinib to topical therapies alone. In the comparison of abrocitinib to dupilumab, ICER notes the "I" rating as "any situation in which the level of certainty in the evidence is low," whereas ICER states the "P/I" rating for abrocitinib compared to topical therapies alone as "demonstrating a moderate certainty of a small or substantial net health benefit, with a small (but nonzero) likelihood of a negative net health benefit." We disagree with these evidence ratings and respectfully recommend that ICER elevate the evidence rating of abrocitinib compared to dupilumab to a "Incremental or Better/B+" rating, defined as "moderate certainty of a small or substantial net health benefit, with high certainty of at least a small net health benefit." The rationale for this proposed change is based on the following evidence available in the literature: 1. In the JADE (JAK1 Atopic Dermatitis Efficacy and Safety) COMPARE phase 3 clinical trial (NCT03720470), abrocitinib was directly compared to dupilumab at week 2 with respect to itch response (PP-NRS4). Statistical superiority of 200 mg abrocitinib and numerically higher response of 100 mg abrocitinib was demonstrated for this endpoint.[2] In addition, a post-hoc analysis presented at the 2021 American Academy of Allergy Asthma & Immunology congress, showed that treatment with abrocitinib 200 mg provided numerically greater and more rapid responses than dupilumab across stringent efficacy endpoints (EASI-90, IGA-0, DLQI-0/1, etc.).[3] Response rates relative to placebo in the abrocitinib 100 mg and dupilumab groups were similar.[3]	The net health benefit considers not just efficacy but also risks. This comment primarily focuses on the benefits but does not take into account the potential risks or harms. It is when considering both, we have arrived at the current recommendations. These are also reflected in existing black box warnings for oral JAK inhibitors as well as ongoing concerns expressed by the FDA as part of their consideration of oral JAK inhibitors including abrocitinib for this new indication.

	2. Furthermore, in a recently published network meta-analysis (NMA) of systemic therapies for moderate-to-severe AD which used fixed-effects and random-effects Bayesian NMA models, abrocitinib 200 mg once daily (QD) was shown to have higher rates of EASI response compared with dupilumab 300 mg every 2 weeks (Q2W) in both monotherapy and combination therapy networks.[4] Specifically, in the monotherapy network, abrocitinib 200 mg QD was estimated to have a >97.5% probability of superiority over dupilumab 300 mg Q2W with respect to EASI-50, EASI-75, and EASI-90. In the combination therapy network, abrocitinib 200 mg QD had the highest observed EASI-50, EASI-75, and EASI-90 response rates and was estimated to have a 96% probability of superiority over dupilumab 300 mg Q2W. We believe these probabilities, which were based on all clinical evidence available at the time of this NMA's systematic literature review, would surpass the threshold for "high certainty of at least a small net health benefit" of abrocitinib over dupilumab, consistent with a "B+" rating. 3. In addition to clinician- and patient-reported outcome measures collected in randomized clinical trials, patient preference is an important consideration of net health benefit not traditionally captured in NMAs or economic models. A recently published study sought to quantify patient preferences for systemic AD treatment attributes and differentiate between systemic treatments using a discrete choice experiment.[5] The results indicated that patients significantly preferred an oral daily administration over a biweekly injection and also preferred treatments with more rapid effect of itch relief. We believe both characteristics of abrocitinib should be considered as part of the net health benefit rating.	
31.	Similarly, we respectfully ask ICER to elevate the evidence rating of abrocitinib compared to topical therapies alone to a "B+" based on the following evidence available in the literature, whereby superiority to placebo was consistently shown: 1. Across the abrocitinib JADE monotherapy trials included in ICER's assessment (MONO-1[6], MONO-2[7], Phase 2b[8]), patients were permitted to use topical non-medicated emollients. Abrocitinib 200 mg and	See above.

	100 mg consistently and significantly improved signs and symptoms of moderate-to-severe AD compared with placebo. Namely, more patients treated with abrocitinib achieved primary and key secondary IGA, EASI-75, and itch score responses compared with patients treated with placebo. In addition, when considering both commonly-used and higher threshold efficacy endpoints, a post-hoc pooled analysis of the adult cohort of these 3 monotherapy trials found that higher proportions of patients treated with abrocitinib (200 mg, 100 mg) versus placebo achieved PP-NRS4 (47.1%, 34.7% vs 14.8%), EASI-75 (62.3%, 41.9% vs 12.2%), PP-NRS 0/1 (31.7%, 20.1% vs 4.8%), or EASI-90 to <EASI-100 (29.3%, 15.9% vs 5.9%) responses at week 12.[9] 2. Abrocitinib combination studies had similar patterns. In JADE COMPARE, all treatment groups were required to use emollients twice daily and therapy with a medicated topical (applied once daily) was started on day 1 of the treatment period. Both doses of abrocitinib demonstrated superiority compared to placebo when assessing IGA response at week 12 and 16 ($p < 0.001$), EASI 75 response at week 12 and 16 ($p < 0.001$), and itch response (PP-NRS) at week 2 ($p < 0.001$).[2] In the JADE TEEN trial in adolescents, abrocitinib QD (200 mg, 100 mg) was compared to placebo in combination with standardized medicated topical therapy and found that at week 12, more patients treated with abrocitinib (200 mg, 100 mg) versus placebo achieved IGA (46.2%, 41.6% vs 24.5%; $p<0.05$ for both), EASI-75 (72.0%, 68.5% vs 41.5%; $p<0.01$ for both), and PP-NRS4 (55.4%, 52.6% vs 29.8%; $p<0.01$ for 200 mg vs placebo) responses.[10] 3. Similarly, in the recently published NMA cited above, across both abrocitinib doses and monotherapy/combination studies, abrocitinib was estimated to have a 97.7%-100% probability of superiority over placebo/placebo + topical therapy with respect to IGA and PP-NRS response.[4] We believe these probabilities exceed the threshold for "moderate certainty of a small or substantial net health benefit, with high certainty of at least a small net health benefit" and the size of the efficacy differences between abrocitinib and placebo arms	

	represents a “substantial” net health benefit, consistent with a "B+" rating. 4. In the 2017 evaluation of dupilumab for moderate-to-severe AD, ICER rated the clinical evidence for dupilumab relative to treatment with emollients with or without continued failed topical treatments a “B+” rating, despite a similar number of key trials (SOLO 1, SOLO 2, LIBERTY AD CHRONOS, Thaci 2016, and Blauvelt 2016) and follow-up length (16 weeks) as the available abrocitinib data package.[11] Similarly, ICER identified 5 RCTs of abrocitinib varying in duration from 12 to 16 weeks of treatment comprising the evidence base of the current comparative clinical effectiveness assessment, all of which demonstrated superiority of abrocitinib compared to placebo on the primary endpoints. Finally, a recently presented integrated safety analysis included 2856 patients in the all-abrocitinib cohort (pooled from 6 studies including a long-term extension study); 1248 had ≥24 weeks and 606 had ≥48 weeks of abrocitinib exposure.[12] Results of this integrated safety analysis were consistent with results in individual trials. Based on this analysis, abrocitinib was well tolerated, with a safety profile appropriate for long-term treatment in this population.	
32.	**2. Inappropriate speculation on treatment population** ICER notes on page 40 of the DER when describing “Key Model Choices and Assumptions” that “The patient population is assumed to exclude patients over 50 with increased cardiovascular risk, as JAK inhibitors will likely not be approved in that population.” We disagree with including speculation such as this in ICER’s evidence report and recommend its removal as this is a decision ultimately made by the FDA.	At present we have chosen not to include a scenario where the drug increases mortality. If the manufacturer believes strongly that we should add a population at increased cardiovascular risk, we could consider including this population in the next version of the report.
33.	**3. Inclusion of the cost-consequence analysis as a scenario analysis rather than as a base case analysis** As part of its base case analyses, ICER includes a cost-consequence model estimating the cost per patient-reported outcome (PRO). ICER includes one measure for itch (PP-NRS) and three measures for sleep (POEM, SCORAD, ADerm-IS), wherein the data are derived from a subset of manufacturer submissions. ICER also notes that the analysis was conducted for a specific PRO, only if the data were provided for each EASI responder category. While we acknowledge the importance of measuring PROs in this specific patient population and have done so	Thank you for your comment regarding the cost-consequences analysis being included in the base-case. ICER reviews always attempt to include clinical outcomes that are relevant to patients in the base case. Throughout the patient engagement process, we heard that these PROs were important to patients, and we have therefore included them in the main report to the best of our ability given the data that were made available by manufacturers. This information serves as supplementary information to the cost-

	extensively in abrocitinib's JADE clinical program, we do not believe the cost-consequence analysis should be included as part of ICER's base case results for the following reasons: 1. In its description of the cost-consequence results, ICER notes that "the average incremental change in score over the five-year time horizon is presented where data was available by health state, as no commonly meaningful threshold or translation for these measurements was identified." Without a common threshold for interpreting these results, it will be difficult for payers and policymakers to interpret the output and use it to make meaningful decisions when it comes to patient access, especially when reported in the same context as the cost-effectiveness (CE) results (i.e., cost per quality-adjusted life year [QALY] gained, health benefit price benchmarks [HBPB]). 2. Moreover, in a report by the National Institute for Health Research, the NHS notes that while cost-consequence analyses can present a broader range of health and non-health costs and benefits, there are a number of disadvantages; specifically, the NHS notes that cost-consequence analyses: (1) do not provide specific guidance on cost-effectiveness thresholds, (2) have limited generalizability given disaggregated outcomes and lack of common thresholds across outcomes, and (3) lack transparency for decision-making purposes.[13] Given the above, we respectfully request that ICER move the cost-consequence analysis to the scenario analysis portion of the report and subsequently provide a meaningful interpretation of this analysis to aid patients, policymakers, and payers in understanding the outcomes and applicability to the AD treatment landscape.	utility analysis—for which the interpretation with regards to cost-effectiveness thresholds and generalizability is well established.
34.	**4. Discontinuation probability of emollients** On page 41 of the DER, ICER notes that a per-cycle discontinuation probability of 25.40% was assumed for emollients/standard of care (SOC) in the CE model; this discontinuation probability is sourced from the ECZTRA 1 and ECZTRA 2 phase 3 clinical trials of tralokinumab.[14] We have several criticisms of this input assumption in the CE model:	Thank you for this comment. For the discontinuation rates used in the current model, we looked for discontinuation data conditional on patients responding in the initial 16-week period. In the available trials for therapies included in this model, ECZTRA 1 and 2 were the only trials that provided these data for the placebo arm in the extension period without the use of topical corticosteroids. We considered using the same source for discontinuation

	1. This discontinuation probability is only representative of the placebo arm from the trials of tralokinumab. Because there are other interventions compared to emollients/SOC in ICER's analysis, it is inappropriate to base the discontinuation probability off of one intervention's placebo arm. We respectfully request that ICER provide justification for why only the tralokinumab phase 3 clinical trials were considered to inform the emollients/SOC discontinuation rate. 2. The discontinuation probability from ECZTRA 1 & 2 (25.40%) is considerably lower than the SOC discontinuation rate assumed in ICER's 2017 evaluation of dupilumab (65.80%), wherein ICER assumed that discontinuation in the SOC arm was equivalent to the placebo arm of the dupilumab clinical trial.[11] We recommend that ICER consider conducting a sensitivity analysis for the discontinuation probability assumed for the emollients/SOC arm of the CE model given the significant differential between these two rates.	as in the 2017 model, however the 65.8% discontinuation value was not conditional on having a response in the first 16-week cycle.
35.	**5. Data inconsistencies** Pfizer has identified several inaccuracies and opportunities for clarification, listed in Appendix A with their exact location for ease of correction. We recommend these be addressed in the subsequent version of the Evidence Report.	Thank you. We have reviewed your feedback and addressed any inaccuracies in the revised report.
36.	**6. Comments on Draft Voting Questions** As part of this review period, ICER also provided Draft Voting Questions in anticipation of the Policy Roundtable portion of the public meeting scheduled for July 23, 2021. After reviewing the questions, we have the following feedback: • **Question 9:** States "Patients' ability to achieve major life goals related to education, work, or family life"; however, AD has a substantial impact on activities of daily living and other aspects of patients' and caregivers' lives beyond "major life goals". We recommend adding outcomes to the list to reflect "Patients' [caregivers'] ability to achieve day-to-day goals and activities." • **Question 11:** We respectfully ask ICER to provide additional context and clarification around the intended interpretation of "health inequities." Participants in the Policy Roundtable have a wide variety of backgrounds and experiences and we are	We appreciate your feedback, and we will work to clarify the wording of our voting questions with the CEPAC members prior to the meeting. For questions 15 and 16: per our Value Assessment Framework, we will not take votes on "long-term value for money" in certain circumstances when there is no known net price. Although baricitinib and upadacitinib have not yet been approved for atopic dermatitis, these agents have been approved (and have a reported net price) for other indications. We will revise the wording of these questions to make this clearer to the CEPAC members.

	concerned the question may be too vague for interpretation. • **Question 12:** We believe the question "What are the relative effects of the JAK inhibitors as a class versus dupilumab on patients' ***ability to manage and sustain treatment given the complexities of the regimens***?", in particular the bolded language, is vague and leading in nature and therefore should be clarified and rephrased. • **Questions 15 & 16:** We request that it be noted why only 2 of the 4 systemic therapies are included in this section of the Voting Questions (e.g., we assume it is because their prices are not publicly available at this time).	
Sanofi/Regeneron Pharmaceuticals		
37.	**As it remains unknown which doses of the JAK inhibitors will be approved by the FDA, Sanofi/Regeneron recommend that ICER acknowledge this uncertainty and include a caveat when presenting the results of the updated NMA.** For instance, a draft report of the upadacitinib HEADS-UP study, in which 30 mg was the only dose evaluated, was added to the report's NMA. It is currently unknown if this dose will be approved by the FDA. Should the final approved dose not be 30 mg, this could impact the NMA findings. We recommend that ICER acknowledge the possibility that the NMA results will not be valid if a dose is not approved by the FDA.	We believe that the information presented in the NMA findings reflect the pivotal phase III trials for all of the treatments studied and their comparators. It is clear within the report (in the tables and footnotes, when applicable) where different doses are included and the results of those different doses.
38.	**Sanofi/Regeneron believe that the multinomial model is not appropriate for the NMA**. A multinomial model may be used to address possible abnormal estimates across Eczema Area and Severity Index (EASI) response thresholds. Abnormal estimates may be due to the independent modelling of the categories and/or due to high missing data on any given EASI responses. In this particular NMA, there are no such issues, therefore we do not see the justification for the multinomial model. Further, the disadvantage of the multinomial model is the strong assumption that the treatment effect of achieving each EASI response threshold is the same across all EASI cut-offs, that is, the model assumes that the relative increase in an EASI-75 response would be exactly the same for an EASI-50 or EASI-90 responses. This is an influential assumption that is not supported by the evidence from the individual studies. The attempt to increase precision using a multinomial model in this case is inappropriate and the point estimates could be biased. **Sanofi /Regeneron recommend that ICER models the EASI responses separately, as was done in the first ICER NMA draft report.**	We believe that the multinomial model more appropriately reflects the EASI scores. We have added additional details and rationale on the NMA methodology in the revised report.

39.	Sanofi/Regeneron agree with ICER that **safety is of utmost importance when assessing the value of treatments for AD**. **Dupilumab's long-term safety has been well established, in both children as young as six years of age and adults.** This is supported by a robust and ever-growing body of real-world evidence, as well as widespread use in clinical practice.	We agree that the draft report reflects this statement.
40.	Sanofi/Regeneron agree that long-term safety is critical and needs be supported by long-term evidence. Therefore, we do not agree with **speculative statements included in the report referring to the safety of treatments evaluated. For example, on page 29, ICER states** *"though dupilumab is an IL-4 receptor alpha antagonist, it inhibits IL-4 and IL-13 signaling and suggests that long-term safety data may also apply to tralokinumab"*. This statement is not supported by evidence. We recommend deleting this sentence from the report.	We have revised this sentence in the revised report to clarify our intended meaning.
41.	Given ICER's recognition of the importance of long-term safety, **Sanofi/Regeneron disagree with the exclusion of adverse events in the cost-effectiveness evaluation**. Ignoring adverse events as a factor in the cost-effectiveness analyses may underestimate the cost and overestimate the benefit of treatments associated with important safety concerns. **Sanofi/Regeneron recommend that ICER takes into account important** adverse events observed with JAK inhibitors as described in the boxed warnings of their US prescribing information: serious infections, malignancy, and thrombosis.	We acknowledge serious safety concerns about the JAKs throughout the report, however, the frequency in which these are reported does not align with our inclusion criteria for the cost-effectiveness model. This is also discussed in more detail in the contextual considerations section of the report.
42.	Sanofi/Regeneron believe that, in addition to long-term safety, **the long-term efficacy and durability of effect of treatments for AD should be demonstrated in clinical practice**. As the standard of care in AD, dupilumab's long-term efficacy is well established and further supported by real-world evidence.	Thank you for this statement.
43.	Sanofi/Regeneron agree with **ICER's acknowledgement of the importance of type 2 co-existing diseases in AD and the recognition that dupilumab "*has proven efficacy in treating certain patients with asthma or chronic rhinosinusitis*".**	Thank you for this statement.

#	Comment	Response/Integration
Patient/Patient Groups		
National Eczema Association		
1.	Cost-consequence analysis for depression/anxiety: During our call with the modeling team and in our April 2021 comment letter we articulated the importance of anxiety/depression outcomes from the patient perspective. In the current health state model structure, patients either remain in a non-responder state or transition to one of three responder states. We commend the inclusion of itch and sleep into the cost consequence analysis, as these are outcomes of importance to patients. However, given the significant mental health burden of AD, which often correlates with uncontrolled disease[4-6], existing literature could have been used to additionally estimate the potential benefits of reduced anxiety/depression across the therapies in the responder states.	We agree that depression and anxiety are important outcomes as reflected in the evidence section of the report. However, we feel these data were inadequate to extrapolate and use in our report as you have described for the economic analyses. We were able to include HADS (hospital anxiety and depression scale) in our cost-consequence analysis for LEO Pharma's tralokinumab, which was the only therapy to provide anxiety/depression data by health state.
2.	Pediatric/adolescent scenario analysis and separate voting questions: Based on our call with the modeling team we anticipated ICER would consider adding a pediatric-focused scenario analysis, as completed and ongoing clinical trials for abrocitinib, upadacitinib and tralokinumab have included ages 12 and up, as well as the potential for off-label usage. While omitted for the current report, we suggest this remains an opportunity for the final report or for more specific discussion prior to the final vote. Clinical benefits in the pediatric population may provide greater value when considering the potential spillover benefits to their adult caregivers.	Thank you for this recommendation. We present evidence for adolescent and pediatric patients in the report, but do not perform economic analyses. In terms of the voting questions, we have included adolescent patients along with adult patients. However, there is insufficient evidence for pediatric patients for the new therapies studied, and for this reason they are not included in our voting questions.
3.	Consideration of out-of-pocket costs: While the "average" eczema patient experiences substantial financial difficulties due to the well documented economic burden of this disease[10], patients of lower socioeconomic status are particularly vulnerable.[11-14] Without explicit consideration for health plan policies that may place certain AD therapies on tiers with higher out-of-pocket cost-sharing, lower socioeconomic status patients could be impacted more than those with more expendable income. Knowing this information, different scenarios could be modeled to account for costs and benefits differences impacted by changes in out-of-pocket expectations.	We agree that the costs of these new drugs may have serious consequences for individual patients and we include a discussion of this in our contextual considerations section. Estimating patient out of pocket costs is challenging given the variability in insurance designs in the U.S. and the complex relationship between deductibles, co-pays, and co-insurance therein.

4.	Highlight the revisions made through each validation step: In the draft report, model validation was described on page 51 that includes steps the research team took to refine the model and data used. It would add clarity for the audience to highlight which revisions were made based on patient group or other stakeholder feedback. This would acknowledge the input and engagement of external stakeholders as well as improve the transparency of the validation steps and general value assessment process.	We appreciate the extensive input for our draft report. However, the large amount of input we receive from many sources, some of which is overlapping, makes it impractical for us to track all changes. We produce this document to allow stakeholders to assess how their comments are addressed and any changes made in the revised report.
5.	Add a column for modified societal perspective costs from table E4.2 to Table 4.9: Rather than separating the modified societal costs, we recommend following best practices and including societal costs results alongside the base case costs in the main results table.	ICER has a specific base case and guidelines related to when the modified societal perspective is to be included as a co-base case. We detailed these guidelines, a priori, in the model analysis plan and they can be found in our Value Assessment Framework.
6.	Section 4: "Long-Term Cost Effectiveness" – Inappropriate for a 5-year base case analysis: The incremental cost-effectiveness analysis for the base case focuses on a 5-year time horizon. Using the phrase "Long-Term" in the title of this section implies a lifetime analysis typically chosen by economists looking to capture the more complete picture. Reviewers should be reminded this is a truncated analysis. ICER should remain consistent in its use of "short-term" and "long-term" as it does on page 8 of the ICER Value Framework methods document describing the rationale of a "short-term" 5-year time horizon typically chosen in its budget impact analysis.	We have added a sentence within the CEA section to give further clarity on the definition of "long-term," meaning that evidence from clinical trials was extrapolated beyond the duration of trial follow-up to a time horizon of five years. We acknowledge that a lifetime time horizon is longer than the base-case horizon of five years for this evaluation. Given that our cost-effectiveness research questions are related to addressing the value of interventions within our scope (and not subsequent lines of treatment) and because the interventions have no known relationship to changes in disease progression or to mortality, the cost-effectiveness findings at the end of five years should be closely aligned with that of a lifetime time horizon but with the added benefit of retaining clinical face validity.
7.	We would like to recognize the hard work of the ICER team in synthesizing the evidence and estimating the value of JAK inhibitors and monoclonal antibodies for the treatment of AD. We understand the limits created by the value framework with a pre-specified focus on the value to the health system (or health system perspective) rather than the patient or society. The report acknowledges the	ICER's goal is to improve affordability for all patients through aligning price with value. We agree that addressing the full range of clinical and economic burdens, access to treatments and health inequities is very important. However, in some cases data

	significant burden AD places on "all aspects of patients' lives and those of their family and caregivers" in the first paragraph of the Executive Summary. We appreciate that ICER recognizes these burdens. While the current draft report falls short in addressing or incorporating the full range of clinical and economic burdens of AD in the value assessment, we hope to continue working with your team to provide additional context for the final report and for future evaluations that impact the AD patient community. Specifically, through this process we recognized that out-of-pocket costs, patient affordability, and access are not currently incorporated into the value assessment in a meaningful way. Through our patient engagement activities, we have identified significant health disparities in AD that may have an impact on different components of care. While ICER may not be responsible for the ultimate access decision or formulary determination, recognizing the potential consequences (intended and unintended) on formulary design and access may be an area of opportunity for future evaluations. Lower socioeconomic patients are often the most significantly impacted by more restrictive managed care mechanisms, so special considerations may need to be made to address these populations.	about these aspects are not available to the extent that we can factor these inputs into our model. In these cases, we include these aspects into our "Potential Other Benefits and Contextual Considerations" section, and provide the Independent Appraisal Committee the opportunity to vote on these aspects during the public meeting. We appreciate your collaboration throughout the review process, and we hope to keep working with patient advocacy organizations on assessing the best ways to incorporate these important aspects into our reports.
8.	Contextual Considerations Questions Section (Q6): "Acuity of need" – We read acuity to imply a serious AD crisis leading to urgent/emergency care and/or hospitalization. Please clarify this terminology if ICER means to focus on whether the patient has severe AD, or other intended focus.	This question is intended to assess short-term risk of death for patients without treatment. Decision-makers may wish to give added priority to treatments for conditions that present a high short-term risk of death. This category captures what some ethicists have called a "rule of rescue" and the sense that even relatively small absolute gains in lifetime may be of higher priority when patients otherwise have very little time left before they are likely to die. You can read more about the rationale behind the potential other benefits and contextual considerations here.
9.	Contextual Considerations Questions Section (Q8): Please clarify what may be included in the "Other" category during voting.	This is part our voting questions template and we can add an "other category" if needed.
10.	"Long-term Value for Money" Section: Questions 15 & 16: It may be confusing using the phrase "long-term value for money" for baricitinib and upadacitinib on	We have added a sentence to the report within the CEA section to give further clarity on the definition of "long-term," meaning that evidence from clinical trials

	question 15 and 16 when the base case incremental cost-effectiveness analysis focuses on 5 years. The long-term analysis was only included as a scenario analysis and if the committee focuses on Table E4.4, they might see a very different answer.	was extrapolated beyond the duration of trial follow-up to a time horizon of five years. We acknowledge that a lifetime time horizon is longer than the base-case horizon of five years for this evaluation. Given that our cost-effectiveness research questions are related to addressing the value of interventions within our scope (and not subsequent lines of treatment) and because the interventions have no known relationship to changes in disease progression or to mortality, the cost-effectiveness findings at the end of five years should be closely aligned with that of a lifetime time horizon but with the added benefit of retaining clinical face validity.

#	Comment	Response/Integration
Other		
Consumer Action and PIRG		
1.	We want to frame the concerns for you regarding this anticompetitive practice from the consumer perspective so that the panel can understand that rebate walls have gained the attention of policy makers and antitrust enforcers; how rebate walls limit patient choice regardless of price; and why the panel should consider the competition and patient impacts of the rebate wall when it analyzes the cost effectiveness of each new drug for the atopic dermatitis market. Each patient experiences atopic dermatitis differently and there is no typical patient or treatment approach. Because of this, preserving treatment choice is critically important for patients. ICER is currently conducting an economic evaluation of several JAK inhibitors and monoclonal antibodies seeking FDA approval to treat moderate to severe atopic dermatitis. Soon, JAK inhibitors including abrocitinib (Pfizer); baricitinib (Olumiant®, Eli Lilly); upadacitinib (Rinvoq®, AbbVie); and ruxolitinib (Incyte Corporation) and a monoclonal antibody: tralokinumab (LEO Pharma) will be approved by the FDA and will be marketed to patients. The key to the successful adoption of any of these newly launched prescription drugs by patients and healthcare providers is insurance coverage. If prescription drugs, particularly those as expensive as biologic treatments, are not widely reimbursed by insurance, patients will not have access to more affordable treatments. Rebate walls are likely to play a role in the adoption of some of these newly launched drugs. Specifically, our concern is that AbbVie uses a rebate wall to protect Humira and recently has been using rebate walls to help place Rinvoq on drug formularies to the detriment of rival drugs.	Thank you, we hope to further address this during the policy roundtable at the public meeting on July 23.
2.	**Because AbbVie will be entering the atopic dermatitis space with its new JAK inhibitor**, we would like to highlight how AbbVie has used rebate walls in the autoimmune space. On May 15, 2021, congressional leaders sent a letter to the FTC requesting that the FTC investigate how AbbVie's use of rebate walls may have maintained Humira's market power by excluding rival drugs from preferred positions on drug formularies.2 The letter further noted that "market experts have also raised concerns about AbbVie leveraging its market power to bundle rebates across indications to deny preferred positions on drug formularies to biosimilar and brand name rivals to Humira."3 These rebate walls are exclusionary contracting practices that AbbVie uses to limit	Thank you, we hope to further address this during the policy roundtable at the public meeting on July 23.

	the ability of rivals from gaining preferred formulary access or block them from getting on formulary at all.4 AbbVie provides conditional lucrative financial incentives to payors in the form of an "all or nothing" conditional sales volume-based rebate across Humira's ten indications in exchange for preferential formulary access and denying or limiting formulary access to a rival drug (i.e., step therapy). Rival drugs with only one indication and little to no patient volume cannot match the breadth of Humira's rebate so the payors are economically coerced to accept AbbVie's offer.	
3.	**AbbVie's rebate walls foreclose competition5 and harm patients by increasing costs and restricting patient access to more effective and affordable prescription drugs.6** Dr. Wayne Winegarden, director of Pacific Research Institute's (PRI) Center for Medical Economics and Innovation, claims that rebate walls cause patients to suffer in the form of artificially inflated prices which result in higher coinsurance payments, or out of pocket expenses that are usually a percentage of the list price, as well as reduced choice. 7 Dr. Winegarden calculates that ending rebate walls would save patients more than $6,000 of out-of-pocket savings for expensive biologics like Humira that run approximately $70,000 per year.8 Importantly, rebate walls cause patients to miss out on obtaining more effective treatments sooner by having to step through older incumbent drugs prior to using new more effective treatments. This raises the costs for patients and health plans because patients need to try older drugs and fail before gaining access to more effective and affordable treatments from the beginning.	Thank you, we hope to further address this during the policy roundtable at the public meeting on July 23.
4.	**AbbVie's rebate walls protect Humira and helped with the launch of two of its new immunology drugs.** AbbVie's rebate wall involves the coupling of volume-based rebates across Humira's ten indications with penalty provisions, resulting in the withholding of hundreds of millions of dollars from payors that put rival drugs on their formularies.9 On September 12, 2019, twelve consumer and public interest groups and four unions signed onto a letter to the FTC expressing concerns about the anticompetitive effects of the AbbVie-Allergan merger and identified AbbVie's use of rebate walls as a competitive concern.10 AbbVie's rebate wall kept Humira in the preferred position on formularies while impeding the ability of new drugs indicated for moderate to severe psoriasis from obtaining the preferred position on formularies even though many of the new drugs are clinically superior and lower cost than Humira.11 On May 5, 2020, FTC	Thank you, we hope to further address this during the policy roundtable at the public meeting on July 23.

	Commissioner Rohit Chopra raised concerns in his dissent of the FTC's approval of AbbVie's acquisition of Allergan that the FTC had evidence suggesting that AbbVie used its bargaining leverage and rebates on Humira to help with the launch of its new branded drugs.12 Indeed, AbbVie used rebate walls based off of Humira's prescription volume to compel payors to put its new psoriasis drug, Skyrizi, in a preferred position on payors' drug formularies.13 These arrangements also prevented these more efficacious drugs from being placed on a preferential position on the formulary forcing patients to go through costly step therapy before having access to the most effective drug for their particular diagnosis. Moreover, AbbVie's rebate wall has essentially been used to preserve formulary spots for both of its new drugs, Skyrizi and Rinvoq.15	
5.	**Federal Trade Commission Is Concerned About Rebate Walls** On May 28, 2021, the Federal Trade Commission (FTC) issued a report on rebate walls to Congress and committed to investigating exclusionary practices that "threaten to delay new entry" and "deny patients access to competing treatments."16 In its report, the FTC outlines the framework for a legal analysis of drug company rebate practices and noted that "a variety of stakeholders have identified rebate wall issues" and that "the Commission is closely attuned to pharmaceutical manufacturer contracting practices, including rebate strategies."17 Both FTC Chairwoman Rebecca Slaughter and Commissioner Rohit Chopra issued their own statements noting that the FTC needs to give more attention to rebate walls, but that the normal FTC investigatory process would likely take too long to avoid competitive harm in the near term.18 And, reportedly, the FTC has been investigating Johnson & Johnson's ("J&J") use of anticompetitive rebate walls to protect its blockbuster drug, Remicade, a drug used to treat rheumatoid arthritis, and to stifle the entry of Pfizer's biosimilar, Inflectra.19 So, the use of rebate walls is not limited to AbbVie. Others may be using the practice to stifle the entry of not just branded drugs, but for biosimilars and generics as well.	Thank you, we hope to further address this during the policy roundtable at the public meeting on July 23.
6.	Policy Makers Are Concerned About Rebate Walls On September 17, 2019, nine Senators, including *then* Senator Kamala Harris, wrote a letter to the FTC regarding the AbbVie/Allergan merger and they recognized that rebate walls harm competition and reduce consumer choice.21 The letter noted that "rebate traps or rebate walls can have the effect of preventing alternative drugs,	Thank you, we hope to further address this during the policy roundtable at the public meeting on July 23.

	including more affordable biosimilars and generics, from competing."22 On June 10, 2020, Senators Klobuchar and Blumenthal as well as Congressmen Cicilline and Jeffries asked the GAO "to conduct an assessment of the prevalence of rebate traps in pharmaceutical markets and their effects on pharmaceutical pricing, competition, and innovation."23 They noted that rebate walls can "be used in harmful ways to strategically exclude competing products. So-called "rebate traps" (or "rebate walls") may stifle pharmaceutical competition and product development, potentially limiting patients' access to lower-cost generic drugs and biosimilars, as well as new innovative drugs."24 On July 17, 2020, the U.S. House Committee on Appropriations included language in its report accompanying H.R. 7668 urging "the FTC to prioritize investigations into manufacturers that erect rebate walls to block competition from new branded therapies, biosimilars, generics, and other innovative products."25	
7.	**Given AbbVie's history and the interest from policy makers and antitrust enforcers in its use of rebate walls**, we are concerned that AbbVie could use rebate walls to advantage Rinvoq and disadvantage its rivals for the treatment of moderate to severe dermatitis. Accordingly, we hope that ICER's comparative value assessment of atopic dermatitis considers the market realities that rebate walls exist in the autoimmune space and how AbbVie's rebate wall could create barriers to more cost-effective therapies by foreclosing their access to drug formularies. The problem is that the most cost-effective products are unlikely to be available to patients if they cannot get on a drug formulary because of a rebate wall.	Thank you, we hope to further address this during the policy roundtable at the public meeting on July 23.
Partnership to Improve Patient Care (PIPC)		
8.	**ICER's model is not sensitive to or reflective of the outcomes that matter most to patients.** In ICER's *Patient and Caregivers Perspective* section of the draft evidence report, it is clear that the primary symptom of concern for AD patients is itch. Patients express that itch can lead to a host of additional problems including skin pain and infections, as well as disrupting sleep and causing anxiety and depression. It is primarily through itch and pain, that AD can have a profound impact on life activities, interpersonal relationships, and the ability to be productive at work. Patients highlighted the need for therapies to address itch and pain that work quickly, provide sustained relief, and are safe for long-term use.	Thank you for these suggestions. We based our modeling approach on available data, recommendations from clinical experts and patient advocacy organizations. Considering all of the input received from stakeholders mentioned above, we opted to use the EASI score for the model and focus on additional aspects in the Potential Other Benefits and Contextual Considerations section.

	Other than discontinuation rate, none of these aspects of importance raised by patients was incorporated into the model. The cycle in the model was 16-weeks, so any benefit from a therapy that resulted from a quick response as compared to a slower or delayed response would be missed in the ICER model. Similarly, long-term data was not used in the construction or execution of the ICER model. We would encourage ICER to rework the model to ensure the benefit of expedient relief is captured. Despite the emphasis patients put on the importance of itching on their quality of life, the ICER model is structured solely around Eczema Area and Severity Index (EASI) score, which combines coverage, location and severity weighted equally by clinicians – not patients. Recent studies have suggested that itch-specific measures have weak-to-moderate correlations with EASI. There are more sensitive resources available that do capture a more accurate picture of the patient's experience with itch and pain, and we would encourage ICER to look to these for its model. For example, the model could be built on a combination of EASI and PP-NRS or used patient itch questionnaire - numerical rating scale and verbal rating scale (PIQ NRS, VRS) or frequency of itch. For example, ICER states that more patients achieved a ≥4-point improvement in PP-NRS with upadacitinib 30 mg than dupilumab (55% vs. 36%). But since the ICER model is based solely on response as defined by change in EASI score, upadacitinib is considered to be 'less effective' than dupilumab. Subsequently upadacitinib has almost twice the efficacy of the comparator in terms of the one outcome that matters most to patients but still the model shows these two treatments to at best be equal in efficacy, and at worse, less effective than the comparator. We would highly encourage ICER to rework its modeling to ensure it is capturing the outcomes that matter most to patients.	
9.	**ICER's model continues to use the discriminatory Quality-Adjusted Life Year (QALY) and relies on population averages and does not take into account patient heterogeneity.** We would like to reiterate that the QALY innately discriminates against people with disabilities and chronic illnesses and is an inappropriate tool for assessing value. We would encourage ICER to look to more sensitive mechanisms that do not rely on population level averages	We appreciate the concerns about relying solely on QALYs. The quality-adjusted life year (QALY) is the gold standard for measuring how well all different kinds of medical treatments lengthen and/or improve patients' lives, and therefore the metric has served as a fundamental component of cost-effectiveness analyses in the US and

	and do a better job incorporating the outcomes that matter to the specific patient population in question. In addition to its reliance on the QALY, ICER compares all treatments it is assessing to placebo or dupilumab, under the assumption that both the index and comparator drugs are similarly effective for each patient. This is an example of when the value assessments only looking at the "average" patient will not reveal accurate or useful information on actual efficacy of treatments. For many patients dupilumab will not work, will stop working after treatment initiation, or will be discontinued due to side effects. For all three of these groups, the comparison to dupilumab is irrelevant. We would encourage ICER to abandon its reliance on population level averages and address the question of value from the perspective of patients who have very particular needs from their treatments.	around the world for more than 30 years. If evidence shows that a treatment helps lengthen life or improve quality of life, these benefits are comprehensively summed up to calculate how many additional QALYs the treatment provides, and this added health benefit is then compared to the added health benefit of other treatments for the same patient population. To complement the use of the QALY, ICER's reports also include a calculation of the Equal Value of Life Years Gained (evLYG), which evenly measures any gains in length of life, regardless of the treatment's ability to improve patients' quality of life. In other words, if a treatment adds a year of life to a vulnerable patient population – whether treating individuals with cancer, multiple sclerosis, diabetes, epilepsy, or a severe lifelong disability – that treatment will receive the same evLYG as a different treatment that adds a year of life for healthier members of the community. By understanding a treatment's cost per evLYG, as well as its traditional cost per QALY, policymakers can take a broader view of cost-effectiveness and be reassured that they are considering information that poses no risk of discrimination against any patient group.
10.	**ICER's inputs are opaque, and we would encourage more transparency.** The cost-effectiveness calculations in ICER's model are largely driven by the choice and application of the health utility weights within the QALY. In the past ICER has been urged by various stakeholders to be more transparent. Unfortunately, this specific report seems to take a step backwards and is less transparent than many previous reports, as many of its inputs are blacked out. It is very difficult for stakeholders to make comments on data choices we cannot clearly see. We would encourage ICER to be transparent about its choice of utilities and make a	The redacted data in the report and supplement are academic-in-confidence data provided to us by manufacturers. Per our guidelines for accepting and using "in-confidence" data, "Academic-in-confidence data will be redacted from all external and public ICER documents until the earlier of: (a) publication or presentation of such data by the data owner or study investigators; (b) 18 months following the date of the public ICER meeting; (c) for reports that are not subject to a public meeting, 18 months following report publication. Following

	concerted effort to share more, not less, data with stakeholders as it continues performing assessments.	any of these dates, ICER will unmask all redacted information from reports, presentations, and other public documents."
11.	**ICER uses randomized clinical trial data when real world estimates of utilities for health states, particularly for active disease, are likely to be more representative of the population of need.** As a general rule, real-world cohort-based estimates of utilities, especially for active disease states (non-response) will provide more accurate data than relying on randomized clinical trial data. Clinical trials are known to recruit healthier patients than those people who make up the real-world population of need. There is also the problem of the placebo effect in randomized clinical trials on patients in the comparator arm. Finally, patients in RCTs tend to receive far more non-treatment specific care and attention; symptom management, and interaction with clinicians than the average patient in a real-world setting. As such, quality of life measures in patients non-response states are often higher for patient in randomized clinical trials than in real world cohort studies. Given the availability of real-world estimates of utilities, we would encourage ICER to use this available data instead of relying on utilities from randomized clinical trials. Literature based values for utilities have been preferred in the vast majority of AD models produced in the last decade. A recent review of studies measuring health utility weights in AD patients showed a fairly consistent conclusion that untreated moderate to severe AD had a fairly consistent estimate of 0.61.	This report uses peer-reviewed (and academic-in-confidence) data that are currently available and highlights the limitations of these data as well as the qualitative input of a range of stakeholders. Per our review process guidelines, we can update our report findings after the public meeting once new relevant data (including real-world evidence) emerge.
Patients Rising Now		
12.	<u>People-Centered Perspectives</u> Atopic dermatitis – commonly known as eczema – is a complex immune disorder affecting the skin. The draft report does a reasonably good job of describing many of the clinical and personal challenges faced by people with atopic dermatitis. But it is also clear that better treatments for atopic dermatitis are needed because of great variability in how the condition affects individuals and people who have various co-morbidities. As the draft report states: • "Despite available treatments, many individuals do not respond to multiple different topical and systemic therapies supporting the need for new treatment options."[i]	Thank you for these suggestions. We agree that atopic dermatitis impacts various aspects of a person's life. When assessing the clinical effectiveness and cost-effectiveness of new treatments, our reports always take into consideration available data, recommendations from clinical experts and patient experts. Based on all of the above, we select a narrower scope for our reviews since it would be difficult to factor in all relevant drugs for all ICER reviews. However, when it is relevant to do so, we do also conduct class reviews.

	• "There was broad recognition that current therapies do not address all of the needs of patients with atopic dermatitis."[ii] Better treatments are needed not just to improve clinical outcomes, but perhaps more important, to improve patients' productivity and quality of life. As described in the draft report: "For students it can affect school attendance and lead to distraction when in class, negatively impacting developmental milestones. Similarly, atopic dermatitis can affect work through missed days, decreased work performance (presenteeism), missed promotions, limited career options, and even disability from one's chosen profession. The net result is a financial impact on individuals and families over the course of one's life in terms of educational and work advancement opportunities delayed or lost."[iii] Unfortunately, that reality is minimally recognized in the draft report's analyses and conclusions. There are similar important aspects of how atopic dermatitis affects people and their treatment choices that the draft report fails to acknowledge or incorporate into its analysis and conclusions. First, a key data point cited in the draft report highlights the personal financial toll of atopic dermatitis: "The overall costs associated with atopic dermatitis are estimated to be $5.3 billion in the US, including over $1 billion in health care costs."[iv] This means that the personal (i.e., non-health care costs) are about 400% greater than the health care costs. This four-to-one ratio quantifies the serious limitations of the draft report, its analyses, and its conclusions since it focuses almost exclusively on the costs that are less than 20% of the actual impact of the disease. Second, although the draft report discusses how atopic dermatitis significantly impairs an individual's work and life activities, it fails to capture the full consequences of the "social embarrassment and isolation"[v] resulting from a person's skin appearance, and how that leads to "psychological distress including loss of self-esteem, anxiety, depression, and suicidal ideation."[vi] Specifically, the draft report does not explore research about atopic dermatitis leading to greater suicide attempts (although it is unclear if the condition causes an increase in deaths from suicide) or other mental, emotional, or behavioral health issues.[vii]	Furthermore, we hope to bring up many of these points at the public meeting for this review. Economic modeling cannot always capture all of the details you have listed due to lack of data and high-quality peer-reviewed evidence. However, all of these aspects are important and we will highlight these in the discussions at the time of the CEPAC meeting on July 23rd.

	Third, while the draft report – like much of ICER's work – focuses on a small group of treatments, for people with atopic dermatitis and their clinicians, the actual range of treatment options is much wider and more complex. This discrepancy is apparent when comparing the draft report's scope with that of the two actual systemic reviews and technology assessments summarized and referenced in Section D5 of the Supplemental Material.[viii] One of those reviews evaluated "20 different medications," and the other "13 different approved treatments in Europe," in contrast with only six treatments included in the draft report. For clinicians, patients, policy makers, and others concerned with improving the quality and efficiency of health care within the populations of their purview (e.g., the management of Medicare, state Medicaid programs, private health insurance, Veterans Affairs' health care, Department of Defense health care, or the Indian Health Service), the question is not about evaluating small subsets of treatment options, but rather how to develop and implement appropriate policies for ensuring quality and efficient health care for the population for whom they are either paying for their health care or actually delivering their health care services and treatments. In contrast – as we've noted before – ICER's work is illusionary in that it assumes a unified, single health care system, and it assumes that there is a single health care budget for that "system." And lastly, in the subgroup analysis, the only differentiators are age and disease severity. However, there are some indications that women and Black Americans are more likely to have severe atopic dermatitis.[ix] Even though the available data may be limited or not definitive, given the inherent underrepresentation of women and people of color in clinical trials, and the disparities and inequities they continue to experience in access to health care in the U.S., we strongly believe that the draft report should at least address the important issues for those subgroups, namely potential issues related to the need for new treatments, and challenges accessing them. And in this area, we note that the draft report states, "Given the large impact of atopic dermatitis in African-Americans and the importance of skin appearance on outcomes of treatment more broadly, few trials included a sizable number of patients with darker skin complexions, and we are not aware of any trial that has reported outcomes among those with darker skin complexion."[x] So while ICER appears to be aware of this issue, we suggest that it be more explicitly stated in the draft report, and that the need for better and more	

	extensive data collection on those subgroups, and greater inclusion of people of color in future research, be stressed by ICER.	
13.	Data, Modeling, Assumptions, and Uncertainties Because the draft report does a deep numerical dive into the available research for six different medicines, it contains an extensive amount of data. However, just because there are numbers, and those numbers are compared and plugged into formulas for evaluative purposes, does not make the resulting "output" insightful, useful, or even correct. We are reminded of the old adage: "Not everything that counts can be counted, and not everything that can be counted counts."[xi] Breaking this down into its two parts, we see that the first part relates to the reality that patient concerns and perspectives are often hard to measure and are often not robustly evaluated in clinical research. For atopic dermatitis treatments, we are gratified that there are so many different patient-focused metrics as described in the draft report's Supplemental Materials Definition section.[xii] However, of those 11 different outcome measures, the draft report focuses on two that are investigator-measured (i.e., EASI and IGA), rather than patient-reported or primarily related to quality of life. This selection of measures may be because of the structure and compatibility of data across trials, but it underscores that the way data is collected and chosen for evaluation drives both thinking and conclusions. To that point, we appreciate that uncertainties about metrics such as EASI are discussed in the draft report, e.g., "…we assumed that levels of EASI response are associated with differences in health-related quality of life." However, there may be differential effects of the treatments modeled on conditions such as itch and sleep that are not completely captured by generic quality of life instruments. However, available data did not support the use of treatment-specific utilities. Additionally, there may be incremental effects of some of these treatments on quality of life in sub-populations of people with atopic dermatitis, such as those with co-occurring asthma or chronic rhinosinusitis, which are not explicitly captured in the current model."[xiii] Because of the importance of those uncertainties, they should have been explored in greater depth and earlier in the draft report, particularly since one researcher stated that the use of such measures "in clinical practice is not recommended," and that "both objective and subjective assessments of disease severity are important to assess, consideration of	We agree that this report contains a large amount of data and our evidence review requires synthesizing it to the best of our ability. Recognizing this, we have updated the presentation of data in the revised report with the intent to make it clearer and more accessible. In terms of the outcomes used in the economic models, we based our modeling approach on available data, recommendations from clinical experts and patient advocacy organizations. Considering all of the input received from stakeholders mentioned above, we opted to use the EASI score for the model and focus on additional aspects in the Potential Other Benefits and Contextual Considerations section. Our review of the literature and input from clinical experts led us to assume that atopic dermatitis does not affect mortality in the economic models. In the revised report, we have included a scenario where patients can receive the new therapy in addition to topical treatments.

clinical characteristics such as disease recurrence or persistence, as well as location of the affected areas, should be considered in the overall judgement of disease severity and consideration of therapy choice."[xiv]

And more generally concerning ICER's assessment approach, a recent review of books on the topic of evaluation metrics[xv] produced the following insights and quotes that are very illuminating:

- "Seduced by their seeming precision and objectivity, we can feel betrayed when the numbers fail to capture the unruliness of reality."
- "As Tim Harford writes, data 'may be a pretty decent proxy for something that really matters,' but there's a critical gap between even the best proxies and the real thing—between what we're able to measure and what we actually care about."
- "To simplify the world enough that it can be captured with numbers means throwing away a lot of detail. The inevitable omissions can bias the data against certain groups."
- "Numbers are a poor substitute for the richness and color of the real world."
- **"Numbers don't lie, except when they do."** [emphasis added]

Another problematic assumption in the draft report is the relationship between atopic dermatitis and mortality. The draft report states, "We assumed that atopic dermatitis disease and treatment did not affect mortality,"[xvi] and one of the Long-Term Cost Effectiveness analysis' assumptions is "Atopic dermatitis disease and treatments do not affect mortality."[xvii] However, research indicates higher rates of suicide attempts, and overall higher mortality, i.e., one analysis "found that patients with atopic eczema had an 8-14 percent increased risk of death due to infectious, digestive, and genitourinary causes. They noted that increased mortality risk was mainly in those with the most severe or more active atopic eczema. Patients with severe atopic eczema had 62 percent higher overall risk of death. These findings are consistent with previous studies."[xviii]

The draft report also primarily compared trial data that looked at monotherapy, but advancement and actual practice may include a combination of treatments, including systemic and topical. Once again ICER may be looking at the theoretical that does not reflect reality. As the report itself describes in discussing its modeling, "the NMA

	analyses that informed our effectiveness estimates in the model were derived from phase II and III RCTs that compared the treatments of interest to placebo with only the added use of topical emollients at 16 weeks. Therefore, the incremental value of these treatments may not be generalizable to patients using topical steroids and/or calcineurin inhibitors."[xix] Overall, the extensive data, charts, graphs, and comparative analytics across six different treatment options contained in the document made the draft report very user unfriendly. In other words, for unsophisticated readers, the content is probably indecipherable, leaving those individuals to look at the conclusions and assume that ICER's internal and external teams got everything correct. And for sophisticated readers and analysts – such as those who decide clinical care, formulary placement or reimbursement policies – there remains the question about how the information in the draft report fits in with the much larger array of treatment options for atopic dermatitis (including possible combinations of treatments), or the much larger issue of managing access and coverage for immunomodulator medicines. On both points, the draft report clearly fails usability tests in multiple and different ways.	
14.	Please explain how the New England CEPAC is both a "core program of ICER" and "an independent committee."	You are welcome to read more about our independent appraisal committees on our website.
15.	The draft report states that "ICER does not provide health benefit price benchmarks as part of draft reports because results may change with revision following receipt of public comments," however, that is not true. Health Benefit Price Benchmarks were included in ICER's recent draft report about Alzheimer's treatments. And further – as we pointed out in comments to that draft report – ICER's draft reports should absolutely include benefits price benchmarks from a societal perspective, particularly in this draft report because (as noted above), there is a 4:1 ratio in societal to health care costs. To add to the draft report's inconsistencies in this area, the Long-term Cost Effectiveness Supplemental Information goes into some detail about analyzing the situation from a societal perspective,[xx] but here too the draft report ignores the evidence about increased mortality related to atopic dermatitis. This is another example of ICER making up its own arbitrary rules but only following them when it sees fit to do so.	Because of the initially compressed timeline for ICER's review of aducanumab (before the FDA extended the review period), ICER consulted with the manufacturer at the beginning of the review process and both parties agreed that Health Benefit Price Benchmarks would be included in the aducanumab Draft Report. This was a specific alteration in the usual ICER review process to deal with the compressed timeline.

16.	The draft report states that as part of building the comparative clinical effectiveness model the assumption was made "that background topical medication is not an important effect modifier."[xxi] Does this mean that ICER believes that topical medications are ineffective? We would appreciate ICER specifically responding to this point and to the clinical logic behind that assumption as it relates to ICER's modeling in the draft report and hence the draft report's conclusions.	That is not the meaning of "effect modifier." https://www.amazon.com/Users-Guides-Medical-Literature-Evidence-Based/dp/0071794158 and https://www.amazon.com/Clinical-Epidemiology-Robert-Fletcher-MSc/dp/1451144474 are good texts discussing basic terms used in clinical epidemiology. Please let us know if that is not specific enough.
17.	There is no discussion about the biological mechanism of action of atopic dermatitis, aside from it being related to "problems with the body's immune system"[xxii] or as an "allergic condition,"[xxiii] while also noting that people with atopic dermatitis also commonly have allergies and asthma. Such general and imprecise language does a disservice to readers. According to Mt. Sinai Medical Center, atopic dermatitis is an autoimmune disease at the molecular level,[xxiv] and the Immune Deficiency Foundation also discusses atopic dermatitis within the spectrum of autoimmune skin diseases.[xxv] The draft report should include more discussion about the underlying cause of atopic dermatitis, and if the draft report's writers and reviewers disagree with the conclusions noted above, then those disagreements should be explained.	Thank you for this comment. We believe that we have written a comprehensive background section which includes references to relevant resources which describe the underlying causes of atopic dermatitis. In addition, the mechanism of action of the therapies studied are also presented and referenced.
18.	Given the extensive data density in the draft report, it is critical that the language be crisp, clear, and correct. However, there are several places in the draft report where words are missing, the meaning is unclear, or the text is complex and hard to decipher. Such poor writing (or faulty proofreading or copyediting) does a severe disservice to readers and ultimately to anyone who might use ICER's reports for anything substantive. For example: a. In this sentence, we believe the word "report" is missing: "Concerns about lack of long-term data for dupilumab, noted in ICER's 2017, have been alleviated over time based on published data and widespread use in clinical practice."[xxvi] b. And this sentence is misleading: "Non-pharmacologic treatments are recommended to maintain and prevent flares." That is, we do not believe that treatments are recommended to maintain flares.	Thank you. We have revised wording in certain parts of the report and supplement.

19.	In the draft report, the acronym AD is used to refer to Atopic Dermatitis, but it is not in the list of acronyms on page vii of the draft report nor could we find it specified in the text of the draft report. While that may seem obvious, in the previous draft report AD was used for Alzheimer's Disease, and that abbreviation was noted in on page viii of that draft report.	Thank you for pointing this out. We have revised this in the report and we now use "atopic dermatitis" throughout the report without the abbreviation.
Paul Langley		
20.	As you will no doubt recall, you are aware of my concerns that the ICER reference case framework for value assessment fails to meet the standards of normal science. That is, your reports lack credibility in the claims made for the value of products; they cannot be evaluated empirically nor can the claims be replicated. Your models also violate the fundamental axioms of measurement theory in confusing ordinal scales with interval and ratio scales, and simple logic in driving claims by assertions and assumptions.	Thank you for your feedback. ICER works with numerous clinical experts, modeling experts and academic institutions to provide a diverse and exhaustive approach to our value assessment work. We have developed the framework for our assessments with the help of several stakeholders in addition to the ones we have mentioned above, and we continue to welcome feedback on how to further improve our methodology.
21.	While you might view your standards and reports, and the application of lifetime incremental cost-per-QALY calculations and the application of cost-per-QALY thresholds as the state of the art in health technology assessment, the problem is that the entire exercise is essentially a waste of time. The QALY, for example, as you have been informed on a number of occasions, is a mathematically impossible construct with a paper in *F1000Research* and a letter to *Value in Health* pointing this out. As noted in the latter, we have now experienced 30 wasted years in health technology assessment, with ICER perpetuating this charade. The key point is that in the case of new and emerging therapies for atopic dermatitis we have too little data to make even a reasoned, and scientifically valid, claim for pricing and budget impact. This should be put on hold until more data become available instead of rushing in to invent modelled claims.	The quality-adjusted life year (QALY) is the gold standard for measuring how well all different kinds of medical treatments lengthen and/or improve patients' lives, and therefore the metric has served as a fundamental component of cost-effectiveness analyses in the US and around the world for more than 30 years. If evidence shows that a treatment helps lengthen life or improve quality of life, these benefits are comprehensively summed up to calculate how many additional QALYs the treatment provides, and this added health benefit is then compared to the added health benefit of other treatments for the same patient population. To complement the use of the QALY, ICER's reports also include a calculation of the Equal Value of Life Years Gained (evLYG), which evenly measures any gains in length of life, regardless of the treatment's ability to improve patients' quality of life. In other words, if a treatment adds a year of life to a

		vulnerable patient population – whether treating individuals with cancer, multiple sclerosis, diabetes, epilepsy, or a severe lifelong disability – that treatment will receive the same evLYG as a different treatment that adds a year of life for healthier members of the community. By understanding a treatment's cost per evLYG, as well as its traditional cost per QALY, policymakers can take a broader view of cost-effectiveness and be reassured that they are considering information that poses no risk of discrimination against any patient group.
22.	Let me consider the assertion regarding your belief that the EQ-5D preference instrument has ratio properties. For a measure to have ratio properties there must be no possibility whatsoever that the instrument can generate negative values. The true zero is a universal reference for any measure that claims to have ratio properties. We might believe it if you could prove, not assert, that there is no possibility of a respondent to the symptoms and response levels of the instrument reporting negative values.	Thank you for this comment.
23.	Clarification on your use of preference scores requires more information than has been provided in your draft evidence report. Unfortunately, we have no idea as to what these scores actually are for mild, moderate, and severe stages of AD. They are blanked out. All we have is the Delphic utterance from the internationally respected CHOICE expert group at the University of Washington, College of Pharmacy that in constructing your imaginary assumption driven claims for the pricing and recommendations for atopic dermatitis therapies were '*weighted by a single set of health state utility values from pooled manufacturer data to derive quality-adjusted life-years (QALYs)*.' Seeking further clarification on these utility scores the process is described by the University of Washington expert group as follows: > We derived health state utilities for the non-responder and responder states by pooling utility estimates from manufacturer submitted data. We estimated weighted average utility values for each health state, combining estimates from all treatments with data available by health state. We considered therapy-specific health state utility values to capture benefit beyond EASI	The redacted data in the report and supplement are academic-in-confidence data provided to us by manufacturers. Per our guidelines for accepting and using "in-confidence" data, "Academic-in-confidence data will be redacted from all external and public ICER documents until the earlier of: (a) publication or presentation of such data by the data owner or study investigators; (b) 18 months following the date of the public ICER meeting; (c) for reports that are not subject to a public meeting, 18 months following report publication. Following any of these dates, ICER will unmask all redacted information from reports, presentations, and other public documents." Regarding the specific questions you have about the model—in addition to the atopic dermatitis report and supplement, you can find further details and rationale

	score, however the available evidence did not support differential utility scores by treatment (p. 42). No further details are given. This is unfortunate because if the protocols for the various AD trials are reviewed (Clinicaltrials.gov: ECZTRA 1&2; MEASURE UP 1 & 2; AD UP; and SOLO 1&2) there is no evidence from the list of secondary outcomes for each of these of any health related quality of life or just quality of life instrument that is designed to generate either direct or indirect preference scores. At best, we have the ordinal Dermatology Life Quality Index (DLQI) in two trials (ECZTRA 1 & 2 and SOLO 1 &2) which simply provides an aggregate of 10 4-level Likert scales (scores 0 – 30). Other than that I have no idea how the University of Washington Expert Group then proceeded to create utility values for a ratio scale with a true zero and a range of 0 = death to 1 = perfect health. I presume, as these are all secondary endpoints for the various protocols that they were all powered to create a 'composite' utility scale. Can you confirm? It might also be pointed out that if these various inputs from manufacturers are patient reported outcomes with ordinal properties, then the calculations vaguely described by the University of Washington expert group are mathematically impossible (with a further concern that they lumped together utilities from different instruments). Ordinal scales can only support non-parametric assessments. I am not sure if the Washington expert groups understands the need to conform to the axioms of fundamental measurement in statistical and econometric analysis (let alone building imaginary simulation models); if so, this is a major concern that ICER and the University of Washington should address. As a renowned university research group I would have thought their training would have included measurement theory (and some elementary logic to include Hume's Problem). Given this, it might be pointed out that in your previous review and imaginary modelling for Dupilumab in moderate to severe AD you provide EQ-5D-3L utility values (source Sanofi data on file). For patients with moderate disease (IGA), the utilities ranged from 0.684 (baseline) to EASI 50 0.892, EASI 75 0.895 and EASI 90) 0.907 while for severe disease (IGA4) the baseline was 0.536 to EASI 75 0.535, EASI 75 0.090 and EASI 90 0.911.	in our Model Analysis Plan which is available here.
24.	What I find puzzling is that there are a range of preference scores for AD available from the literature; perhaps your	Thank you for this feedback. In addition to the atopic dermatitis report and

	expert group did not think a systematic review worthwhile? These are well documented and include the impact of demographic factors as well as comorbidities typically associated with AD as well as systematic reviews. Of particular note is the recent study by Silverberg et al utilizing the *AD in America Survey* sampled from the long standing GfK knowledge panel (n=8,217) Applying the SF-6D preference instrument yielded a mean AD score of 0.69; mild AD 0.73 and moderate to severe AD of 0.63 As this is an ordinal scales these mean values for the SF-6D are actually incorrect; they should have reported medians or modes. A study by Anderson et al utilizing the EQ-5D-5L and a visual analog scale (VAS), covering the US and selected European countries found for the US overall a EQ-5D-5L score of 0.77 for moderate AD and scores between 0.69 and 0.42 for severe AD. The VAS yielded, for the US, scores of 75.0 for mild AD, 67.8 for moderate AD and in the range 63.5 to 55.4 for severe AD (out of 100). Returning to your belief that preference scales, such as the EQ-5D-5L, are in fact ratio scales in disguise with a true zero, it is worth noting that in the Andersen study 26 persons with AD were reported with negative EQ-5D-5L values ranging from -0.003 to -0.53 (Figure 1). Presumably, these can be ignored in the belief held by the University of Washington expert group, that ordinal preference scores have undeniable ratio properties. Negative scores are merely inconvenient inconsistencies.	supplement, you can find further details and rationale in our Model Analysis Plan which is available here.
25.	If your team at the University of Washington had probed a little further they would have encountered a patient centric need fulfillment quality of life instrument which meets the standards for fundamental measurement. This is the Quality of Life Index for Atopic Dermatitis (QoLIAD) first developed in 2004, it has been revised and used to create interval scores in AD trials including most recently Dupilumab in moderate to severe atopic dermatitis.	Thank you for this recommendation.