Nadofaragene Firadenovec and Oportuzumab Monatox for BCG-Unresponsive, Non-Muscle Invasive Bladder Cancer: Effectiveness and Value

Response to Public Comments on Draft Evidence Report

November 6, 2020

Table of Contents

#	Comment	Response/Integration
Manufacturers		
FerGene		
1.	The rates of complete response (CR)/high-grade recurrence-free survival (HG-RFS) for nadofaragene firadenovec in ICER's cost-effectiveness analysis (CEA) is inconsistent with the clinical trial results We urge ICER to use the estimates based on the complete long-term data from Kaplan-Meier (KM) curves of durability of response (DOR) to accurately inform the proportion of patients remaining in CR/ high-grade recurrence-free survival (HG-RFS) for nadofaragene firadenovec. ICER's current approach relied on incidence of RFS data over a short term. These estimates are rough approximation, have low precision and are inappropriate for model estimation as the actual proportion of patients remaining in CR/HG-RFS at each specific time point are not accurately reflected. In contrast, the DOR estimates have high precision. It reflects the proportion of patients who remain in CR/HG-RFS precisely at each month. In addition, the DOR curves included longer-term data: up to month 27 for the carcinoma in situ (CIS) ± Ta/T1 cohort, and up to month 30 for the HG Ta/T1 cohort, versus the 12-month data from incidence rates. The additional data over 12-month with the DOR curves provides better fit for the long-term trajectory of CR/HG-RFS. ICER's current approach substantially underestimated nadofaragene firadenovec's efficacy when compared to that using the KM curves of DOR. For example, at month 27, the deviation from the observed phase 3 trial data furthered to 61% (18% by KM curves vs. 7% by Incidence estimation). ICER used inconsistent approaches to estimate the CR/HG-RFS rates for oportuzumab monatox and for nadofaragene firadenovec. For oportuzumab monatox, ICER used point estimates that matches the KM curves based on the trial observation. However, for nadofaragene firadenovec, ICER used the short-term incidence data, and as discussed above significantly underestimated the nadofaragene firadenovec's efficacy and deviates substantially from the trial observation.	We preferred to use the primary study outcomes of complete response (for patients with CIS) and high-grade recurrence-free survival (HGRFS) (for patients with HG Ta/T1). We had concerns regarding censoring occurring in clinical trials due to adverse events. The Kaplan-Meier estimates for DOR requires an assumption that censoring was not associated with treatment outcome. We believe that this assumption was not maintained and that the Kaplan-Meier estimates were biased.
2.	For consistency and to use all available data that better reflect trial observations, we suggest ICER to apply the complete available long-term DOR KM data for both nadofaragene firadenovec and oportuzumab monatox. In addition, we suggest that ICER select the generalized gamma model to extrapolate the long-term efficacy. ICER's current approach only used two incidence data points to extrapolate the long-term clinical probabilities after year 1, and arbitrarily applied the exponential model (i.e., P=1-e-kt) to extrapolate the long-term efficacy. However, using Akaike information criterion (AIC), which is the most standard statistical method to evaluate model fit for non-linear parametric models and is widely used to select best-fit models by health technology appraisal agencies2 and in prior ICER evaluations,3,4 the generalized gamma model is shown to fit the observed trial data much better than the exponential model (e.g., AIC 153.4 vs. 249.2).	Given that we did not have similar data for other treatments, we did not select the generalized gamma model to extrapolate long-term efficacy. However, we adjusted the calculations for response and progression after 12 months to better reflect long-term efficacy in a manner that could be replicated for other treatments. We believe that allowing these time-varying probabilities, and not fitting a curve to the data, better reflected the data that was made publicly available in all trials evaluated in this report.

3.	Nadofaragene firadenovec meets ICER's definition of B+ evidence rating ICER defines B+ rating as "Incremental or Better" – moderate certainty of a small or substantial net health benefit, with high certainty of at least a small net health benefit. Nadofaragene firadenovec has been designated a breakthrough therapy and assigned a fast track designation by the Food and Drug Administration (FDA).5 The criteria for breakthrough therapy designation were based on clinical evidence demonstrating that nadofaragene firadenovec treatment results in substantial improvement on at least one clinically significant endpoint over available therapy.6 The FDA guidance for drug development in Bacille Calmette-Guerin (BCG) Unresponsive NMIBC states that "Randomizing patients with BCG-unresponsive NMIBC to a placebo or minimally effective drug as a concurrent control raises ethical concerns. Currently, single-arm trials are appropriate for assessment of therapies for patients with BCG-responsive disease". The efficacy of nadofaragene firadenovec has been demonstrated in clinical trials. In a phase 3 study, accepted for publication by Lancet Oncology, nadofaragene firadenovec reported CR/HG-RFS rates of 53.4% for CIS ± Ta/T1 patients and 72.9% for HG Ta/T1 patients at 3-month, and the durability of response among patients who achieved CR/HG-RFS is 41% in CIS ± Ta/T1 and 51% in HG Ta/T1 at 18 months.1 The efficacy of nadofaragene firadenovec has exceeded the clinically meaningful thresholds suggested by the expert panel consensus that informed the FDA guidance on drug development for BCG unresponsive NMIBC patients.7,8 In a phase 2 study published at Journal of Clinical Oncology,9 nadofaragene firadenovec has demonstrated promising efficacy for patients with HG NMIBC after BCG therapy. Two phase 1 trials, published at Journal of Urology and Annals of Surgical Oncology, demonstrated that nadofaragene fivadenovec is well tolerated with promising efficacy.10,11 The totality of the above evidence and information supports the B+ rating for nadofaragene firadenovec.	We appreciate the arguments presented and have changed the rating from C+ to C++ to reflect the possibility of a substantial benefit. The lack of a placebo or active comparator control, though meeting FDA guidance, results in uncertainty about the magnitude of benefit of this new agent. The standards set by the FDA represent criteria that are based upon historical data and may not reflect current practice. In addition, the short-term outcomes reported in a relatively small number of individuals leads to uncertainty about whether the benefits observed will be seen in routine practice and whether long-term outcomes will continue to be favorable. Since most patients treated will end up having a recurrence over time, it remains to be seen whether delaying potentially curative therapy with cystectomy leads to greater long-term disease related mortality.
4.	Medical costs of health states in BCG unresponsive NMIBC are significantly underestimated We suggest that ICER use the more recent cost estimates based on SEER-Medicare data (Yang et al., 202012) for medical costs of patients in the cost-effectiveness model. ICER used medical costs for NMIBC recurrence and muscle-invasive bladder cancer (MIBC) collected in 1991-1999.13,14 Furthermore, the data were sourced from only 208 patients in a single medical center in Texas. The majority of the patients had less severe disease (only 28% with a prior history of recurrence) than HG BCG unresponsive NMIBC. The Yang et al. study, in contrast, has used more recent and more representative SEER-Medicare data (2008-2015), with medical costs reported specifically for HG NMIBC patients with adequate BCG treatments.12 Estimates from Yang et al. (i.e., $25,820 for NMIBC recurrence and $59,774 for those with progression) are substantially higher than the annual cost estimates considered by ICER (i.e., $5,832 for NMIBC recurrence and $28,108 for MIBC).	Thank you for this suggestion and for making this publication available to us for review. While we thought this was a well-conducted retrospective evaluation of patients without versus with progression, after careful consideration, we chose not to use the results of this study for inputs in the model for the following reasons: 1) Both groups (no progression and progression) included cystectomy costs; and 2) the progression group's cost estimates included those with metastatic disease. As these are separate states in our model, the inclusion of these costs for those with "Progression to MIBC" would produce highly biased estimates for this Markov state, as it would double count the costs of cystectomy and metastatic disease. The inclusion of these outcomes, along with inclusion of non-bladder cancer-related costs explain why there are differences between the costs we chose to use in the model and those reported in Yang et al.
5.	Clearly label the comparator arm in the cost-effectiveness analysis (CEA) as "hypothetical treatment" instead of "usual care" to avoid confusion and potential misinterpretation ICER used a hypothetical comparator arm in the CEA and "intentionally left this comparator undefined." However, ICER labeled this hypothetical treatment arm as "usual care" in its draft evidence report. The term could be highly misleading as this is not the "usual care" in real clinical practice. The term "hypothetical treatment" should be used instead of "usual care" to correctly characterize the comparator arm used in CEA. In addition, due to the hypothetical nature of the comparator arm in the CEA, ICER should clearly state the limitations of its CEA results in guiding real world decision making.	We agree with calling the comparator a "hypothetical treatment," since it does not well represent usual care.

6.	Present the clinical effectiveness evidence and cost-effectiveness results by comparable study design and patient population to avoid potential "apple to orange" comparison. In the draft evidence report, ICER acknowledged "differences in patient populations and study design make any direct comparisons exceedingly difficult." ICER also recognized that heterogeneity in patient characteristics could lead to differences in expected treatment outcomes. For example, ICER mentioned that "failure types such as BCG-relapsing are associated with better outcomes compared with other reasons for BCG failure." ICER also acknowledged that prior treatments and their intensity could "lead to differences among studies in terms of patients and how resistant to subsequent treatment their NMIBC is likely to be." However, in the draft evidence report, ICER summarized the efficacy results from various treatments in one table (Table 4.16) without clear separation by study design nor by patient population. Retrospective observational study with much less severe disease than BCG unresponsive are grouped together with clinical trials with BCG unresponsive NMIBC patients. Similarly, for cost-effectiveness assessment, the results are summarized in tables (Tables 5.10-5.13) for various treatments with very different study design and patient population. To avoid misinterpretation, a modified table format that clearly states the differences in study design and patient population is needed. [SEE COMMENT FOR SUGGESTED TABLE] In addition, statements should be added to the tables to clearly state that the differences in patient characteristics and study design could significantly affect study outcomes independent of the treatments and any comparison of efficacy across different study designs and/or patient population is not warranted. Below are suggested mock table templates (Tables 1-3) for ICER's considerations:	To make our data presentation clearer, we have revised how we present results from the different studies. In the evidence section of the report (section 4), we have removed table 4.16 and instead present the data in graphical form for nadofaragene, oportuzumab, and pembrolizumab separately. This is also used in the executive summary section of the evidence report. In the economic modeling section, we separated results by method of drug delivery (instilled vs. systemic) and study types used to estimate the primary effect of drugs (prospective vs. retrospective). This was done for tables 5.10-5.13 and corresponding tables in the executive summary.
7.	Include full and complete adverse event (AE) and associated costs in CEA ICER arbitrarily included only three common non-grade 3-5 AEs (i.e., urinary tract infection, rash, and pruritus) in the CEA for oportuzumab monatox and pembrolizumab. However, as presented in the clinical effectiveness section Table 4.6 (P22) and Table 4.9 (P26) in ICER's draft evidence report, 21% patients treated with oportuzumab monatox and 29% patients treated with pembrolizumab experienced grade 3-5 AEs. ICER's current approach to model AE and the associated costs substantially underestimated the cost of treating grade 3-5 AEs and could be highly misleading on the safety of the treatments without including the full and complete serious AEs and associated costs. Corrections are needed in ICER's revised CEA model to fully account for these AEs and associated costs. Some inconsistent numbers are noticed in ICER's draft evidence report as well. For example, the type of AEs and their proportions used in the CEA are inconsistent with the numbers reported in the clinical effectiveness section of the draft evidence report for oportuzumab monatox, and the US Prescribing Information (USPI) for pembrolizumab. For oportuzumab monatox, the clinical effectiveness section reported 32% of patients have urinary tract infection (P22), while the CEA considered 12% of patients with this event. Many common AEs that were highlighted in the clinical effectiveness section or the USPI for pembrolizumab were not considered in the CEA, including: fatigue (29%), diarrhea (24%), hematuria (19%), cough (19%), arthralgia (14%), nausea (13%), constipation (12%), peripheral edema (11%), hypothyroidism (11%), and nasopharyngitis (10%) for pembrolizumab; and pain or burning on urination (26%) and hematuria (25%) for oportuzumab monatox.	The clinical section and economic section had different criteria and reasons for reporting adverse events. Importantly, the economic report only included those adverse events that were grade 3-5 and were deemed to be treatment-related, because those adverse events were likely to be treated and incur additional costs. The one exception to this a priori decision was urinary tract infection (UTI), in which all treatment-related adverse events were included regardless of severity. All treatment-related UTIs were included because all patients with UTIs would likely receive treatment. The rationale behind the a priori selection of adverse events to include in the model and decision criteria for doing so has been better elucidated in the revised report.

8.	Comments on the draft voting questions Before the voting questions, clearly defined best supportive care is needed for both the voting panel and the public to make informed decisions. Additionally, clear evidence summaries on the efficacy, safety, tolerability, patient adherence/discontinuation, and frequency of administration for nadofaragene firadenovec, oportuzumab monatox, pembrolizumab, gemcitabine with or without docetaxel, and best supportive care, separated by study design and patient population, are needed before the voting questions. Strength/source of the evidence needs to be provided, e.g. peer-reviewed journal publication, congress presentations, investor report/social media postings, number of patients included in the study. For draft voting questions 1 - 7, substitute the "net health benefit" with efficacy, safety, tolerability, patient discontinuation, and frequency of administration, respectively, to better inform the various aspects of the differences in treatments. e.g. expand question 1 into 1a – 1e as follows: 1a. Is the evidence adequate to demonstrate that the efficacy of nadofaragene firadenovec (Adstiladrin®, FerGene) is superior to that provided by best supportive care? 1b. Is the evidence adequate to demonstrate that the safety of nadofaragene firadenovec (Adstiladrin®, FerGene) is superior to that provided by best supportive care? 1c. Is the evidence adequate to demonstrate that the tolerability of nadofaragene firadenovec (Adstiladrin®, FerGene) is superior to that provided by best supportive care? 1d. Is the evidence adequate to demonstrate that the patient discontinuation of nadofaragene firadenovec (Adstiladrin®, FerGene) is superior to that provided by best supportive care? 1e. Is the evidence adequate to demonstrate that the frequency of administration of nadofaragene firadenovec (Adstiladrin®, FerGene) is superior to that provided by best supportive care? Before the Potential Other Benefits and Contextual Considerations section of the voting questions, provide the following summary table to better inform the voting panel and the public: • Unmet need in HG BCG unresponsive NMIBC • Cost-effectiveness threshold as reference points for cost-effectiveness determinations • The levels of absolute quality-adjusted life-year (QALY) measure or proportional QALY shortfall that would be considered small/medium/large health loss.	We will emphasize to the CEPAC voting panel and the public that there is no established comparator for patients with BCG-unresponsive NMIBC. This is the basis for the FDA permitting trials without a comparator. However, in the cost effectiveness modeling section, we emphasize that a hypothetical treatment is the comparator. This reflects the fact that patients may decline or be ineligible for cystectomy, the doctor will then select the best available treatment if the patient cannot be enrolled in a trial. ICER's Evidence Report and presentation are available to the CEPAC – these are not repeated in the voting questions. ICER asks the CEPAC voting members to focus on the net health benefit that combines these separate items, similar to what doctors and patients do as part of routine clinical practice.
9.	Other suggestions Table 4.2 in the draft report: Nadofaragene firadenovec has reported the 12-month HG-RFS without mandatory biopsy as follows: CIS ± Ta/T1: 28 (27.2%); HG Ta/T1: 23 (47.9%). These numbers should be included for consistency as the numbers reported for other new treatments are measured without mandatory biopsy.	In the evidence section of the report, we have added information to table 4.2 that highlights patients with 12-month recurrence based upon the mandatory biopsy alone.
10.	P18: Update the Progression to MIBC section to: "8 (5.3%) of 151 patients in the overall study population progressed to muscle-invasive bladder cancer (MIBC) during the full available follow-up (median of 23.62 months)".	We have revised this sentence to clarify the follow-up period.
11.	Table 4.10: Revise the proportion of patients with low-grade Ta/T1 only from Skinner et al. (2013) publication to 11% (5/47).	Thank you. We have corrected this proportion.

Merck		
1.	Calculate the transition probabilities from NMIBC to MIBC for pembrolizumab based on treatment-specific progression free survival (PFS) Recommendation: We strongly recommend revising the transition probability from NMIBC to MIBC for pembrolizumab from 2.40% to 1.36%, estimated based on the published 12-month pembrolizumab-specific PFS of 96.9% [1]. The input calculation for pembrolizumab is detailed in Appendix Table 1. Concerns and Rationales: PFS is an important factor in determining the incremental cost effectiveness of treatments. We have significant concern that the current analysis was based on incorrect transition probabilities from NMIBC to MIBC for pembrolizumab, resulting in inaccurate clinical effectiveness and cost-effectiveness results .The average of the transition probabilities based on the PFS for the two intervention drugs was used as a proxy to populate the transition probability for pembrolizumab in the draft report, assuming that PFS data for pembrolizumab was not available. In fact, the 12-month PFS for pembrolizumab was reported as 96.9% [1] and thus should be used to populate this transition probability. This point was raised in our response to ICER's model development plan on August 21, 2020 (refer to section 1.8). Using the transition probability (1.36%) derived from the published 12-month PFS for pembrolizumab is more appropriate than what was used in the draft report (2.4%), because the former approach leads to a predicted PFS curve more aligned with the observed PFS curve from the clinical trial KN057 than that of the latter approach (Appendix Figure 1). Pembrolizumab should have the lowest transition probability (1.36%), compared with nadofaragene firadenovec (2.2%), and oportuzumab monatox (2.6%), because pembrolizumab had the highest 12-month PFS (96.9%) vs. nadofaragene firadenovec (95.1%) and oportuzumab monatox (94%). We agree with this comment and recommendation. We have adjusted the transition probabilities for all treatments to reflect the actual values for each drug.	We agree with this comment and recommendation. We have adjusted the transition probabilities for all treatments to reflect the actual values for each drug.
2.	Calculate the transition probabilities from Complete Response (CR) to NMIBC for pembrolizumab based on duration of response (DOR) Recommendation: We strongly recommend using median DOR to derive the time-constant transition probability for pembrolizumab. For consistency, this approach should also be applied to the two interventions and other comparators when median DORs are available from the respective trials. The recommended inputs for pembrolizumab (base case and 2 alternative scenarios) are presented in Appendix Table 2. Note that this approach could still be conservative in estimating pembrolizumab's long-term effectiveness, given the possible durable treatment effect implied by the flattened tail observed in the DOR curve beyond 12 months from KEYNOTE-057 (Appendix Figure 2).	We preferred to use the primary study outcomes of complete response (for patients with CIS) and high-grade recurrence-free survival (HGRFS) (for patients with HG Ta/T1). We had concerns regarding censoring occurring in clinical trials due to adverse events. The Kaplan-Meier estimates for DOR requires an assumption that censoring was not associated with treatment outcome. We believe that this assumption was not maintained and that the Kaplan-Meier estimates were biased.
3.	Concerns & Rationales: First, pessimistic assumption was made in the draft report when interpreting the complete response data. When the number of patients in CR reduces over time, as shown by "number at risk", it can be due to either an event (loss of CR) or a censor (e.g., reach the end of study cutoff, start new treatments, or have non-evaluable assessments), as illustrated in Appendix Figure 2 and Figure 3 [2]. The current approach pessimistically assumes that all censored patients experienced recurrence, which overestimates the transition probabilities from CR to NMIBC. With this approach, the median DOR that the model predicted for pembrolizumab (12 months, as shown in Appendix Table 3) is much shorter than what was reported from the KEYNOTE-057 (16.2 months), indicating that the current model lacks internal validity. We recommend using KM estimates, as illustrated in Appendix Table 2, as KM estimation is a typical approach to deal with censoring.	We agree that including censored patients as having had the outcome (i.e. loss of CR or loss of HGRFS) could bias the results against treatments and produce pessimistic results. However, censoring before 12 months occurred in clinical trials primarily due to adverse events. The Kaplan-Meier estimate for DOR requires an assumption that censoring was not associated with treatment outcome. We believe that this assumption was not maintained and that the Kaplan-Meier estimates were biased. In dealing with censored patients, we had to choose between a potentially overly optimistic estimate produced by Kaplan Meier estimates and a potentially pessimistic estimate assuming censored patients had the worst outcome after censoring. Given that we assigned the hypothetical comparator a 0% CR probability at 3 months, we believed that assigning a potentially pessimistic outcome to censored patients would still produce an incremental cost-effectiveness ratio that favored treatments.

4.	Second, inconsistent approaches were used for populating the transition probabilities from CR to NMIBC for the CIS population in the draft report. Specifically, percentages of patients in CR over time were used for pembrolizumab, whereas high-grade recurrence free survival (HGRFS) probabilities were used for other regimens. According to the FDA guidance, CR and DOR are the recommended primary efficacy endpoints for patients with high-risk NMIBC with CIS since these patients have active disease at baseline, whereas recurrence-free survival is recommended for patients without CIS (as disease was resected before trial entry)	We agreed with this comment and used CR, when available, as our outcome for patients with CIS for the evidence report.
5.	We have major concerns that using inconsistent approaches for calculating the transition probabilities have led to model predictions that are contradictory with the trial efficacy results. Specifically, the model predicts lower LY and QALY for pembrolizumab (6.22, 4.74, respectively) compared to those for oportuzumab monatox (6.28, 4.80), which was inconsistent with the clinical trial results that pembrolizumab had a slightly higher CR at 3 months (40.6% vs. 40%), and much longer median DOR (16.2 vs. 9.6 months) than oportuzumab monatox. Appendix Table 4 demonstrates that different transition probabilities (from CR to NMIBC) were derived for the same treatment (i.e., oportuzumab monatox), when calculated using difference approaches (i.e., CR and HGRFS, respectively). To deal with the above-mentioned censoring, endpoint and consistency issues, we strongly recommend using CR and DOR KM estimates (specifically median DOR) to derive the transition probabilities from CR to NMIBC for pembrolizumab and other interventions.	The changes made to assessing transition probabilities for recurrence and progression after 12 months, along with using CR as the appropriate measure for assessing recurrence in patients with CIS have corrected many of the apparent inconsistencies.
6.	Remove the cost-effectiveness analysis of gemcitabine + docetaxel in CIS population Recommendation: We propose to remove the cost-effectiveness analysis of gemcitabine + docetaxel for the CIS population until robust data become available. Concerns and Rationales: Gemcitabine + docetaxel are not appropriate comparators for the CIS population. These regimens are not recommended in the clinical guidelines for this population due to a lack of rigorously conducted clinical trials in this setting.	We included studies of gemcitabine with or without docetaxel based upon expert input. We recognize that the quality of the studies for these drugs is below that for the newer drugs reviewed. Because of this we did not perform indirect comparisons and we do not compare the cost-effectiveness of the various drugs evaluated. Including gemcitabine with or without docetaxel in these analyses highlights their potential efficacy and cost-effectiveness. As such, future studies should consider whether these drugs produce similar results to the newer instilled agents. Furthermore, the experts viewed intravesical chemotherapy as an option for patients who declined or were ineligible for cystectomy and for whom a clinical trial was not available. We believe that this also reflects clinical guidelines. Specifically, we are not aware of guidelines that specifically recommend against using these regimens. Rather, their potential role remains uncertain due to limitations in the available literature. Based upon this input, we felt that these regimens may be considered for certain patients with NIMBC who were unresponsive to BCG.
7.	The studies identified via the literature review were all retrospective in nature; some included a heterogenous population of patients (i.e., a mix of CIS and non-CIS patients) with varying risks of recurrence and progression. These limitations make it impossible to draw robust conclusion on the comparative efficacy of gemcitabine + docetaxel versus usual care for the CIS population, and thus invalidates the cost-effectiveness analysis.	We recognize and highlight the limitations of studies of gemcitabine with or without docetaxel. We also highlight the limitations of the studies of nadofaragene, oportuzumab and pembrolizumab that while prospective are not randomized or have a usual care comparator. Because of the limitations of the available clinical trials, we have not attempted to perform indirect network metanalysis for any of primary drugs or comparators. We have attempted to highlight trials where there are similarities in populations studies (e.g. the percent of patients with CIS +/-Ta/T1 or Ta/T1 alone) and with similar outcomes and follow-up periods. We believe that our analyses reflect the guidance of our experts and that we have not overstated our findings.

8.	The cost-effectiveness analysis for the CIS population relied heavily on one retrospective study [4] to inform efficacy inputs for gemcitabine + docetaxel. Key issues included: It is inappropriate to use the 3-month HGRFS as a proxy for the CR rate, as explained above under point 2. However, CR rates and DOR were not reported from the study, making it impossible to populate the model in a consistent way as other regimens. The adjustment factor (rate ratio) used in the draft report lacks clinical justification. The rate ratio calculation was arbitrary and can vary by the selected time points. In addition, the adjustment factor was derived from an overall population, and thus not applicable to the CIS sub-population. Efficacy inputs were solely based on studies for gemcitabine plus docetaxel, and therefore should not be used to represent the efficacy for gemcitabine without docetaxel.	We appreciate the concerns raised in this comment. Given the large number of limitations identified in studies of nadofaragene, oportuzumab, pembrolizumab, and gemcitabine with or without docetaxel, it was not possible to compare any of these drugs to each other. As such, when interpreting the calculated incremental cost-effectiveness ratios for each of these drugs, it is important to also consider the limitations of the studies and data that were included in the analyses.
9.	The significant limitation of the data and the use of inappropriate endpoint have led to clinically implausible predictions of the model in the draft report. Specifically, the model predicted that gemcitabine + docetaxel has a median DOR of 4 years, and that patients on average would stay in CR for 5 years during an average of 11 life years (Appendix Table 3). These model results are not aligned with clinical insights and other published data, which suggested much lower efficacy for gemcitabine + docetaxel [5]. Two additional impactful calculation errors are described in Appendix Table 5.	We have corrected the identified error in the time to calculate the transition probability from CR to NMIBC after 12 months. We reviewed the RR adjustment factor. Our estimate for Steinberg 2015 differs from that suggested in the supplied Appendix Table 5 of the response letter. We did not make changes to the adjustment factor.
10.	Revise key model inputs (i.e., drug cost and transition probability from CR to NMIBC at 6 months) for oportuzumab monatox Recommendation: The following model inputs for oportuzumab monatox should be revised. The frequency of drug administration was inconsistently reported in different sections in the draft report. The correct dosing schedule is every other week for maintenance.	Thank you for identifying this error in the oportuzumab dosing schedule. We have corrected this input in the model.
11.	The drug cost for oportuzumab monatox should be $4,317 per dose (instead of $2,826 per dose in the draft report), calculated based on the total number of doses of 38 per year (instead of 58 per year in the draft report).	Thank you for identifying this error in the oportuzumab cost estimate. We have corrected this input in the model.
12.	The transition probability from CR to NMIBC at 6 months for oportuzumab monatox should be 23.8% (instead of 20% used in the draft report). It should be calculated as (1-32%/42%)*100% = 23.8%.	Thank you for identifying this error in the model. Changes made to using CR as the primary outcome for patients with CIS made this input obsolete.
13.	Comments on comparison between pembrolizumab and treatments other than usual care. Recommendation: We propose to remove the sentence (page 69) on the draft report, that for pembrolizumab, 'the QALY gains appeared to be smaller than those seen with any of the other treatments'. Rationales: This statement implies to compare the clinical effectiveness of pembrolizumab with nadofaragene firadenovec, oportuzumab monatox, and gemcitabine + docetaxel. This contradicts the conclusion from the draft report around large uncertainties in comparative benefits and harms among pembrolizumab and other therapies (see Section 4.4, Comparative Clinical Effectiveness). Thus, it is premature to compare and draw any conclusions on QALY comparison between pembrolizumab and other therapies.	We agree with this comment. This language has been removed from the report.

Sesen Bio		
1.	The gemcitabine plus docetaxel data presented in the report was generated from retrospective studies and extreme caution should be taken regarding the outcomes. All the studies evaluating gemcitabine have been performed prior to the 2018 FDA definition of BCG-unresponsive patients hence the data is not comparable to that of the Oportuzumab monatox Phase 3 trial. The gemcitabine/docetaxel combination studies published by Steinberg et al., 2015 and Milbar et al., 2017 contained heterogenous patient populations comprised of BCG-unresponsive, BCG naïve or BCG intolerant patients and low-grade patients. The patient population in the Daniels et al., 2020 paper is defined as BCG-failure and not characterized as BCG-unresponsive. Therefore, the data from these three studies was generated with an easier patient population to treat compared to Nadofaragene Firadenovec, Pembrolizumab and Oportuzumab monatox. We believe that the data cannot be used for direct comparison with Oportuzumab monatox, Nadofaragene Firadenovec and Pembrolizumab.	We have included evidence for gemcitabine with or without docetaxel based on input from our clinical experts. We highlight the details of the studies and the differences with the studies of oportuzumab and nadofaragene. As noted here, we have not directly compared the outcomes of gemcitabine with or without docetaxel with either oportuzumab, nadofaragene or pembrolizumab.
2.	The only suitable comparator data is from the paper published by Steinberg et al., 2020 in which a cohort of 71 CIS and 34 PAP patients are defined as BCG-unresponsive. However, despite intriguing data, the following citations from the paper "limitations include the retrospective nature and moderate follow-up" and "might be influenced by selection bias given that physician discretion was utilized to determine those who received treatment" clearly indicate that the data should be interpreted with extreme caution. Moreover, the comparison with Oportuzumab monatox data is difficult since the number of prior BCG cycles and the proportion of BCG-unresponsive CIS patients with Ta or T1 papillary disease are not indicated. As mentioned in the ICER report, the difference in the number of prior BCG cycles and proportion of patients with CIS + T1 disease does not allow comparison between trials (page 36). All together, given the retrospective nature of the studies, the heterogeneity of the patient population and the presence of low-grade patients, the gemcitabine docetaxel data presented in the report may not represent the outcomes of a clinical trial enrolling only BCG-unresponsive patients. An appropriate clinical trial should be performed to assess the efficacy and safety of gemcitabine/docetaxel combination as per 2018 FDA guidance. Of note, a meta-analysis performed by Merck showed that the historical rate at 3 months for BCG-unresponsive CIS patients treated with a single chemotherapeutic agent was 21% (CI 95%: 15, 27%) (5). Sesen Bio believes that the data from single chemotherapeutic agents should be used for comparison since Oportuzumab monatox was used as a monotherapy.	We have included evidence for gemcitabine with or without docetaxel based on input from our clinical experts. We performed a literature review to identify studies that met our screening criteria and highlight their methods and results. We do not see why including studies of gemcitabine with docetaxel both given intravesically would be problematic.
3.	Sesen Bio agrees with ICER that the evidence rating is insufficient to compare gemcitabine/docetaxel with Oportuzumab monatox. For this reason, Sesen Bio would like to ask ICER to remove question 5 from the voting list.	The CEPAC panel that will review and rate the evidence presented will vote on this question as is ICER's standard practice.
4.	Pembrolizumab Phase 2 trial enrolled 62.5% of OUS patients. The report should mention that out of 96 patients evaluated after Pembrolizumab treatment, 62.5% (60 of 96) of patients were enrolled outside of the US and only 37.5% (36 of 96) in the US. The CR rate for the OUS cohort was 47% vs. 30.6% for the US cohort (page 24 of the briefing book) (5). The report should also specify that the median duration of response was 16.2 months for all patients, however there is no data specifically for the US cohort.	We include information about studies recruiting from inside and outside the U.S. However, we present data from all patients enrolled since the investigators followed the same protocol, treatment algorithm and outcome assessments.
5.	Sesen Bio agrees with ICER that the evidence rating is insufficient to compare Pembrolizumab with Oportuzumab monatox. But more importantly, since most of the data for the US cohort is unknown, Sesen Bio would like to suggest that ICER removes question 7 from the voting list.	The CEPAC panel that will review and rate the evidence presented will vote on this question as is ICER's standard practice.

6.	The lifetime total cost of usual care was estimated at $190,000 by ICER. However, there is no reference provided for this figure and corresponds to a hypothetical usual care. The reported cost-effectives per QALY are premature. Without a reference, Sesen Bio cannot comment on the lifetime total cost of usual care estimated at approximately $190,000 for CIS and papillary patients by ICER. This number is lower than the cumulative cost of care over a 5-year period of $366,143 for high-risk NMIBC published by Mossanen et al., 2019. Furthermore, the article reported that the primary driver of cost was progression to MICB contributing to 92% of the overall cost for high-risk disease. Therefore, Sesen Bio believes that the long-term data of the trial will increase the cost-effectiveness of Oportuzumab monatox. As the data matures, Sesen Bio is convinced that Oportuzumab monatox will be recognized as a viable alternative to cystectomy by urologists as a cost-effective alternative by payers. Based on multiple rounds of market research, payers view Oportuzumab monatox as cost effective, specifically due to outcome data such as time to cystectomy, overall survival and progression-free survival, as well as the favorable safety profile.	This estimate was directly calculated from the model. Therefore, a reference was not needed. We used one-year costs for recurrence and progression from Mossanen et al as inputs in our model. These resulted in costs that were well aligned with previously published estimates from other models (see the section of the report on Model Validation in Section 5 of the draft report). Upon thorough review of Mossanen et al, it is not completely clear why five-year total costs are so different in this model from other previously published models, but appear to be due to differences in the cost of cystectomy and progression to MIBC and/or metastatic disease in this model. Unfortunately, this paper does not separate out costs included in the MIBC state that were attributable to cystectomy and metastatic disease for comparison with our model inputs, which were derived from Leow (2014) and Malangone-Monaco (2020). It does appear from the model figure and description that the costs associated with cystectomy could have been applied to multiple cycles for patients with MIBC in Mossanen et al, thereby double-counting cystectomies and falsely elevating the cost of being in the state "Progression to MIBC and Cystectomy" after five years.
7.	As outlined in the FDA guidance, avoiding cystectomy is a key secondary endpoint for NMIBC therapies and Sesen Bio believes that the report should further discuss this point. The 2018 FDA guidelines indicates that “the goal of therapy in patients with BCG-unresponsive NMIBC is to avoid cystectomy” (7). Radical cystectomy not only has a tremendous impact on the quality of life for a patient, including catheterization and urinary diversion, but it also comes with significant costs to the healthcare system. From the data in our Phase 3 trial, 76% of patients treated with Oportuzumab monatox are estimated to remain cystectomy-free for 3 years. Additionally, responders have a statistically significantly higher probability of remaining cystectomy-free at 2 years than non-responders (88% vs. 61%), which could change the lives of patients and provide significant savings for the healthcare system. The Pembrolizumab Phase 2 study only enrolled patients who were ineligible or refused to have a cystectomy. As a consequence, cystectomy data was not included as a secondary endpoint. Therefore, Pembrolizumab is only approved for “the treatment of patients with Bacillus Calmette-Guerin (BCG)-unresponsive, high-risk, non-muscle invasive bladder cancer (NMIBC) with carcinoma in situ (CIS) with or without papillary tumors who are ineligible for or have elected not to undergo cystectomy”. Sesen Bio believes that any cystectomy data from the Phase 2 trial should be taken with caution since a selection bias may have been introduced by enrolling a population in which 95% of patients refused to have a cystectomy (ODAC, FDA presentation, slide 19) (8). As BCG-unresponsive patients are facing a difficult decision with lifetime implications on quality of life, Sesen Bio believes that the cystectomy data obtained with Oportuzumab monatox will be a differentiating factor from other intravesical or systemic therapies.	We agree that new therapies that provide an alternative to cystectomy will be welcomed by patients and their providers. The challenge is that it remains unclear whether therapies such as oportuzumab and nadofaragene delay the need for cystectomy without increasing the risk for more advanced bladder cancers that no longer may be cured by cystectomy. Only longer-term outcomes data will be able to answer that question. With regards to comparing outcomes across trials, we made clear that differences in study design as well as the lack of a placebo or usual care comparator limit our ability to perform direct or indirect network meta-analysis.
8.	Minor comments: Reporting of the adverse events (AEs): The data reported for Nadofaragene firadenovec are treatment-related adverse events (TRAEs) (Table 4.3, page 19); therefore, table D5 should be edited accordingly. However, all treatment-emergent adverse events regardless of causality are reported for Oportuzumab monatox (Table 4.6, page 22). For consistency and fair comparison, Sesen Bio recommends that ICER reports either all AEs or TRAEs for both products.	We have revised the presentation of the adverse events to clarify differences in how they are reported among the different trials.

9.	Progression to MIBC for the gemcitabine/docetaxel combination study (Steinberg et al. 2020): The report should clarify that it was 4% of patients from the entire patient population that progressed to MIBC i.e. 276 patients (11/276 = 4%) (page 33). This number is misleading and does not exclusively represent the percentage of the BCG-unresponsive patients that progressed to MIBC.	We have revised the text to make the denominator population clearer.
10.	Table 5.8, page 52: Using $164,337, the net price per dose provided for Oportuzumab monatox in the table implies 55 to 60 doses per year which is not correct. Patients will receive up to 36 doses in the first year (12 doses from Week 1-6, 6 doses from Week 6-12 and 18 doses from Week 14-52) and up to 24 doses in the second year.	See responses above.
11.	Figure 5.3, page 60: Replace Nadofaragene firadenovec with Oportuzumab monatox in some of the probability listings.	These figures and labeling have been updated.
#	**Comment**	**Response/Integration**
Patient/Patient Groups		
Cancer Support Community		
1.	As we have noted in previous letters, we believe this value assessment is premature, particularly as pricing is not yet available for nadofaragene firadenovec and oportuzumab monatox. ICER notes that this makes it "difficult to determine whether treatment for BCG-unresponsive high-risk NMIBC will be considered cost-effective." ICER elected to substitute the annual price of pembrolizumab and noted that "as a result, determining an appropriate and fair health-benefit based price for this heterogenous group of patients will be difficult, made even more so by not having evidence on potential comparators."	We recognize that for newly approved treatments there are often limited data available. However, patients, clinicians and insurers are still faced with decisions about how best to use these new agents once approved for use. As such, we view comparative clinical effectiveness research, and cost-effectiveness modeling as a useful and important way to identify the key inputs that impact the effectiveness and cost of a new therapy. Our reports use data that are currently available and highlight the limitations of these data as well.
2.	ICER recognizes the "profound impact of BCG-unresponsive NMIBC on quality of life" and the "large burden" placed on patients dealing with this disease. The disease is a chronic condition for many resulting in significant quality of life, logistical, psychosocial, and financial burdens for patients. We appreciate ICER's recognition of these burdens on both patients and caregivers. ICER states that guideline-concordant care includes radical cystectomy as the "gold standard treatment" yet it is often declined (due to quality of life issues) or unfeasible (due to comorbidities). While "few patients progressed to metastatic disease or died during the short follow-up period…it is possible that these treatments may lead to more patients avoiding potentially curative cystectomy and therefore progressing to metastatic disease or dying of bladder cancer." We recognize the need for longer term follow-up with patients treated with nadofaragene firadenovec and oportuzumab monatox, however we also wish to re-emphasize the critical quality of life components inherent for patients whose only treatment option is cystectomy. While disease recurrence or progression over time is a possibility, the trade-off for patients who wish to avoid the significant health, quality of life, psychosocial, logistical, and financial issues that can accompany cystectomy must be seriously considered. As a result of this difficult decision for patients, overall survival may not be the endpoint of most concern for them and must be weighed alongside all of the issues that may accompany cystectomy.	We agree with this comment and sought to highlight this difficult tradeoff for patients contemplating next steps with NMIBC who are unresponsive to BCG. Avoiding cystectomy for the reasons cited may lead patients to select bladder sparing treatments even if these agents may not work or may only delay progression. The fact that most patients will eventually have recurrent disease over time highlights the need for new and more effective therapies.
3.	An additional item of note is ICER's reference to the dosing schedule of nadofaragene firadenovec and that a less frequent schedule is "an advantage during the COVID-19 pandemic where minimizing office visits is desirable." ICER goes on to say that "it is also likely that decreased frequency of dosing will decrease the burden of treatment and travel-related costs for patients, as well as family and caregivers." We would like to emphasize that less onerous dosing schedules are likely desired by many patients and caregivers, regardless of the pandemic. While the risk of contracting COVID-19 in a clinical setting is certainly a warranted concern, it is important to recognize the impact of dosing frequency and setting when considering value.	We agree with this comment. Our statement about this being an advantage during the COVID-19 pandemic was meant to highlight how this issue is even more important now. It was not intended to imply that it isn't an issue when the pandemic is over. We have revised our statement to clarify this point.

4.	Finally, we are resubmitting our open input comments on bladder cancer treatment as well as our Cancer Experience Registry findings to help inform the voting panel's deliberations on this review.	Thank you.
Partnership to Improve Patient Care		
1.	ICER continues to conduct studies prematurely PIPC echoes the Cancer Support Community and other stakeholders in the belief that this report is being undertaken prematurely. ICER has chosen again, in the absence of sufficient evidence, to prematurely assess the value-based price of these drugs. No respected health technology assessment agency anywhere in the world evaluates new drugs before phase III data is available and the relevant drug regulation agency has approved its use. Despite this, ICER has made it common practice to prematurely assess the cost-effectiveness of drugs. Without a drug being approved and a price established, it is irresponsible to evaluate its cost-effectiveness.	We recognize that for newly approved treatments there is often limited data available. However, patients, clinicians and insurers are still faced with decisions about how best to use these new agents once approved for use. As such, we view comparative clinical effectiveness research, and cost-effectiveness modeling as a useful and important way to identify the key inputs that impact the effectiveness and cost of a new therapy. Our reports use data that are currently available and highlight the limitations of these data as well.
2.	PIPC has concerns about the sources and construction of ICER's health state utility inputs The health state utility values for the model seem to be taken from a single study undertaken in the UK where quality of life data was collected as part of the BOXIT trial. The approach taken in this study was to estimate utility loss increments, not to actually estimate utility values of certain health states. This method is a valuable way to capture variance in disease states and comorbidities, but it must be approached correctly. The problem with ICER's use of these utility values is that these incremental utilities have been applied individually to create proxy health states for the ICER model. In reality, many of these utility loss increments will be relevant to most, if not all, patients, so the use of individual utility loss increments – rather than combinations of utility loss increments is likely to significantly overestimate the health utility levels of people in more severe states of disease. For example, in the ICER model, patients with inoperable advanced metastatic bladder cancer seem to have an HSUV of 0.7. This is a magnitude of quality life higher than people with arthritis, dermatitis or migraine. It is highly unlikely that this an accurate summation of quality of life for people suffering the late stages of incurable cancer, and further demonstrates the flawed logic of a QALY-based model. The result of this overestimation is that the value of reducing time spent in these health states – the stated goal of most new treatments for any disease – will be undervalued.	This interpretation of the purpose and use of utilities from the BOXIT study is incorrect. Table 2 of the report by Cox et al clearly shows the Estimated Health State Value estimates from the study. These are the utility inputs that were used in the model.
3.	Mixed data sources for measures of effectiveness are likely to lead to biased estimates in the ICER model. ICER chooses to compare retrospective data to randomized clinical trial data in order to compare effectiveness across drugs. Whenever possible, ICER should compare equivalent data sets for consistency. The review of the phase II and III trials shows a complete response (CR) for gemcitabine ± docetaxel of no greater than 39%, and HGFRS at 12 months ranging from 21-28% in populations with a high proportion of CIS ±HIG Ta/T1. Yet the ICER model uses a much higher figure that comes from a retrospective chart review of selected patients of 60-69%, and a figure of 75.2% for complete response. ICER acknowledges that these response rates are peculiarly high yet still chooses to use this data instead of comparable source data from trials. Retrospective data is incredibly valuable when used correctly, but the issue here is that there are not equivalent data sets for new drugs or therapeutic approaches. There is strong empirical evidence that the relative effectiveness of new therapies tend to improve over time, as physicians and providers develop better understanding of when, to whom and how to incorporate therapies into everyday treatment plans. This learning-by-doing leads to a rise in effectiveness, as has been shown to exist in oncology for multiple tumors. Comparing efficacy rates from a phase II or III trial with a retrospective case review is not a reasonable comparison.	We agree that differences among the trials, both prospective and retrospective, including the lack of placebo or active comparators, do not allow us to compare firadenovec nadofaragene directly or indirectly and oportuzumab monatox to each other or to the comparators. However, based upon the input of our experts we felt it was important to examine the cost-effectiveness of the different drugs using a hypothetical treatment comparator. In doing so, we selected data from trials of gemcitabine with or without docetaxel that we felt were most appropriate. We emphasize the limitations of this data but also recognize that the effectiveness and costs of the chemotherapeutic agents make them worthy of consideration.

4.	ICER uses the discriminatory Quality-Adjusted Life Year (QALY) As PIPC has voiced many times in the past, we are concerned with ICER's continued use of the Quality-Adjusted Life Year (QALY). The QALY is known to discriminate by devaluing treatments designed for individuals with disabilities and chronic illnesses. In a 2019 report, the National Council on Disability, an independent federal agency, found that use of the QALY is contrary to United States civil rights laws and due to its implications for disability discrimination. The report specifically focuses on the United Kingdom's use of the QALY, highlighting cancer patients' lack of access to novel treatments and worse outcomes. PIPC encourages ICER to abandon the use of the QALY for this assessment and all those moving forward.	We appreciate the concerns about relying solely on QALYs. They are not used in the assessment of the comparative net health benefit and they are also only one component of the value assessment. The QALY has served as a fundamental component of cost-effectiveness analyses in the US and around the world for more than 30 years. 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.
Patients Rising Now		
1.	People-Centered Perspectives As is well known, people with cancer face a variety of health care and life concerns; being diagnosed with cancer can be a very distressing and jarring event. The draft report describes this at the onset: "Bladder cancer can have a large effect on patients' lives, particularly if the cancer does not respond adequately to standard therapy." While the patient perspectives discussed in Section 2 are useful, they seem to be only derived from "two patient advocacy groups and a patient treated for bladder cancer." ICER should have engaged with as broad an array of patients as possible. Although the patient advocacy groups may have provided access to more patient insights, the draft report does not include that level of specificity about the input ICER received.	As part of our scoping process, ICER engages patient advocacy groups, clinical experts, the manufacturers and payors to better understand their perspectives and concerns. Our process includes multiple formal and informal opportunities for patients and caregivers to engage with our review, and actively reaches out to disease specific patient organizations for input in each review from beginning to end of the process. Nevertheless, we welcome specific suggestions as to how to improve our current process.
2.	Diagnostic and Treatment Complexities and Opportunities Cancer is widely recognized to be a category of disease rather than a single disease. Different cancers present very different concerns and challenges for patients. For example, glioblastoma is very hard to treat and most people live only a few years; squamous cell skin cancer is very common and easily treated or cured (if it hasn't spread too widely); and prostate cancers have varying degrees of severity and aggressiveness. We point this out since not only do different cancers represent different clinical outlooks and life choices for patients, but as biomedical science has advanced, it is clear that even a single "type" of cancer is really an amalgam of many different subtypes – often characterized by specific genomic and biomarkers or mutations. Perhaps the best example of that variation is breast cancer. In contrast, the treatment of bladder cancer is currently guided by its clinical and pathological presentation, including traditional pathology markers of cellular changes, organ penetration, and spread. This situation for bladder cancer is important because researchers hope and expect that genetic markers for characterizing bladder cancer will be discovered and validated, and targeted therapies will then be developed. However, increased barriers to accessing treatment, insurance coverage, or reimbursement would slow down those advances, ultimately harming the care for people who develop bladder cancer in the future, and thus society as a whole.	We agree with this comment and hope that insights into the genetic basis and cellular mechanism underlying bladder cancer will lead to advances in diagnostics and therapeutics.

3.	The diagnostic determinations for people who currently have bladder cancer are critical for guiding appropriate care choices – as is well illustrated from just one diagram in the NCCN's 118-page July 2020 "Clinical Practice Guidelines in Oncology" for Bladder Cancer reproduced below. This diagram is relevant because it conveys the complex clinical decisions that people with bladder cancer must make with their clinicians to determine the treatment options and care plan that are best for them. That shared decision-making process is critical to ensure the patient receives the most appropriate treatment for himself or herself.	We agree with this comment that the decision making for patients with NMIBC resistant to BCG is complex. We have attempted to highlight this in our section on controversies and uncertainties.
4.	Uncertainties and Assumptions The draft report summarizes – and attempts to analyze – the clinical trial data for two experimental treatments. The draft report states that the "review focused on clinical benefits, as well as potential harms (treatment-related AEs) of these agents compared to each other and to systemic pembrolizumab and intravesical gemcitabine ± docetaxel." However, the draft report then indicates it was not possible to conduct such direct comparisons, and the entire review was done via modeling with significant uncertainty in the assumptions, making it hard to see the value of the conclusions. The extent of the limited and problematic data underlying the draft report's "analysis" is stated in the draft report itself: "Feedback received during this project recommended against comparing nadofaragene firadenovec or oportuzumab monatox to each other or to the comparators. Differences in study population, design and outcomes were felt to be too great to compare results. The lack of a placebo or standard treatment group in the Phase III trials examined make this particularly challenging."	As stated in the draft report, our goal was to compare these new agents to each other and to the selected comparators. We did this and concluded that the nature of the evidence did not permit a direct or indirect comparison of results. We believe this is important in that there soon may be a number of FDA approved agents for NMIBC resistant to BCG. As such, additional studies comparing these new drugs to each other and potential other agents is warranted. We agree that not being able to make direct or indirect comparisons makes for less robust conclusions. However, we still believe the information provided will be of value.
5.	We also note that there have been recent reports about pembrolizumab indicating that it may not be as effective as previously thought for treating bladder cancer. Since receiving FDA approval in 2017 for use in bladder cancer, pembrolizumab has had its approved label for bladder cancer modified several times. This is important not only for the treatment of individuals with bladder cancer, but it points out – once again – the ever-evolving nature of biomedical science and best practices for clinical care. It is one reason why the NCCN updates its guidelines so frequently, and why ICER's process of doing reviews before there is sufficient data, and cross-compound comparisons without actual data is dangerous – particularly when ICER is reticent to update its own work when new data is available.	In our report, we have highlighted the limitations of the drugs being reviewed, including pembrolizumab. For any new drug, a key issue is uncertainty about whether the benefits and side effects seen in the published trials are similar or different from those seen in actual clinical practice. That said, waiting for new data is problematic because patients, clinicians and payors are all having to consider the pros and cons of the drug as it is being used, with the uncertainties cited.
6.	The application for nadofaragene firadenovec received a complete response letter from the FDA in May 2020 concerning some manufacturing issues, so it is unclear when this treatment will be available for patients. And for oportuzumab monatox, according to the company, it is "on track to complete the BLA submission in the fourth quarter of 2020 and anticipates potential approval in mid-2021." Both those points of information should be included in the report.	Issues pertaining to when a new drug may become available goes beyond the report's emphasis on the drug's evidence of benefits and side effects and its potential cost-effectiveness. This information may be a point of discussion at the public meeting.
7.	While Section 3 correctly notes that neither of the two agents have been approved by the FDA, we did find a preliminary clinical use policy from national carrier Centene from February 2020, which stated its policy would be effective upon the date of FDA approval and that its use criteria "will mirror the clinical information from the prescribing information once FDA-approved." We point this out to indicate that health insurance companies – in this case one that provides commercial as well as Medicaid plans – are thinking ahead and preparing for coverage decisions about new treatments prior to FDA approval. Clearly, they are doing this using their internal review and evaluation processes, and not relying on ICER to do this for them. As we've repeatedly pointed out, doing that makes sense since they need to determine what is appropriate for the population of people for whom they are providing health insurance, rather than some generalized assessment about the "cost-effectiveness of different care pathways for broad groups of patients.	We have updated Chapter 3 to include a summary of Centene's preliminary clinical policy for nadofaragene firadenovec.

Other		
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 claims made for the value of products; they cannot be evaluated empirically nor can the claims be replicated. 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.	Cost-effectiveness analyses including cost per QALY estimates have been used for decades by academic researchers, international health technology assessment agencies, and pharmaceutical manufacturers. The results of these models may be calibrated with other data and analyses and are often replicated by other researchers.
2.	This conclusion rests on the failure to recognize the limitations imposed by the axioms of fundamental measurement. You focus on constructing simulated QALY claims yet we know that the utility score (typically the EQ-5D-3L/5L) is an ordinal measure. 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.	This comment is based on the premise that the EQ5D is considered an ordinal scale. However, the widely held belief is that the EQ5D-3L and -5L can estimate 243 and 3125 unique health states and is widely accepted to possess interval scale properties. A report by Weinstein, Torrance, and McGuire (Value in Health 2009; 12: S5-S9.) described criteria needed for multi-attribute utility instruments to be considered for use in estimating QALY. The EQ5D is described as one of the scales that meet the necessary minimum criteria.
3.	Unfortunately, the draft evidence report for bladder cancer, with the model developed by Professor Touchette and colleagues in the College of Pharmacy Modelling Group, University of Illinois at Chicago, also apparently believe (or at least they have an understanding) that the EQ-5D-3L utility scale has 'ratio' properties. There is no defense of this position or a proof for this belief. If ICER and the University of Illinois group wish to explore this further, I would recommend a recent peer reviewed paper by myself and a colleague (note in particular the peer-reviewers' comments). Perhaps the Illinois group could provide a proof that the EQ-5D has ratio properties (and even demonstrate that it has by default interval measurement properties).	We agree that the EQ5D itself is not a ratio scale. We disagree that a ratio scale is necessary for estimation of utility for use in producing QALY estimates. Ratio scales are necessary only when needing to multiply or divide values along the continuum of the scale. The requirements of calculating a QALY in our model required only that the scale produced an equal magnitude of difference for each point on the scale. Therefore, an interval scale was required. As discussed above, the EQ5D possesses interval scale properties and is considered as meeting the criteria for producing QALY estimates when multiplied by time.
4.	You may recall that in the public comment window for ulcerative colitis, I raised a number of questions designed to establish the basis for your belief in the ratio scale property of the EQ-5D; specifically your ability to provide a proof of this claim. Your response to these questions indicated that you could not provide a proof. Your response reads: "We (and most health economists) have the understanding (emphasis added) that the EQ-5D (and other multi-attribute instruments) do have ratio properties. The EQ-5D value sets are based on time trade-off assessments (which are interval level) with preference weights assigned to different attributes. We fail to see why this should be considered as an ordinal (ranked) scale. ICER believes that the dead state represents a natural zero point on a scale of health-related quality of life. Negative utility values on the EQ-5D scale represent states considered worse than dead."	While we do not agree that ratio properties are necessary for estimation of utility for use in producing QALY estimates, some have argued that the QALY may satisfy ratio scale properties. For a discussion of the scale properties of the QALY model (including TTO), please see: Roudijk et al., Medical Decision Making 2018; 38(6):627–634.

5.	A detailed rebuttal of this rather strange and inconsistent response has been published strange response. Rather than repeat these comments (although it might be noted that the TTO does NOT have interval properties), ICER should be asked once again to provide a proof that the EQ-5D, which features in the bladder cancer report, has a ratio scale. It is somewhat self-defeating to maintain that the EQ-5D-3L has a natural zero and in the next sentence point out that EQ-5D can create negative utility values. ICER cannot have it both ways: a pseudo-ratio scale with negative utilities and a natural zero point? It is not clear what a natural zero point means. In the case of the EQ-5D-3L the zero is simply an artifact of the equation or algorithm that creates the utilities. Unlike, for example, a true zero in measuring weight (i.e., you can't have negative weights). If ICER or the academic group at the University of Illinois are not sure of this, they might refer you to the standard textbook on health technology assessment (Drummond et al. 4th Eds. pg. 148) . As detailed in a number of my publications, the I-QALY is an impossible construct which means, by extension, that your reference case value assessment framework is invalid . It is up to you, but I would think you should advise your audience in ICER subscribers and the various formulary assessment groups, and PBMs of these limitations on your imaginary modelled recommendations.	We believe it is logical to assume that individuals would consider the quality of life of being dead as equal to zero, and that people can conceive of some states as being "worse than dead."