Belimumab and Voclosporin for Lupus Nephritis
Response to Public Comments on Draft Evidence Report

March 12, 2021

Table of Contents

#	Comment	Response/Integration
Manufacturers		
Aurinia		
1.	**Assume 1.5 years as the average length of LUPKYNIS treatment duration.** **ICER's three-year period for therapy duration overestimates the average time patients will receive LUPKYNIS and should be adjusted to a mean duration of 1.5 years.** LN is characterized by highly objective treatment goals based on a simple noninvasive metric, the UPCR. Therefore, response to therapy is relatively simple to assess. In fact, the 2019 European update to the recommendations for the management of LN states "Evidence of improvement in proteinuria should be noted by 3 months and at least 50% reduction in proteinuria by 6 months". The dual MOA of LUPKNIS results in an expected rapid decline in UPCR in responders as detailed in both of our pivotal trials. This combination of objective, easy to assess response with a rapid mechanism of action results in LUPKYNIS US PI language guiding clinicians to consider discontinuation of LUPKYNIS if therapeutic benefit is not observed by 24 weeks. In addition, the US PI highlights that safety and efficacy have not been established beyond one year and clinicians need to consider the risks and benefits of longer durations of therapy. We believe that standard HCP LN treatment practice combined with product characteristics reflected in the US PI language will result in a mean LUPKYNIS treatment duration of 1.5 years vs the 3-year assumption in ICER's draft evidence report. Underpinning this, a survey of 96 treating U.S. physicians suggests that the majority would keep patients on treatment for no more than 1.5 years after achieving a complete renal response, as shown in Table 1. Given this information, we recommend that ICER apply an 18-month maximum average treatment period for LUPKYNIS in their model which accounts for treatment duration across responders and non-responders.	We received multiple comments about the treatment duration in the model, including American College of Rheumatology (ACR) suggesting that the treatment duration for CR/PR patients is likely to be longer than 3 years. Also, Black Women's Health Initiative (BWHI) suggested that the model should include different treatment durations for responders and non-responders. Our clinical experts agreed that the treatment duration would likely be different for those achieving response and those who did not. As such, the model is updated to incorporate a differential treatment duration of 18 months for patients in AD state. The treatment duration for CR/PR states is 3 years (as before).
GlaxoSmithKline		
1.	**Update WAC pricing and labels for both therapies** As a research payers, GSK welcomes organization, ICER has set out to objectively evaluate the clinical and economic value of prescription drugs, medical tests, and other health care and health care delivery innovations. As the evidence that ICER produces may	The revised report uses the most recent WAC and estimated net prices for each treatment, and utilizes dosing and other information from the label for the recently approved voclosporin.

#	Comment	Response/Integration
	be used to inform policy decisions made by US ICER's decision to use the most up-to-date and accurate information regarding WAC pricing, available from publicly available pricing sources (e.g., Red Book1). Further, GSK would like to emphasize the importance of including the most recent information available on safety, dosing, and available formulations, using the prescribing information for both medications.2,3 Lastly, in some instances, information presented by ICER does not align with the available data (e.g. Table 4.3). GSK posits that this is due to pooling of different trials, but asks that ICER provide a clarifying footnote in those cases. *As such, GSK welcomes ICER's decision to use the most recent WAC pricing and recommends the use of the most recent label in its review for both belimumab and voclosporin.*	Table 4.3 now includes a clarifying note about the origin of the reported values.
2.	**Steroid tapering rules and definition of non-responders in trials** Throughout its Draft Evidence Report, ICER refers to small differences in definitions of Complete Response (CR, AURORA/AURA-LV) and Complete Renal Response (CRR, BLISS-LN). GSK requests that ICER change its language as GSK finds differences in these endpoints to be meaningful, especially with respect to steroid dose requirements used to identify responders. In the Draft Evidence Report, ICER states that, in the short-term model, a minimum relative increment in utilities was applied for the proportion of patients on low-dose steroids in the BLISS-LN clinical trial. In the short-term model for voclosporin, no increment in utilities was used during weeks 0 to 8 of the trial, an increment related to low-dose steroid use was applied from weeks 8 to 16, and an increment related to no steroid use from week 16 onwards was used. GSK believes that this does not take into account the different definitions of non-responders used in each trial. • In the BLISS-LN clinical trial, prednisone dose was required to be reduced to ≤10 mg/day by Week 24 and be maintained through Week 104; otherwise, a patient was considered to be a non-responder.4 • It is GSK's understanding that in the AURORA clinical trial, while a stringent steroid taper was recommended by	Steroids consideration in the short-term model had minimal impact on both cost and QALYs predicted and would not change the results of cost-effectiveness estimates for any of the treatments. Moreover, considerations of steroid in the model lowered incremental cost-effectiveness ratio for belimumab only, since per-protocol analysis was applied for voclosporin (because of lack of data), assuming equal steroid tapering in both treatment and comparator arms of AURORA trial.

#	Comment	Response/Integration
	protocol, patients were only considered responders (i.e., renal responder) if the steroid dose was no more than 10 mg prednisone for ≥3 consecutive days or for ≥7 days in total during from Weeks 44 through 52.3 Therefore, the application of an increment of utilities related to low-dose steroid use from weeks 8 to 16 and the increment related to no steroid use from week 16 to 44 in voclosporin's short-term model may not accurately reflect voclosporin's steroid-sparing benefit. Additionally, GSK asks ICER to elaborate on rate of steroid reduction, as there are meaningful differences between the clinical trials under considerations: • In the BLISS-LN clinical trial, patients could be started on one to three pulses of 500 mg to 1000 mg of IV methylprednisolone at the investigator's discretion. The starting dose of oral prednisone after induction could have been up to 60 mg/day, and those achieving Partial Response (PR) and CRR were required to be taking ≤10 mg/day from week 24 to 104.4 • In the AURORA clinical trial, patients were started on 250 mg or 500 mg per day (depending on body weight) of IV methylprednisolone on Days 1 and 2 followed by an oral corticosteroid taper starting at 20 mg or 25 mg per day to achieve a target dose of 2.5 mg/day by week 16. Those patients achieving CR were required to be taking low dose steroids from week 44 to 52.3 Without data for mean steroid dose in the AURORA clinical trial, or both number and proportion of patients observed to achieve ≤2.5 mg/day beyond week 16, it is not possible to ascertain the degree to which applying the steroid tapering protocol might inaccurately reflect the steroid-sparing benefit attributable to voclosporin. GSK believes that these differences between steroid tapering protocol and definition of treatment failure in the AURORA clinical trial ought to be considered in the short-term model. *GSK suggests that ICER acknowledge the imbalance of their modeling approach, accounting for the steroid utilization associated* *with both medications.*	
3.	**Budget impact numbers and calculations** GSK believes that there may be some inaccuracies in the LN patient population numbers presented by ICER in the Potential Budget Impact portion of the Draft Evidence Report. In section 7.1 of the Draft Evidence Report, ICER states that there are 13,700 eligible	Thank you for pointing out this discrepancy. The results in section 7.2 were incorrect due to a copy error. The results in the revised report have been corrected to reflect the eligible population of 13,700. Details on the derivation of the eligible population estimate are found in section 7.1.

#	Comment	Response/Integration
	patients per year for each of the five years in the analysis. In section 7.2, however, ICER concludes that, for belimumab, *"Approximately 74% of the approximately 11,800 eligible patients could be treated in a given year without crossing the ICER budget impact threshold…"* GSK would like further clarification from ICER to understand why the eligible patient population for belimumab decreased from 13,700 to 11,800. *Specifically, GSK would recommend ICER include a more detailed version of its patient population "funnel" calculation in an effort to maintain optimal transparency.*	
4.	**Clinical data considered in evidence review** GSK believes that the following evidence be considered and acknowledged by ICER as it applies to the clinical evidence review: • While mentioned in the draft report, GSK would like to emphasize the BLISS-LN secondary endpoint of time to renal related events or death. Importantly, the risk of renal-related event or death at any time was 49% less in the belimumab group compared to the placebo group (HR=0.51; 95% CI: 0.34, 0.77; p=0.0014).4 • GSK appreciates that ICER recognizes BLISS-LN was a 2-year study, and importantly, a higher proportion of patients achieved the primary endpoint (Primary Efficacy Renal Response, PERR) at end of the 2-year treatment period with belimumab than with placebo, both added to standard of care. Furthermore, we would like to highlight previously referenced evidence by Petri et al (2020) that retrospectively evaluated long-term outcomes of patients who were modified PERR responders at 24 months in the Hopkins Lupus Cohort.5 Results showed that achieving mPERRb at 24 months was associated with an increased likelihood of long-term renal survival and chronic renal insufficiency–free survival in patients with LN. • Additionally, GSK would like to draw attention to belimumab extra-renal effects in the lupus nephritis population as presented at ACR 2020 and referenced in our evidence response. In addition to demonstrating efficacy in renal outcomes in BLISS-LN, positive effects were observed for the overall SLE activity in lupus nephritis patients. Patients in the	Thank you for your input. Bullet 1: This is specifically highlighted in the 3rd paragraph under belimumab in the clinical benefits section. Bullet 2: Yes, thank you for the reference to this late 2020 conference abstract. However, the same research team in a peer-reviewed article in the Journal of Rheumatology concluded that "Proteinuria within the first year of diagnosis of SLE is one of the most important predictors of end stage renal disease. Our data also confirm African American ethnicity, younger age at SLE diagnosis and low C3 as strong predictors of renal failure." This remains an area of active investigation and controversy. Bullet 3: We agree that this is a potential other benefit of belimumab and have highlighted that possibility in the Potential Other Benefits section of the report. However, we did not find any published data to support additional benefits beyond the renal benefits in patients with lupus nephritis. We eagerly await publication of the results of secondary quality of life outcomes listed as part of the phase 3 trial of belimumab that were not reported in the published report of the trial. We have added a summary of the abstract results reported in November 2020 at the ACR convergence meeting to the clinical benefits section.

#	Comment	Response/Integration
	belimumab group had a 43% lower risk of experiencing a severe SFI flare compared with patients in the placebo group (HR: 0.57; 95% CI: 0.39, 0.84). Furthermore, the percentage of patients with low SLE activity as defined by SLEDAI-S2K score <4 was greater in the belimumab group (27.8%) than the placebo group (18.8%) at Week 104 (OR: 1.76; 95% CI: 1.11, 2.78; p=0.0164).6 *GSK requests that ICER consider and incorporate the additional clinical evidence for belimumab cited above into the evidence* *review.*	
5.	**Areas in need of further clarification** Lastly, GSK would like to suggest two corrections to ICER's Draft Evidence report: • On page 4, ICER states *"Our search identified one randomized trial of belimumab in patients with LN, the pivotal phase 3 trial BLISS-LN, with outcomes at 104 weeks, and one uncontrolled trial."* GSK assumes that this statement refers to Davidson et al., 2018,7 but asks ICER to clarify and to provide in-text reference. Additionally, GSK would like to emphasize that Davidson et al. is not a study evaluating belimumab, but an evaluation of a retrospective cohort that applies BLISS-LN endpoint definitions. Therefore, Davidson et al. should not be referred to as a belimumab clinical trial but as an observational study in ICER's report. Additionally, we also request that Petri et al. (mentioned previously) be included in the evidence base. • On page 5, ICER states *"The primary differences in the study populations for the two drugs were that the AURORA trial* *excluded patients with an eGFR<45 ml/min and required background therapy exclusively with MMF, whereas the BLISSLN trial had no eGFR exclusion threshold and allowed background therapy with either MMF or cyclophosphamide. (Table D4.2)"* The exclusion criteria reported by ICER do not match the BLISS-LN study protocol, which states that *"estimated eGFR < 30 ml/min/1.73 m2 at the screening visit was an exclusion criterion."* GSK asks that ICER correct exclusion criteria in the draft report.4	Bullet 1: We have split out the references as requested, but we were referring to Furie 2020, not Davidson 2018. Thus, the additional comments on Davidson 2018 do not apply. Petri et al. above, as described, is not a study of belimumab and should not be included here. Bullet 2: We have made the correction. Thank you for pointing out the error.

#	Comment	Response/Integration
Patient/Patient Groups		
Black Women's Health Imperative		
1.	***BWHI applauds ICER's decision to perform a scenario analysis for Black patients, and urges it to include this analysis in the final LN report.*** The differential impact of LN on women of color appears to be complex and multifactorial. As ICER noted, "[d]isparities in outcomes between White and non-White LN patients persist even when adjusting for socioeconomic factors, signaling the possibility of both biological differences and the impact of systemic racism in the health care system and society" (Draft Report, at 2). Systemic racism has impacted Black, Latinx, and other people of color with respect to (a) reliable access to health care, income potential, and food and housing security; (b) inclusion within clinical trials; (c) prevalence of significant comorbidities and poor health outcomes. ICER's decision to include a scenario analysis for Black patients is an ***important*** step toward acknowledging and reducing health disparities associated with race and systemic racism; the manner in which ICER performed this scenario analysis was a ***bold, but essential, step*** that forges a path toward leveraging health economics to close inequities in care and health outcomes due to systemic racism. In particular, we appreciate that ICER: - Applied general, rather than ethnicity-specific utility values, to avoid "discounting" value associated with treatment effectiveness that would result from incorporating race-specific differences in income potential; and - Utilized cost values in its long-term model that were independent of ethnicity. We agree with ICER's decision to focus its subpopulation specificity on treatment impact and disease burden and to treat racial divergence in care cost and income potential as extraneous variables rather than legitimate inputs. As ICER acknowledges, the relatively small number of Black patients included in clinical trials of the evaluated treatments reduces the precision of ICER's calculations. Similarly, including a scenario analysis of one subpopulation may blur variable treatment response and disease burden across other subpopulations. Ideally, ICER would have sufficient data to assess and evaluate divergence in disease burden and treatment response for Black, Latinx, and Asian patients that would impact the value of each alternative treatment. BWHI, however, agrees with ICER's inclusion of the scenario analysis and	ICER acknowledges that data available on ethnic minorities are limited and of low quality. Thus, by including population analyses in the report ICER stresses the importance of studies involving different ethnic groups and discourages using data without statistical significance for policy or reimbursement decision making.

	urge it to include it in its final report. From a medical decision-making standpoint, the analysis highlights information that is likely to further ICER's goal of informing clinicians and patients as they weigh the benefits of various treatment options. The scenario analysis also provides insight into an equitable approach for valuing emerging treatments in Black, Latinx, and other underserved populations that could encourage enhanced efforts from clinical trial sponsors to enroll study participants that mirror disease state patient demographics. Moreover, ICER's methodology appropriately declines to "discount" the lives and health of non-White patients by implicitly recognizing that race-specific variability in cost of care, health outcomes, and economic potential are influenced by longstanding inequities that would be both legitimized and perpetuated if included as model inputs.	
2.	***We urge ICER to include language in its final evidence report that highlights the potential imprecisions in the base case scenario due to divergence between clinical trial populations and real-world LN patient demographics.*** ICER appropriately acknowledged that the primary source of heterogeneity was anticipated to be race/ethnicity as non-White patients typically present with more severe LN that progresses more rapidly. With respect to the scenario analysis, ICER noted that "[t]hese results are highly uncertain and highlight the need for better data on the relative effectiveness of these treatments among racial and ethnic groups who constitute the majority of patients with LN in the United States." We believe that this observation is sufficiently important to warrant inclusion in discussion of the base case scenario given that: the prognosis of patients with LN is worse in Black patients and Latinx patients, and progression to ESRD in Black and Latinx LN patients is almost nine and four times greater than in White patients, respectively. Black and Latinx patients are also more likely to rely on Medicaid coverage and far less likely than White patients to receive treatment aligning with the standard of care (SoC) (Feldman, et al). The treatment experience for the placebo cohorts in clinical trials, therefore, likely exceeds the care that Black and Latinx patients actually receive in the community. Although the "general" success rate of SoC may approach ICER's 50 percent approximation, treatment failure is more common than success in non-White patients. Although it is difficult to quantify these factors with precision, they are sufficiently important to warrant inclusion as a cautionary statement in the base case scenario. This additional cautionary statement is	ICER acknowledges that both belimumab and voclosporin trials do not include representative samples of ethnic groups. ICER already highlighted uncertainty regarding population sub-groups analysis in the report and will include a statement on need for better representation of LN population by ethnicity in the conclusions.

	supported by sufficient evidence, and would be helpful to both clinicians and payers who may otherwise reserve either of these newer therapies for patients actually failing the SoC and unintentionally subject the majority of their Black and Latinx patients to unacceptable side effect profiles and, more importantly, to continued disease progression.	
3.	***We support ICER's use of data sets that more accurately reflect the demographics of the LN population.*** ICER's Model Analysis Plan suggested reliance on disease models that did not reflect the real world demographics of the LN patient population. We appreciate that ICER has augmented its modelling with Davidson et al. (2018) to more accurately reflect the diversity of ethnicities in the US LN population.	Thank you for your comment.
4.	***BWHI is concerned that ICER's assumption of treatment continuation in non-responding patients skews their associated costs.*** We urge ICER to align assumptions on treatment duration with the FDA-approved labeling statements. The voclosporin label, for example, suggests that "[i]f the patient has not experienced therapeutic benefit by 24 weeks, consider discontinuation of LUPKYNIS" and that clinicians should "[c]onsider the risks and benefits of LUPKYNIS treatment beyond one year in light of the patient's treatment response and risk of worsening nephrotoxicity." Just as we expect that non-White patients could benefit from newer treatment options that replace the unsatisfactory efficacy and side-effect profile of the existing SoC, we believe that assessing treatment effectiveness and adjusting care plans accordingly are essential to quality care. Use of reduction in proteinuria to assess treatment response is not only appropriate within the context of ICER's review, but could be an important part of an emerging SoC that could close the racial disparities on patient outcomes by offering a standard, objective means of assessing treatment adequacy, follow-up plans, and the need for treatment plan modifications to maximize improved outcomes while reducing unwarranted side effects, adverse events, and excess costs. We expect that treatment duration assumptions for belimumab are likely complicated by its utility in Lupus beyond LN, and urge ICER to consult with the manufacturers of the assessed treatments on the expected duration of treatment in both responding and non-responding patients, and to revise its assumptions to better align with those expectations and the FDA-approved labeling.	ICER referred to received clinical advice and clinical evidence to define expected treatment duration in clinical practice for LN patients. ACR suggested that treatment of patients in CR/PR state is likely to continue longer than 3 years, however, a 3-year timeline used in the model which will underestimate the costs for patients in CR/PR state. We adopted the treatment duration of 18 months in AD state to account costs of likely longer treatment duration in clinical practice.

5.	***BWHI appreciates that ICER's model reflects the importance of reduced steroid exposure.*** Corticosteroid use contributes to development or worsening of health conditions that already disproportionately impact Black and Latinx patients, including hypertension, obesity, diabetes, and osteoporosis. High-dose steroids are also associated with a wide array of side effects impacting overall health and quality of life, including mental health issues, weight gain, and changes in appearance. Moreover, the costs of managing adverse events associated with longer-term use of corticosteroids (60 days or more) can actually be higher than disease-related medical costs. We strongly support ICER's application of a positive increment in utilities and a reduction in costs for patients treated with low-dose steroids or no steroids.	Thank you for your comment.
6.	***BWHI appreciates that ICER included childbearing potential in its set of contextual considerations and urge it to work toward mechanisms that more fully account for and quantify this important outcome for inclusion in future assessments.*** LN impacts women at the peak of their career and childbearing potential. The existing SoC includes treatments that are associated with ovarian toxicity, teratogenicity, infertility, and miscarriage. We appreciate ICER's recognition of this important consideration as the inability to start a family can have a profound impact on the lives of women impacted by LN. We also urge ICER to consider incorporating impact on childbearing potential within model inputs for future assessed treatments.	Thank you for the comment. We will be watchful of academic or industry efforts to quantify this potential benefit, but we are unaware of anything at the current time. This is why we have a large range for our value-based pricing considerations that start from a presumed ceiling price and go *up* from there. We want considerations like this to be recognized by payers and others, and we hope this will be highlighted during the discussion at the public meeting.
7.	***ICER's contextual considerations and "other benefits" are particularly important in assessing treatment value in LN.*** ICER noted that the reviewed treatment options might be associated with benefits and considerations not reflected in the model. We agree that ICER's inclusion of quality of life factors specific to LN was hampered by the lack of clinical trial data, and urge ICER to examine alternative data sources that might assist in identifying patient-preferred outcomes and the impact of emerging treatments on those outcomes as it continues to refine its processes. We also urge ICER to ensure adequate discussion and consideration of quality of life factors within its discussions leading to panel votes and ICER's final report.	Thank you for the comment. We agree that the contextual considerations and other benefits are key elements when assessing value and need to be deeply and thoughtfully integrated into value assessment beyond the numbers produced by the economic analysis. We look forward to more detailed reports of the impact of these therapies on quality of life in patient with lupus nephritis. This will be an important consideration to highlight in our discussion of value prior to votes at the public meeting.
8.	**BWHI appreciates ICER's discussion of QALY-associated shortfalls.** QALY use, without the separate subpopulation analyses presented in the Draft Report, as well as base case analyses incorporating significant adjustments to the underlying model and its inputs, presents the potential to distort value determinations	Thank you. Racial inequities in care access and delivery certainly drive health outcomes, and that's why it is only with great caution that anyone should interpret subpopulation analyses of the effectiveness of treatment. Our version of MCDA, which does not seek to assign a quantified weight to

	and perpetuate the race-specific inequities LN patients already experience. Clearly, the QALY framework predates the emerging recognition that racial inequities in care access and delivery can drive health outcomes and distort utility values. We urge ICER to continue to work toward aligning health economics with true treatment value for both White and non-White patients, including through use of methods such as multiple criteria decision analysis (MCDA) that could enhance relevance of QALY to patients of all races likely to benefit from treatment or suffer from having it withheld. Once again, we appreciate the opportunity to respond with comments to the Draft Report, and look forward to continuing to engage with your team to improve ICER's ability to capture the value of emerging treatments on the lived experience of women of color.	every possible factor, reflects our concern that the methods for estimating appropriate weights may disadvantage vulnerable patient groups. There may be unintended "winners and losers" so we think it is wiser to highlight these factors, vote on them, and leave a range within which decision-makers can apply them qualitatively as they see best. We believe that QALY shortfall results are one way to help measure the "burden of illness" faced by patients with chronic conditions like LN.
Lupus and Allied Diseases Association, Inc.		
1.	**Health State Utilities:** As we stated in previous comments and above, given the importance of the underlying health state utilities in the economic modeling, we remain steadfast in our desire to see ICER's proposed assumptions for the key health state utilities for AD, CR, PR and ESRD reflect a US lupus patient population instead of being derived from a non-US patient cohort. We have yet to locate additional US sources to consider while continuing our own research, but instead re-emphasize the study that assessed health state utilities for varying types of disease flares across a number of different country populations that highlights the fact that significant discrepancies can exist across health state utilities from country to country by disease flare type and population. We request that ICER address this type of finding.	Pollard et al. measure utility values for patients with SLE in 6 countries, though not in the US. For LN patients, Pollard et al. report only utilities in "severe renal" state, which would not be sufficient to inform the model. No source measuring utilities for model states conducted in the US was identified. Neither did we find a more representative or better-quality study. We believe our estimates represent the best possible source of evidence for utilities. We externally validated the scores to utilities of patients younger than 65 years in a cohort of North American dialysis patients (also measured on EQ-5D score)(Manns, 2007). We also addressed uncertainty in utilities by conducting one-way sensitivity analysis, which varied utility values to the lower- and upper bounds of plausibility.
2.	**Indirect Costs and Other Contextual Considerations:** We appreciate that ICER recognizes the importance of the negative impact lupus nephritis has on an individual's ability to work, to have children, and to advance in their careers as well as the burden to patient, caregiver and society and that potential benefits and contextual considerations are not fully captured by you. However, we suggest that ICER identify a mechanism to better quantify the elements of "other contextual considerations" so that these can officially be added into an economic analysis on their value versus simple statements. We are hopeful that if ICER is able to commit to updating this analysis in the future, the patient, provider, and research community will analyze as well.	Cost-effectiveness from modified societal perspective includes costs of lost productivity for LN patients (absenteeism and unemployment) and additional costs for their caregivers (due to loss of productivity and extra healthcare costs). While it is tempting to assess a complex impact of LN on patients' wellbeing and costs, lack of relevant, good quality, quantitative data, prevents the inclusion of other parameters in the societal perspective. Supplementing the model with highly uncertain, low-quality, unreliable data does not improve the assessment of LN impact but increases uncertainty in modeling predictions and so diminish the usefulness of the model.

	In LADA's previous comments we highlighted the life modifying and often life diminishing impact of SLE/LN on one's ability to attain both educational and professional accomplishments. Although we found only one productivity study so far, we will continue to search for research articles that include additional assumptions to share with ICER. We also request that you review sources from other serious disease drug reviews that ICER has completed again to see if the information may be applicable to the LN review. In addition, we have included our previous resource on productivity to reinforce the importance of value assessment report data that includes productivity and uses co-base case analysis rather than scenario analysis to inform payers in their benefit designs. It emphasizes that excluding productivity undervalues treatments and risks inappropriate restrictions on patient access to treatments.	
3.	We commend ICER for noting the negative effect to women with SLE/LN who are not able to have children or experience motherhood and due to its importance, are including this information again to reaffirm that although it may be difficult to quantify from existing literature, the quality of life impact is colossal. We also reiterate that these additional costs have the potential to increase the societal costs to a level where the cost effectiveness from a modified societal perspective may be warranted as the co-base case when added to the currently in scope indirect costs. We request that ICER formally note the limitations in their final report if literature sources cannot be identified to address these extensive impacts. We are pleased to see that ICER recognizes the access issues faced by people with SLE/LN regarding intravenous infusible therapies such as treatment costs, child, and elder care, and in some geographic areas limited availability of infusion providers/centers and/or transportation challenges as well as the time and travel required to access kidney dialysis or infusion therapy as an obstacle to care for many people. This is further heightened by safety concerns in having to leave their homes for infusion treatments during the COVID-19 pandemic. In addition, we are thrilled that ICER listened to our concerns regarding utilization management payer policies such as step therapy protocols that force	In our discussion of Potential Other Benefits and Contextual Considerations in the report, we have noted fertility impacts as an important consideration. However, we are unaware of any data that would allow us to quantify this for inclusion in the cost-effectiveness analysis.

	patients to try and fail preferred treatments that can be ineffective or result in adverse reactions. We would also like to add prior authorization requirements that delay proper patient care; switching stable patients due to nonmedical reasons resulting in inconsistent coverage, unstable formularies, and disruption in care; and copay accumulators that preclude individuals from using copay cards, coupons, or other cost-sharing programs to cover their out-of-pocket expenses to the list of payer protocols that prevent patients from receiving the most appropriate treatment.	
4.	For individuals living with SLE/LN and other debilitating diseases of unmet need, access to appropriate medication can dramatically improve disease outcome and quality of life. There is ample evidence that new innovative medicines such as targeted treatments, biologics, fusion proteins, and plasma-derived therapies may offer therapeutic advantages over conventional medicines, but these treatments usually cost more than older drugs due to their route of administration by intravenous infusion or injection and because they are not yet available as generics so produced in lesser quantities. Although costlier, these medications can reduce the severity and frequency of disease activity and decelerate its progression, in turn enabling people to lead more productive lives. Basing treatment decisions exclusively on cost rather than also including clinical considerations ignores important variations that can exist among patients in terms of safety, efficacy, and tolerability in drug classes and can discourage drug research and development, especially for diseases of unmet need with limited therapies. Scientific research shows that gender, racial, and ethnic differences in responses to treatments exist, and limiting access will greatly widen already existing health disparities. This is especially relevant given the higher prevalence of both SLE/LN in females and non-Caucasian populations. The determination of the most appropriate medication for a particular individual with SLE/LN must be made on the basis of patient acceptability, prior individual drug response and side-effect profile, and long-term treatment planning – not solely on cost. Many of these individuals already face tremendous challenges in their daily lives and do not need another roadblock to further complicate their medical care. We feel that it is imperative that physicians' rights to make medical decisions in the best interest of their patients are preserved in order to ensure ethical accountability and guarantee patient access. Furthermore, the determination of the most appropriate medical treatment is best	Thank you, we agree entirely with your very important statements on these broad points.

	accomplished by open and transparent communication between the patient and the health care provider who is educated and ethically bound to treat to the individuality o that patient. Given the heterogeneity of SLE/LN and the patient population, we must remain vigilant in safeguardin the doctor/patient relationship while promoting unfettere access to vital life-enhancing and lifesaving treatments.	
5.	We would like to suggest that the degree to which payers have incorporated these findings into their clinical and coverage assessments for Belimumab and Voclosporin, and the extent to which they may or may not be allowing access to the new medications be examined as part of this review. If there is value in these products, but people are unable to access them due to onerous or restrictive coverage, we may actually be advancing the inequities that we are hoping to circumvent, especially across the Medicare and Medicaid programs which cover a large percentage of the overall US LN population	This is a good idea. There are no coverage policies for voclosporin yet, but these will emerge over the coming months. We have promulgated Fair Access Criteria by which patient groups, clinicians, and other policymakers can evaluate coverage to determine if they are reasonable, and we ourselves are currently engaged in a project to evaluate the coverage of 27 drugs. At some point in the future when coverage for LN for both these drugs is available, we might well look at the coverage for both these drugs, and we would encourage you or others to do the same!
6.	Lastly, as you state in Table 5.2. Potential Other Benefits or Disadvantages that the SC formulation of Belimumab may supplant the IV infusion in the real world, we ask that this be revisited in a specified time period to reassess the drug's effectiveness and value as time progresses, especially given safety concerns while the COVID-19 pandemic lingers.	Thank you. We agree and will revisit this if the SC formulation is shown to benefit patients with LN.
Lupus Foundation of America		
1.	**SUMMARY** Our most serious concern with the draft report is the reliability of the data supporting the cost-effectiveness of belimumab for Black lupus nephritis patients compared to that for non-Black patients. We strongly believe the data i inadequate to draw any conclusion regarding the relative cost-effectiveness of treatments for Black patients. To suggest that a treatment is less cost-effective for Blacks without solid supporting evidence may lead doctors and Black patients to believe this treatment is not an appropriate choice for them and may put Black patients at risk of significant health care access challenges for an important new treatment option. While we encourage ICER to objectively report what is known and unknown about subgroup effects for voclosporin and belimumab, we insist that ICER remove the Black subgroup cost per QALY results from the report due to insufficient supporting evidence of cost differentials.	ICER added the analysis for sub-populations on a request of patient organizations. We agree that reliability of sub-analysis for the Black population is low; this remark though is relevant not only for belimumab, mentioned in the comment, but also for voclosporin. The text of the report already comments on high uncertainty of this analysis. To avoid misinterpretations, we deleted the results of sub-population analysis from the table and added a statement that the results of sub-population analyses should not be used in policy and/or reimbursement schemes.
2.	**OVERALL METHODOLOGY** As noted above, we are pleased that ICER's analysis led to a main conclusion that both therapies are cost effective. We believe it is important, however, to point	Quality and representativeness to the US population were the selection criteria for longitudinal data. Davidson et al. study (2018) is the latest longitudinal analysis in the US with relatively (for LN studies)

	out a few aspects of the analysis that could be strengthened or approached differently. Although we recognize that ICER is utilizing studies and data currently available; in some cases, this information does not reflect the real-world experience of people with lupus—with either their disease state or likely treatment experience. As both treatments have been approved recently for the treatment of LN, ICER understandably relied largely on clinical trial data for its analysis. The drawback of this approach, however, is that the trials were designed for the specific purpose of demonstrating safety and efficacy for regulators and not to demonstrate value in a real-world setting. Trials, by design, are not reflective of the general population and test therapies in a highly controlled environment on patients selected because they meet certain criteria. Although this is not necessarily by design, trial participants are typically less diverse than the overall patient population and Black and Hispanic people are historically underrepresented in trials. LFA understands these challenges and is pursuing initiatives to increase diversity in trials given the outsize impact of lupus on people of color. In addition, the data used to develop some of the baseline measures are unlikely to reflect the real-world experiences of people with lupus. First, the studies which ICER relied on to establish a baseline for end-stage renal disease (ESRD) events and death are dated and only include a small number of patients. In particular, the Davidson et al. and Chen et al. references, which use data from patients in the 1980s, are problematic. Secondly, candidates for both therapies are likely to be sicker than patients in the study ICER used for its baseline cost model. Patients in the study were treated with immunosuppressive drugs and corticosteroids, but voclosporin and belimumab are both indicated for patients who have not responded to earlier treatments and whose SLE has progressed to LN. ICER also utilized claims data largely focused on commercial and Medicare Advantage plans, whose patients tend to be healthier than patients with Medicaid or traditional, fee-for-service Medicare. Many LN patients are covered by Medicaid due to the financial impact of their chronic disease and preexisting economic disadvantages. We have a particular concern with the belimumab methodology. Belimumab has been included in the standard protocol for treating non-LN SLE for many years and is now approved for LN. Most LN patients have non-	large population size. The study of Chen et al. was not used directly in the modeling, only the proportions reported, to fulfill the data from Davidson et al. No alternative sources were suggested in public comments received. ICER did a scenario analysis to consider worse longitudinal survival in non-trial population. It is likely that with sicker patients, the drugs will be less cost-effective. Thank you for pointing this out: we added a relevant statement to the report. LN patients can claim to be covered by Medicare if their disease progress to ESRD state. We substituted costs of ESRD state with costs of ESRD alone or disability qualification for Medicare-covered patients. ICER acknowledges that benefits of belimumab go beyond what can be captured in the LN model. This recognition is reflected in the results and conclusion of the report. ICER gives preference to real-world, large population, contemporary data whenever possible, and accepts recommendations on data from different stakeholder groups. When no such data are available, ICER needs to rely on available evidence and acknowledges the data limitation in the analysis.

	nephritic SLE disease manifestations. The QALYs attributed to belimumab should include the treatment effect for both nephritic and non-nephritic disease manifestations. Otherwise, all of the cost is being compared to only part of the effect. We encourage ICER to use real-world, large population, contemporary data whenever possible. Furthermore, recognizing that data for certain aspects of ICER's analysis may be limited, we urge ICER to caveat any conclusions drawn from data that is not aligned with the real-world experience of contemporary patients and/or not statistically reliable.	
3.	**SUBPOPULATION ANALYSIS** As noted in our previous comments and by ICER in the draft evidence report, lupus disproportionately affects women, especially women of color. As such, we commend ICER for efforts to include an analysis of the Black subpopulation in the draft evidence report. We are concerned, however, that the data available on this subpopulation is insufficient to support conclusions, especially the quantitative finding that belimumab might be significantly less cost-effective for Black patients than non-Black patients. The BLISS-LN trial for belimumab did not produce statistically significant findings for Black patients, the only reported subgroup. The odds ratio confidence interval for Black population treatment effect ranges from a negative effect to an effect well in excess of the non-Black population (Furie, 2020, Supplement, Figure S2). Therefore, no subpopulation conclusions can be drawn from the study. For an intuitive understanding of the statistical unreliability of the Black population analysis, one only needs to observe that if just one more Black person in the trial had responded to the treatment or if one less Black person had not responded to the treatment, nearly all of the treatment gap between the Black and non-Black participant groups would have closed (Furie, 2020, Supplement, Figure S2). The draft report also omits information about the Hispanic subpopulation, another group at greater risk of developing lupus and for having worse health outcomes. Although the BLISS-LN trial did not report out on belimumab results for Hispanic participants, the AURORA trial for voclosporin did. The trial found a statistically significant treatment effect for Hispanic patients (Arriens, 2020). This positive result is not mentioned in the draft evidence report.	ICER added the analysis for Black sub-populations (the largest ethnic sub-group) on a request of patient organizations. We agree that reliability of sub-analysis for the Black population is low; this remark though is relevant not only for belimumab, mentioned in the comment, but also for voclosporin. The text of the report already comments on high uncertainty of this analysis. To avoid misinterpretations, we deleted the results of sub-population analysis from the table and added a statement that the results of sub-population analyses should not be used in policy and/or reimbursement schemes. Also, please note that the results for the Hispanic population are worse than for the White population. Using data on CR and PR among Hispanic population from the Aurora trial would results in ICER above the threshold ($168,539 per QALY). While we report the results on Black population to stress the importance of clinical data on sub-populations and existing uncertainty around these values, the analysis for Hispanic population is not reported.

	While it's important that ICER discuss subpopulation data in the final evidence report, and we encourage ICER to do so given the significant impact of lupus on people of color, the discussion needs to include all subpopulations and report the strengths and weaknesses of the underlying studies and not lead readers to unsupported conclusions that could negatively impact access to important new treatment options. Furthermore, cost per QALY differences for subpopulations should only be reported when the data used to estimate the cost difference is statistically reliable.	
4.	**LIMITATIONS OF MODELING TIMEFRAME** Lupus is a chronic disease that, even when treated effectively, must be managed throughout a person's lifetime. The model used by ICER, however, focuses on LN patients being treated for three years and then having a stable disease state for the rest of their life. In reality, people with lupus experience changes in their symptoms over time; worse symptoms during disease flares and improved symptoms at other times. Even if lupus patients are able to go off one or more treatments for a period of time, they may require additional treatment should a disease flare occur. ICER's current model does not account for such changes in symptoms or treatments, which are possible within a three-year time period, and almost certain to occur over the lifetime of a person with lupus.	Partitioned-survival model is a standard methodology to model disease progression. The model does not just have "stable disease" over the patient lifetime, but relies on longitudinal data reporting LN progression over time to estimate the proportions of patients in CR/PR, AD and ESRD over time. The changes in treatments for patients require over their lifetime is reflected in costs of the disease states, in particular costs of AD state or ESRD when the disease progresses. Increase of complexity of the model without data to inform these transitions would mean increase in modeling uncertainty and so uncertainty in cost-effectiveness predictions.
5.	Although people with lupus are likely to experience change in their disease and symptoms over their lifetime even afte receiving treatment with either therapy, there is certainly value associated with a slower disease progression that ma occur as a result of this treatment. In Black and Hispanic populations, which generally progress more rapidly to ESRI and death, a treatment that will move them out of active disease status and minimize long-term damage has even greater value. ICER notes that LN tends to progress more rapidly in Black and Hispanic patients, implying this larger gain, but the subpopulation analysis did not account for th differential. The value assessment for the Black and Hispanic subpopulations should factor in the likelihood of avoiding or reducing the need for more intensive and costl medical care and delaying death from ESRD.	Limited data suggests that LN may progress rapidly in Black population, however, there are no studies that quantify this difference in disease progression between the ethnic subgroups (for it to be included in modeling). A scenario analysis with assumed worse longitudinal survival resulted in a higher incremental cost-effectiveness ratio (lower cost-effectiveness) of drugs for these patients. We did not want to include arbitrarily defined lower survival in sub-group analysis to avoid discrimination of ethnic sub-groups.
Lupus Research Alliance		
1.	While we recognize that the cost-effectiveness for both medications was very positive, we will reiterate our response to the Modeling Analysis Plan (MAP) submitted to you on December 10, 2020. Our concern is that the	Thanks. We agree that LN patients can claim to be covered by Medicare if their disease progress to ESRD state. We have used the ESRD costs of $103,029 and updated them to 2019 values (using

	medical costs being used in the model to assess the economic value of the LN treatments under review is low and could therefore impact the cost-effectiveness. As noted in our submission to ICER dated August 26, 2020, the LRA is working with the National Minority Quality Forum (NMQF) on the development of a data warehouse of lupus claims data - the Lupus Index. As stated in our response to the MAP, we are concerned that the cost data for medical care being used in ICER's model is low based on our evaluation of Medicare data for people with LN. The health care costs for LN patients used by ICER are based on a paper by Bartels-Peculis which are derived from a predominantly commercially insured population – 80% commercial and 20% Medicare Advantage. Using that analysis along with other sources, the annual cost in the end stage renal disease (ESRD) health state reported in Table 4.4 of the Draft Evidence Report is $104,685 based on 2014-2016 data for 1,039 people and inflated to 2019 values. In our review of 2016 Medicare Fee for Service costs, we found 3,624 people with LN and ESRD, with an average cost of $103,029. When adjusted to 2019 values is $111,752 or $7,067 higher than the amount being used in the model.	PHC and PCE indices as recommended in the ICER methods guide).
2.	ESRD is a criterion in which people may become eligible for Medicare. Until January 1, 2021, people with ESRD as their Medicare eligibility criteria could not join a Medicare Advantage plan. It is our belief that in order to assess the cost of people with ESRD, it is essential to include Medicare beneficiaries on original Medicare (or Medicare fee for service). The above noted source used by ICER includes Medicare Advantage. The model used includes insurance plan members with at least one claim with an LN diagnosis code in any diagnostic position who had both medical and pharmacy coverage for the years 2014 through 2016. Using these criteria, they came up with 1,039 patients with an average per year cost of $45,469. We used the Lupus Index to replicate these patient-selection criteria in original Medicare for 2016: Medicare beneficiaries with parts A, B and D (inpatient, doctor and outpatient, and prescription drugs, respectively), with at least one LN diagnosis code in any diagnostic position.	See above.
3.	It is critically important to review administrative data sets for both private and public insurance coverage to determine cost and utilization for people with LN. An analysis of commercial, Medicaid and Medicare data in which LN cases were defined requiring at least two visits	No source reporting costs for model states conducted in the US was identified. Neither did we find a more representative or better-quality study. We believe our cost estimates are based on the most appropriate and representative sources. We

	to a nephrologist or at least two LN diagnoses found the prevalence of LN to be 15, 31 and 40 per 100,000 in each database, respectively (Gandhi, 2013). The result was an estimated 63,256 LN patients in the U.S. Based on this more stringent case definition, the analysis proposed by ICER is based on a data set representing less than 2% of the estimated population with LN whereas the Lupus Index data set represents about 20%.	have used the ESRD costs of $103,029 as suggested and updated them to 2019 values (using PHC and PCE indices as recommended in the ICER methods guide).
Partnership to Improve Patient Care (PIPC)		
1.	**In order to adequately capture the heterogeneity of lupus nephritis patients, ICER should be producing ranges, not averages.** ICER acknowledges that lupus nephritis affects certain populations, in particular women and African Americans, more severely than others. This reality combined with the variance in terms of both disease severity and level of symptoms lupus patients suffer by stage of disease means that the reporting of a single estimate of the cost-effectiveness for each therapy is unlikely to be helpful in informing payor decision-making in practice. A larger point with respect to value assessment reporting is that the archetypal cost-effectiveness model relies heavily on producing effect size based on population averages, and findings specific to minorities are rarely released in final results. It is well established that the generating and reporting of differential value assessment across subgroups leads to substantial health gains, both through treatment selection and coverage. If ICER is to take seriously its role of informing health policy decision makers about the value of new therapies, it needs to move away from the assumption that all patients are the same and that the value to each patient can be determined by the estimation of the average value to a patient archetype.	We agree and always produce ranges of cost-effectiveness findings across different thresholds, while also performing univariate and multivariate sensitivity analyses to explore heterogeneity, and scenario analyses to examine specific subpopulations. Please note that reporting findings across subgroups does not always lead to "substantial health gains." For instance, taking action on findings that suggest relatively poor cost-effectiveness among racial minority groups would be inappropriate in our view, and we have made this point in our report. We hope you would agree.
2.	**The QALY is not an appropriate metric for use in lupus nephritis, and the utilities used do not paint an accurate picture of burden of disease in the United States.** In ICER's one-way sensitivity analyses, two of the inputs to which the cost-effectiveness ratios were most sensitive were (1) utilities for patients with active disease and (2) utilities for patients in complete response. This suggests that these two inputs are amongst the strongest drivers of the cost-effectiveness ratios. PIPC has highlighted the flaws inherent in the QALY on numerous previous occasions. We would like to reiterate the holistic discriminatory impacts of the QALY	We had extensive conversations with patient groups and manufacturers seeking the best possible way to translate treatment effects into quality of life benefits for patients with LN. The QALY and evLYG are important tools to try to help bring fairness into thinking about value across different treatment areas, for we would not want to disadvantage patients whose symptoms are less visible to the public (e.g., depression) or that might be the object of stigma (e.g., epilepsy). Please feel welcome to criticize without suggesting a better alternative, but we are always hoping to

	and note that it is a particularly concerning metric in the study of treatments for lupus nephritis. Recent studies have suggested that the EQ5D is at best a moderate proxy for disease-specific measures of quality of life in lupus, and at worse a weak one. Even setting aside that the use of the QALY generally is not a good fit in studying lupus nephritis, the specific utilities used in ICER's assessment are not a good proxy for a typical American lupus nephritis patient. The data used for these inputs come from an old Swedish cohort and a small Thai study of eighteen patients. This will lead to a misleading assessment for two primary reasons. First, the ICER model is a simple model with a small number of health states, meaning that each mean utility for each health state will hide a considerable level of heterogeneity across a true lupus population. Second, there is a large geographic and demographic variance in the burden of lupus nephritis, which ICER acknowledges. Black patients tend to have worse outcomes, so relying on Thai and Swedish data sets will not paint an accurate picture of the burden of disease or value of effective treatment.	receive constructive criticism with feasible alternatives that might offer advantages. Manufacturers were satisfied we were using the best available data sources. And we can –and have—called for manufacturers to include a more representative patient population in their trials. But we cannot control the fact that they performed their trials without paying attention to this issue, and I hope you will join us in holding them accountable.
3.	**ICER needs to look to a wider set of outcomes in its definition of "value."** The value of a therapy for a condition like lupus nephritis has additional facets of societal value that go beyond simply summing the average patient HRQoL-based utilities of new therapies. These include potentially greater benefits of these therapies for Black patients, the innovative nature of the new therapies' mechanism of action, and the lack of FDA-approved therapies prior to the availability of these drugs. A simple cost-per-QALY measure of value in a disease such as lupus provides a very limited view of actual benefits of new treatments. Numerous other international health technology assessment bodies have widened their scope to address these important aspects of value that stretch beyond the simple cost-utility framework. The two most obvious justifications in the case of lupus are to encourage innovation in a disease space that affects particularly vulnerable and underserved populations, and to encourage innovation in a disease that has very few alternative therapies with a goal of providing patients with needed treatments. Any true measure of "value" must incorporate benefits beyond a narrow health care systems view to the accrual of the resulting reductions in systematic health	Thank you for your comment. We are glad you view our "potential other benefits" and "contextual considerations" as valuable parts of our report and of our deliberation. We believe the factors you list should be important elements in an overall judgment of value and that is why we ask our appraisal committee members to vote on them separately AND to integrate them into their overall value vote. Please join us in signaling to payers and others that these factors are important.

	inequalities and in bridging of the gap in choice of therapies between those diseases that have numerous treatment options and those that have few or none. These are aspects of health care that numerous social preference studies have shown that people and societies value above pure allocative efficiency.	
Patients Rising Now		
1.	Before presenting our comments in those areas, now that voclosporin has been approved by the FDA, and important pieces of information accompanied that approval are available – including some black box warnings – we strongly recommend that ICER redo and reissue its draft report to allow for additional public comment before moving to hold a meeting with its advisory committee, and finalizing a report.	We do not believe the FDA approval has presented important new changes that would require a reconsideration of the draft report. The announced price from the company will be incorporated as part of the changes in the revised Evidence Report.
2.	People-Centered Perspectives We appreciate the outreach that ICER made to patient groups and the information shared in the draft report's Section 2: "Patient and Caregiver Perspectives." And we share ICER's frustration that the clinical trials on the two specific medicines newly approved for treating nephritis in people with lupus did not include evaluations of quality of life or other real-world metrics important to patients. We believe that those deficits highlight the need for additional discussion and advocacy for inclusion of such metrics in all critical trials, rather than potentially leaving that to follow-on studies. We also agree with ICER's observation that having an oral treatment option may be of significant value to patients, particularly those with travel or mobility limitations.	Thank you for your comment.
3.	We note that ICER didn't reference its own very recent work on chronic kidney disease to bring some context about how this condition can affect overall quality of life. We find this omission disappointing. If ICER is so compartmentalized that it cannot recognize its own related reports, then we must question if ICER understands and is capable of promoting team-based care, value-based systems of care, and reimbursement mechanisms to promote those advances that are widely seen as potentially benefitting both patients and the overall U.S. health care system. As we consider the scope of the draft report, we are disappointed in ICER's overall presentation of lupus nephritis as a clinical condition. Like too many clinicians, researchers and analysts, the draft report is too tightly focused on nephritis as a sequela of lupus. We are very concerned about this very narrow scope because people with lupus who may develop nephritis as part of their	Thank you for the comment. The recent CKD final report was posted on 3/5/2021. While there is significant overlap in the experience of all patients with CKD, we feel that CKD due to lupus is qualitatively different from CKD due to other causes such as hypertension or diabetes. Moreover, the other report focused on anemia and not CKD itself. For those interested in the anemia in CKD report, it may be found at: https://icer.org/wp-content/uploads/2020/10/ICER_CKD_Final_Evidence_Report_030521.pdf We agree that patients with SLE suffer from more than nephritis, but the published data on the therapies reviewed in this report focus on renal outcomes and have not reported on the impact of these therapies on the whole person. We encourage Patients Rising Now to pressure manufacturers to capture and report data on more

	myriad manifestations from having lupus are not – and should not be seen as – "kidneys who have lupus." The importance of this type of whole-person focus is clearly stated in the Lupus Patient's Voice report that was conducted in parallel with the FDA's Patient-Focused Drug Development Initiative. The Report "was created by the FDA to allow regulators to more effectively understand, in a systematic manner, the unique perspective of people with diseases such as lupus to better assess the risks and benefits of drugs under review." As that report states, "Lupus is a chronic, **systemic**, and often disabling autoimmune disease **with an unpredictable course and inadequate treatment options**." (emphasis added) The report also discussed the high incidence of other autoimmune diseases in people with lupus, underscoring the need for whole-person considerations in their clinical care. ICER needs to do a much better job of encompassing the whole-person concept of value into its work beyond the discussion in Section 2 of the draft report related to symptoms such as fatigue, and life choices that may be limited because of disease progression. Those discussions are most useful when ICER incorporates those very important issues in its analysis. Unfortunately, in this case, ICER did not do so. We realize that without data, inclusion of such factors is difficult, but that cannot be an excuse for disregarding those factors entirely. And for important issues where there is limited data, that uncertainty should be incorporated into the draft report's analyses, conclusions, and discussions to a much greater extent than ICER has been doing.	holistic patient reported outcomes as part of future clinical trials for treatments of lupus nephritis.
4.	And lastly, given that the FDA approved label for voclosporin contains a black box warning, ICER should include a discussion of the significance of such a warning for patients, and how that information should be considered as part of patients' shared decision-making with their clinicians.	Thank you. We have added the description of the black box warning which was published after our draft report was released.
5.	Modeling, Projections and Assumptions The draft report makes an assumption about the price of voclosporin that was based on a single report's four-years old guestimate. That assumption was clearly very significantly wrong, and for very predictable reasons: The old assumption that voclosporin would be priced at a 10% discount to belimumab, (which was four years away from getting a secondary approval for lupus nephritis), was clearly a broad swath "placeholder" that was the	The revised report has been updated to reflect the announced list and estimated net prices for voclosporin, now that voclosporin has been approved. Our reports always point out that any placeholder prices are only assumptions and provide threshold prices for comparison with WAC and net prices when they become available.

	same as three other potential treatments in the report, and apparently based on the premise that later entrants in a treatment area would be priced at a discount to gain market share. This type of "placeholder" may be appropriate when there is no information about the clinical (and other) benefits of each treatment. HOWEVER, the draft report's Figures E5 and E8 (copied below) clearly show the QALY benefits of voclosporin being separated from standard of care to a greater extent than is the case for belimumab. We also note the different QALY scales on the x-axes in those Figures, and their size in the draft report. (The figures above are the actual size as in the draft report.) Using the same x-axis scale in both Figures and making each Figure the same size in the draft report would have been a much better, clearer representation of the data.	The new report standardizes the figures to have the same x-axis scale in both of them.
6.	Now that voclosporin has been approved by the FDA, its actual list price and reported net revenue per patient have been reported. The estimated revenue of $65,000 per year to the company (which we assume is equivalent to the net price since it is much lower than the reported list price of $144,175 based on $3,950 for a ten-day supply at full dosage), represents a cost per QALY that – according to our analysis of the information ICER included in the draft report – is approximately 25% less than the cost per QALY for belimumab. Given that the definition of value is benefits (which could include clinical, patient, health system, and society benefits) divided by cost, the company's reported pricing for voclosporin seems to be completely appropriate, and since it is orally administered, an even higher net price could be justified. That is, the company's pricing for voclosporin reflects the clinical and other benefits it provides. It could be asserted that ICER's draft report (which was released on the same day as the FDA's approval of voclosporin) provided data and rationale for the company's pricing decisions. In that vein, some may point to ICER as reason for this new drug having a higher price than previously projected. However, as all good analysts and researchers understand, correlation does not prove causation. We are much more inclined to believe that the company understood their own data, could compare it to that of existing treatment options – including belimumab – and derived a price (including expected rebates and discounts, etc.) to determine a price consistent with its value to the patient, society, and	Thank you. We will provide an updated estimated net price in our Evidence Report based on discussions with the company. The way they are presenting their net price is not consistent with the way most companies do it so we will describe that in our Evidence Report.

	the health care system that would also enable it to have favorable reimbursement and coverage by payers and adoption by clinicians.	
7.	And lastly, the newly approved label for voclosporin includes guidance for lowering the daily dosing for people with reduced kidney or liver function. We did not see that adjustment in ICER's modeling assumptions. We would appreciate ICER providing insights about that clinical situation. For example, did ICER include that reduced dosing into its modeling, did ICER not know about such dosage adjustments in the clinical trials or from the deliberations by the FDA's advisors, or was it assumed that the number of people who would be using such lower dosages was not knowable or would be very small, etc.?	Since the first draft of the report considered placeholder price for voclosporin, daily dosing was not included in the calculations. The updated analysis uses reported list price and average daily dose of voclosporin in AURORA trial, provided by Aurinia.
8.	<u>Conclusions</u> Patients Rising Now is pleased that people with lupus – should they have or develop nephritis – now have two new and better treatment options that are both clinically and cost effective. We are glad that ICER's draft report reached a similar conclusion. However, given that voclosporin has now been approved by the FDA, we strongly urge ICER to redo its work on the draft report based upon the now available FDA label and price information, and reissue an updated draft report for further comment by the entire array of stakeholders – particularly patient groups and clinician experts.	Thank you. We have updated the report as you have suggested. We look forward to continued dialog at the public meeting.

#	Comment	Response/Integration
Other		
American College of Rheumatology		
1.	**Actual Cost versus Cost Estimates** We note that the ICER evidence report was drafted before the approval of voclosporin. This new therapy received FDA approval on January 22, 2021. With this approval, the actual cost of the drug is now publicly available. We encourage ICER to revise the cost estimates using the drug's actual price rather than rely on the ICER estimated cost. There is a significant difference between the estimated cost and the actual price which will impact the QALY result. Decision-makers reference ICER analysis for drug benefit formularies and other pharmaceutical drug policies. This analysis must use the exact price information for both prescription drugs being evaluated.	Thank you for your comment. Actual costs of voclosporin are used in the updated ICER model.
2.	**Evaluation Beyond Cost** ACR recognizes that ICER's evaluations are designed to allow for conversations surrounding the cost and clinical evidence of treatments as they enter the market. However, we remain concerned about the assumptions used in this evidence report. We know that decision-makers use this information to make drug policies and make these decisions without data to base long-term assessments. Specifically, we are concerned that the assessment assumes three years of treatments. Studies have shown that while there may be clinical remission, it does not mean a histologic remission. Therefore, there is a considerable risk for ongoing damage if treatment is stopped. It is crucial to consider the overall costs and benefits of these treatments for our patients. Specifically, we note that with the ongoing use of these medications beyond the three years outlined in the report, accrual of damage due to LN will likely be reduced and will prevent longer-term and more costly renal replacement treatments related to End-Stage Renal Disease (ESRD). Without the considerations and discussions of the long term treatment of the ongoing disease, long term ability to minimize corticosteroid exposure, and ability to avoid renal replacement therapy, the report does not provide a clear and precise evaluation of these treatments from a clinical or cost perspective. Further, throughout the evidence report, ICER suggests that belimumab may be less favorable with regards to cost. However, it must be acknowledged that belimumab has known benefits for	While voclosporin and especially belimumab may be used longer than three years for patients with CR/PR, ICER does not consider longer treatment because only short-term (1 and 2 years) data available on clinical benefits of the drugs for LN management. This increases uncertainty of the predictions of long-term clinical benefits of treatments. BLISS-LN trial has demonstrated that CR rate increases up to 12 months and does not changes between the first and the second year. Thus, ICER considers that there is no evidence to assume additional clinical benefits above the trials' endpoints. Considering costs further than 3 years without additional clinical benefits would underestimate cost-effectiveness of treatments. The model outcomes do not instruct the policies about when the treatment should be discontinued for the patients. We agree with a caveat that complete benefits of belimumab for SLE patients cannot be fully assessed with LN model. This is now acknowledged in the report.

	other systemic lupus erythematosus (SLE) features for which voclosporin has not yet been studied. These additional benefits allow us to presume that belimumab has a more favorable long-term risk profile. This lower risk profile coupled with more utility for other SLE factors may make belimumab a valuable option for chronic use. We urge ICER to include a caveat in the report noting the shortcomings of the assumptions and limited analysis that does not account for a more holistic review of belimumab and voclosporin. We fear that without this caveat, insurance companies will put significant restrictions on the use of belimumab based solely on cost without considering the additional benefits of the treatment and the long-term care/cost algorithm.	
3.	**Subpopulation Analysis Concerns** The ACR is concerned with the subpopulation analysis within the document. We believe the document omits published reports on health disparities and poorer SLE outcomes, particularly in the black and Hispanic populations. With limited data on the black population in the clinical trials for either drug, real-world data on this population is nonexistent. Without this real-world data, we fear that the message to the black community and payers will minimize the benefits of belimumab and voclosporin for this population's quality of life. Additionally, we note the minimal mention of the Hispanic population in this document. This population experiences LN earlier with greater severity than the white population. The limited data mentioned provides an inaccurate assessment of these two medications on the black and Hispanic populations. Without more robust data points, any mention of these subpopulations should be removed to prevent unintended consequences when decision-makers considered these treatments in their drug policies.	We agree that the impact of SLE differs in the subpopulations that you highlight. The last sentence of the first paragraph of our background section states the first disparity that you highlight: "The prognosis of patients with LN is worse in Black patients and Hispanic patients." And we give two of the many citations supporting this important disparity. We wholeheartedly agree that there has been insufficient attention paid to studying the impact of these interventions in both Black and Hispanic populations. We hope that this will be brought out in the public discussion of the limitations of the evidence base and the importance of focused research in these populations in order to reduce the uncertainties about the relative benefits and harms of both interventions in these key subpopulations.
4.	**Clarification of dosing** We note that there is a discrepancy within Table 1.1. The document states that Benlysta infusion occurs every two weeks. However, Benlysta is administered every four weeks after a two-week loading dose.	Thank you. We have corrected the dosing in the table.
Paul Langley		
1.	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	Thank you, your concerns are noted. As we have expressed before, we (and most health economists) are confident that changes in the EQ-5D (and other multi-attribute utility instruments) do have ratio

	claims made for the value of products; they cannot be evaluated empirically nor can the claims be replicated. You models also violate the fundamental axioms of measurement theory. While you might view these 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. This is why I have coined the term impossible or I-QALY as you and many others insist in believing that ordinal utilities have multiplicative properties. With classical test theory , instruments are typically comprised of ordinal level items on a Likert response scale (the EQ-5D-3L uses three response levels for each of five attributes or symptoms). They suffer from having an unknown or inconsistent difference between the levels on the scale. This makes these ordinal level items problematic when trying to compare results between patients, as well as violating the assumptions of most statistical tests. This conclusion rests on the failure to recognize the limitations imposed by the axioms of fundamental measurement, in particular the application of conjoint simultaneous measurement to measure non-physical attributes. As it is you continue to focus on constructing simulated QALY claims yet we know that the multiattribute utility score (typically the EQ-5D-3L/5L) is an ordinal scale. It cannot support multiplication which is required to transform modelled time spent in a disease state to its quality adjusted time equivalent. This means the I-QALY is a mathematically impossible construct. By extension, not only are lifetime incremental cost per I-QALY claims impossible, but the attempt to generate pricing recommendations (e.g., the notion of a 'fair price') through the application of nominal cost-per-I-QALY thresholds is similarly a waste of time. Hopefully, manufacturers and health system decision makers will not take this effort seriously. Although you have long maintained that multiattribute utility scores have 'hidden' ratio properties it is clear that they can generate negative values or states worse than death. At the same time, if the EQ-5D -3L is taken as a case study, it should be noted that it lacks dimensional homogeneity is capturing five separate attributes with their own characteristics. The algorithm that is used to create scores is the best fit to the data, with rules to ensure that this occurs. The resulting scores are not unidimensional and lack construct validity. The EQ-5D utility score papers you rely on fail to recognize the ordinal nature of the scores.	properties. The EQ-5D value sets are based on time trade-off assessments (which are interval level), with preference weights assigned to different attributes. We fail to see why this should be considered as an ordinal (ranked) scale. The dead state represents a natural zero point on a scale of health-related quality of life. Negative utility values on the EQ-5D scale represent states considered worse than dead. We do not find that this lacks face validity.

	While the University of Sheffield Modelling group no doubt shares your views on the hidden ratio properties of the EQ-5D-3L, it would have been useful if they had defended their choice of EQ-5D as a 'measure' and not just a score. Since the work of Stevens in the 1940s and the development of Rasch Measurement Theory (RMT) in the early 1960s with the introduction of conjoint simultaneous measurement to address issues of non-physical attributes, it is clear that if we are to emulate the physical sciences then the focus should be on measuring single attributes (measurement precedes statistical analysis). As RMT makes clear, if we are to measure latent attributes then we need a framework for translating ordinal to interval scores. Simply fitting data to observations is not measurement. As you insist on utilizing multiattribute utility scores, two comments are relevant in this model. First, you had to search the literature to find 'appropriate' measures as there are no EQ-5D scores for the lupus nephritis populations in the US; and second, the choice of the EQ-5D utilities, if your previous models are any judge, they yield imaginary modelled utilities are little different between comparator arms. This means that costs will dominate and lead almost inevitably to threshold recommendations for substantive price discounts. May I suggest, with the launch of ICER Analytics that you accompany your report with access to the ICER Analytics Sheffield model. This will allow those interested in experimenting with various assumptions, particularly utility values, to see the impact of competing scores. Of course, this would open the doors to a possible multitude of competing models and pricing recommendations. This opportunity is detailed in a recent commentary.	
2.	Although only reported on briefly, I note you engaged with lupus nephritis patients and caregivers. While I appreciate this, you do not seem to have taken this to the logical conclusion to assess the impact of the two therapies on patient and caregiver needs. As you will appreciate, the symptoms captured in the EQ-5D-3L or other multiattribute instruments may not be relevant in many treatment situations (or only marginally so). This means that an instrument such as the EQ-5D, with scores reflecting the preferences of a community may fail to capture concerns; it lacks sensitivity. Symptoms may improve but the needs of the patient may not be responsive. This, of course, has been recognized for some few decades. It would have been useful if either ICER or the Sheffield group could have considered the	Thank you. I think we share with you the hope that drug makers and clinical researchers will begin incorporating more outcome measures related to caregiver effects of treatment. We frequently highlight this in our policy recommendations.

	extent to which new therapies can better meet the needs of both patients and caregivers. This brings us back to the measurement of latent attributes such as needs based quality of life. We have had techniques available for some 60 years (RMT) with the development of patient and caregiver centric instruments since the early 1990s. These meet the requirements of fundamental measurement, creating interval measures to evaluate response to therapy. In fact, there are instruments in lupus nephritis which are patient centric and meet fundamental measurement requirements I could find no reference to these in your report. This is a major oversight. I note in particular the L-QoL patient instrument (a separate caregiver needs instrument would also have to be developed; as well as for particular sub-groups). The L-QoL, developed some 15 years ago, is focused on combining the theoretical strengths of the need-based QoL model with the Rasch model. Content was derived from in-depth patient interviews with cognitive debriefing to assess face and content validity. Rasch analysis was applied to data from an initial postal survey to remove misfitting items with a second postal survey to assess scaling properties, reliability, internal consistency, and validity. The end result was a 25-item instrument with good item fit and stability, excellent test-retest reliability, internal consistency and strict unidimensionality. Items, scored true/not true, included "I just feel tired all the time,' "life is passing me by" and "I can't enjoy myself when I go out." The L-QoL can be reviewed on the Galen Research website (www.galen-research.com) together with other disease specific measures (http://www.galen-research.com/content/measures/L-QoL_UK_-_First_page_sample.pdf).	