



**Gene Therapy for Hemophilia B and an
Update on Gene Therapy for Hemophilia A:
Response to Public Comments on Draft Evidence Report**

November 2, 2022

Table of Contents

<i>Manufacturers</i>	2
BioMarin.....	2
CSL Behring.....	7
Genentech.....	11
Novo Nordisk.....	17
Sanofi	21
<i>Patients/Patient Groups</i>	24
National Hemophilia Foundation, Hemophilia Federation of America and Coalition for Hemophilia B.....	24
<i>Other</i>	29
Paul Langley (College of Pharmacy, University of Minnesota)[this grid includes only feedback relevant to the current review, please refer to public comment folio on our website for all comments provided]	29

#	Comment	ICER Response
Manufacturers		
BioMarin		
1.	<p>BioMarin believes that ICER’s economic model may underestimate the health benefit achievable with valoctocogene roxaparvovec as measured by quality-adjusted life–years (QALY), due to: Use of non-comparable bleeding data for the emicizumab arm. As acknowledged by ICER in the Draft Evidence Report, the most relevant comparator population for valoctocogene roxaparvovec is Group D of the HAVEN 3 clinical trial of emicizumab, as those study participants had been receiving adequate factor VIII (FVIII) prophylaxis prior to study treatment. However, in the Economic Model section of the report, the input for the annualized bleeding rate (ABR) for the emicizumab arm appears to be based on Group B, derived from “<i>Participants receiving previous episodic therapy with factor VIII</i>”.¹ In Group D, the ABR for all bleeds was 3.3², compared with 2.6 for Group B¹ as currently used in ICER’s economic model. <u>BioMarin recommends that ICER updates the values for emicizumab ABR with HAVEN 3 Group D data to reflect a population most comparable to the study populations of valoctocogene roxaparvovec.</u></p>	<p>Thank you for this suggestion. We had used HAVEN B to be consistent with the prior report, but we agree with the comment, and we have switched to the HAVEN 3 Group D data in the Report.</p>
2.	<p>Over-estimation of bleeds in the valoctocogene roxaparvovec arm. The presumed high ABR at FVIII activity level <5 IU/dL used in the model, based on epidemiological data of mild and moderate hemophilia, and the simplified assumption related to return to prophylaxis lead to over-estimation of bleeds for the valoctocogene roxaparvovec arm. A more reasonable assumption to inform the relationship between bleeds and exposure to transgene-produced factor activity level should be informed by valoctocogene roxaparvovec clinical trial data. Recent ad-hoc analysis from GENER8-1 (NCT03370913) suggested that valoctocogene roxaparvovec-treated patients with FVIII activity level <5 IU/dL who did not switch back to prophylaxis had a median ABR of 3.07 (all bleeds)³ rather than >7 as currently estimated in the model. Patients who switched to prophylaxis had a median ABR of 6.89 (all bleeds), which is consistent with epidemiological data.⁴ The median FVIII activity level at switching was 1.95 IU/dL (by one-stage</p>	<p>We are inclined to keep the base case in part to maintain consistency with the prior report. We will also keep the scenario analyses that speak to potential variance around this assumption. Further, we will acknowledge the limitations in projecting switching patterns.</p>

	<p>assay).³ Based on the above evidence, BioMarin advises ICER to cap ABR (all bleeds) at 3.07 for the remaining cohort rather than an ABR (all bleeds) of ~7 projected from epidemiological data.</p>	
3.	<p>Potential under-estimation of health-state utility benefit associated with valoctocogene roxaparvovec. BioMarin appreciates that ICER has listened to the community voice and considered the broad evidence base by implementing an annual utility increment in the valoctocogene roxaparvovec arm for patients who remain off prophylaxis. We would like to note however that 0.01 utility difference in the base-case model is likely to be a lower-bound estimate of the true utility benefit. The incremental utility difference of 0.01 was determined for treatment administration alone (one-time administration with 10-year benefit vs. 1-2 subcutaneous injections per month) from a patient preference/time trade-off study.⁵ Additional benefits of a gene therapy, including the freedom to pursue sports, and having a more active lifestyle and career choices, have been demonstrated in a range of patient-reported outcomes from GENE8-1 and are likely to translate into additional utility gain.⁵ The same patient preference/time trade-off study also showed that improvements in other coreHEM measures, such as change in mental health and chronic pain, also have a significant impact on patient treatment preference and health state utilities. Nevertheless, health-state utility difference between gene therapy and routine prophylaxis, including emicizumab, warrants more research due to multiple well-recognized issues, such as the sensitivity of the generic utility tools, and the disability paradox in hemophilia. BioMarin is committed to continue working with the patient community and experts to generate further evidence on this topic. Overall, BioMarin agrees with the use of 0.01 as a conservative estimate for the utility difference.</p>	<p>Thank you for bringing this to our attention. There were typographic errors in the supplemental section of the report for different values, we have fixed this. We adjusted all the language to reflect an academic in confidence value for valoctocogene roxaparvovec which we felt was the best evidence-based estimate based on data provided to ICER as cited in the report.</p>
4.	<p>BioMarin is concerned that the given clinical rating (C++) for the net health benefit of valoctocogene roxaparvovec vs. FVIII prophylaxis does not reflect the totality of evidence to date. Based on clinical efficacy and safety results already available from the largest phase 3 program and longest follow up (from the phase 1/2 study) of any gene therapy to date, BioMarin believes that the evidence supports a rating of B+.</p>	<p>Thank you for your perspective. We feel that the concerns around decreasing efficacy over time, potential for long term oncogenesis, and the loss of the potential for future gene therapy raise enough concern to merit the C++ rating.</p>

5.	<p>Efficacy: The net health benefit of valoctocogene roxaparvovec vs. FVIII prophylaxis has been demonstrated in the largest phase 3 trial (GENEr8-1; NCT03370913) and the longest follow up (from the phase 1/2 trial; NCT02576795) of any investigational gene therapy. In GENEr8 1 (N=134), 112 participants underwent a prospective observation period of at least 6 months prior to valoctocogene roxaparvovec administration to establish baseline ABR and FVIII use while receiving standard-of-care FVIII prophylaxis. After 2 years of follow up post valoctocogene roxaparvovec administration, there was a 77% reduction in mean total ABR and an 85% reduction in mean treated ABR over the whole study period compared with baseline levels associated with FVIII prophylaxis.^{6,7} The relative benefit of valoctocogene roxaparvovec on ABR vs. baseline FVIII prophylaxis was clinically and statistically significant ($P<0.001$) and was consistent between year 1 and year 2.⁶ In addition to ABR reduction, valoctocogene roxaparvovec was associated with a 98% reduction in use of FVIII compared with baseline over 2 years of follow up ($P<0.001$).⁶ Notably, participants in GENEr8-1 demonstrated broad and consistent improvements in health-related quality of life sustained through 2 years post gene transfer. Measured by the Hemophilia-specific Quality of Life Questionnaire for Adults, clinically meaningful improvements were reported in the domains Role Functioning, Consequence of Bleeding, Worry, and Treatment Concern. The Hemophilia Activities List assessment tool revealed that the largest activity-related improvement was in the area of leisure activities and sports, while the Work Productivity and Activity Impairment Questionnaire plus Classroom Impairment Questionnaire: Hemophilia Specific assessment tool showed less classroom and work impairment after gene transfer. The observed benefits associated with valoctocogene roxaparvovec vs. FVIII prophylaxis in bleeding control, FVIII use, and quality of life warrant an incremental improvement rating.</p>	<p>Thank you. We agree that there is evidence supporting clinically meaningful benefits, hence the C++ rating, which includes the potential for substantial net health benefits when considering both efficacy and harms.</p>
6.	<p>Efficacy: Although post-administration FVIII activity exhibited a steady decline over time, the rate of decline became shallower after the initial decline, proportionate to the initial peak. Importantly, ICER, in the Clinical Benefit section of the report, appear to have misinterpreted valoctocogene roxaparvovec 1-year data from GENEr8-1 to assume that the majority of trial participants resumed prophylaxis once their FVIII activity levels</p>	<p>Thank you for the information. Note, that you cited a different abstract in your comment, but we found the correct abstract and slide set and have updated the draft report to reflect the conference abstract.</p>

	were <5 IU/dL (page 10). On the contrary, existing data suggest a sustained treatment effect, with only 6 (<5%) participants from GENE8-1 returning to prophylaxis (N=134) within 2–3 years of follow up, and all participants from the phase 1/2 study 6E13 vg/kg cohort remaining off prophylaxis through 6 years.	
7.	<p>Efficacy: The hemostatic efficacy and quality-of-life benefit of valoctocogene roxaparvovec for patients with persistent low endogenous FVIII levels warrants further research. However, we believe it is premature for ICER to conclude that the benefit of valoctocogene roxaparvovec is ‘short-lived’. Given the importance of ABR, FVIII use, and patient’s quality of life as patient-relevant outcome measures when determining treatment effectiveness for hemophilia A, we believe there is high confidence in the evidence base to date to demonstrate incremental net health benefit for valoctocogene roxaparvovec vs. FVIII prophylaxis as a one-time infusion for people with severe hemophilia A who want the opportunity to live bleed-free for years without prophylaxis.</p>	Thank you for your comment. In the report, we do not say “is short lived.” The text reads “could be relatively short lived.” Given the lifelong need for adequate factor levels, both patients and hematologists expressed this concern to us.
8.	<p>Safety: In the report, ICER notes that a high risk of elevated liver enzymes requiring prolonged corticosteroid therapy is a concern with valoctocogene roxaparvovec. Most adverse events from GENE8-1 trial were of grade 1 or 2, and the frequency of adverse events decreased over time.^{6,11} No participants withdrew from the study because of adverse events and none showed development of FVIII inhibitors. No thrombotic events, drug-related cancers, or drug-related deaths have been reported. It is important to note that the GENE8-1 protocol was flexible with regard to oral corticosteroid treatment post infusion, which was initiated at the investigator’s discretion based on elevations in alanine aminotransferase (ALT) above the upper limit of the normal range or falls in FVIII activity post-valoctocogene roxaparvovec infusion. It is acknowledged that investigator conduct in the trial demonstrated a propensity for steroid use.¹¹ The European label of valoctocogene roxaparvovec reflects a more restrained use of steroids.¹² In the absence of an alternative cause for the ALT elevation, a corticosteroid regimen should be promptly initiated at a daily dose of 60 mg prednisone (or equivalent dose of another corticosteroid) for 2 weeks. The daily corticosteroid dose can be gradually tapered in a stepwise manner over 6 weeks. We will monitor the use of</p>	<p>Thank you. We believe that we accurately reported the use of steroids for liver enzyme elevations (79.1%) and their mean duration (34.7 weeks) as reported in Mahlangu 2022.</p> <p>We look forward to evaluating real world data reflecting the European experience when they become available.</p>

	immunosuppressant regimens in the real world in Europe, and it is expected that immunosuppression with steroids will be less intensive than was observed in the clinical trials.	
9.	<p>Suboptimal adherence with prophylaxis regimens compromises outcomes in the real world. One-time administration with gene therapy eliminates major issues with adherence to routine prophylaxis. BioMarin recommends that ICER’s comparative effectiveness review takes real-world outcomes into consideration: adherence with FVIII prophylaxis has been reported to vary widely, with some patients self-reporting adherence to be in the range of 30–96%.¹³ Missed doses have consequences. In the 270-902 study (a non-interventional lead-in study of GENEr8-1), adherence to FVIII prophylaxis in both pre-baseline and on-study period was >90%.¹⁴ Despite high levels of adherence to prophylaxis, participants had continued occurrence of spontaneous and joint bleeding events requiring treatment and leading to impaired physical functioning. It is of note that valoctocogene roxaparvovec has demonstrated clinical superiority over baseline FVIII prophylaxis in the highly adherent cohort from the 270-902 study.¹⁵ Adherence in the real world is expected to be much lower than in the 270-902 study, therefore, one might expect a larger magnitude of benefit if compared with a cohort of patients with more usual, variable levels of adherence to prophylaxis. <u>We believe this is a distinction that merits consideration when defining the clinical benefit rating for valoctocogene roxaparvovec vs. standard-of-care treatments.</u></p>	<p>We agree with the statement in bold. Ideally, we would have randomized comparisons of bleeding rates following gene therapy compared with bleeding rates on factor prophylaxis.</p> <p>It may be that the participants in the trial were more adherent with prophylaxis than the average patient on factor prophylaxis, but adherent patients tend to do better independent of therapy. Thus, they likely did better after gene therapy than less compliant patients would do after gene therapy. This remains an area of uncertainty that contributes to our overall C++ interpretation.</p>
10.	The outcomes-based warranty agreement that BioMarin plans to offer for valoctocogene roxaparvovec is designed to address variability and durability concerns for payers and health systems. BioMarin agrees with ICER that this outcomes-based warranty design should be incorporated and remain in the base base-case modelling.	Thank you for the comment.
11.	<p>Factual corrections: Misinterpretation of valoctocogene roxaparvovec 1-year data.</p> <p>On <u>page 10</u> of the Draft Evidence Report, it is stated <i>“In the GENEr8-1 trial, 16 participants (12.1%) had factor VIII levels < 5 IU/dL and 12 participants (9.1%) had levels < 3 IU/dL. Presumably, the majority of these continued factor prophylaxis, though the</i></p>	Thank you. As noted above, we have updated the report to include the actual numbers of patients on prophylaxis as reported at the July 2022 EAHAD meeting.

	<p><i>details are not reported</i>". It is not accurate to assume that the majority of trial participants with low factor level (<5 IU/dL) continued factor prophylaxis. On the contrary, in the 2-year GENER8-1 results presented at EAHAD 2022¹⁶ and BioMarin's previous data submission to ICER, BioMarin reported that only 5 of 31 participants with factor <5 IU/dL (by chromogenic substrate assay) at week 104 had switched to prophylaxis (1 participant switched to prophylaxis at factor >5 IU/dL). The majority continued to be prophylaxis-free and many did not experience bleeding to date.</p>	
12.	<p>Inaccurate reference to valoctocogene roxaparvovec clinical rating in ICER 2020 review. On page ES2, it is stated that ICER gave valoctocogene roxaparvovec a C++ rating compared with factor VIII prophylaxis in its 2020 review (page ES2). This statement is inaccurate. ICER rated the evidence as "promising but inconclusive" (P/I) in the 2020 review.</p>	<p>Thank you. We have corrected the error.</p>
13.	<p>Inconsistency in the reference to utility increment used in Hemophilia A model. In the Draft Evidence Report, ICER makes three different references to the utility gain implemented in the Hemophilia A model:</p> <ul style="list-style-type: none"> • "Valoctocogene roxaparvovec was associated with a utility gain of 0.01 based on data submitted to ICER" (page 18) • "Valoctocogene roxaparvovec was associated with a fixed utility gain of 0.02 per cycle as long as patients did not switch therapies based on data submitted by BioMarin" (page E4) • "...fixed utility gain of 0.03 was used for both gene therapies in their respective models based on EQ5D data provided by BioMarin and CSL Behring" (page E12) <p>Based on the model results, we believe that ICER referenced the patient preference/time trade-off research and implemented a utility gain of 0.01 per year in the model. Please update these references to ensure consistency in the Updated Evidence Report.</p>	<p>We have updated the model to reflect an updated estimate based on academic in confidence data.</p>
<p>CSL Behring</p>		
1.	<p>Under the uncertainties and controversies section, ICER notes that <i>"it is not yet clear etranacogene dezaparvovec will have the same long-term decline in factor levels that has been observed with valoctocogene roxaparvovec, though the decline appears to be less rapid, if it occurs at all"</i>. CSL</p>	<p>Thank you. We agree with the biological rationale questioning the validity of any extrapolation about the transgene expression of Factor VIII to Factor IX. Hence our statement "if it occurs at all." That said, the follow-up has been relatively short and there is</p>

	<p>Behring appreciates that extrapolating factor activity beyond the follow-up period of a trial is innately uncertain. However, we disagree with this phrasing.</p> <p>First, we note that, mechanistically, there is a significant scientific rationale to anticipate that transgene expression of FIX will be more sustained than transgene expression of Factor VIII (FVIII). For example, both etranacogene dezaparvovec and valoctocogene roxaparvovec target hepatocytes to produce FIX and FVIII, respectively.³⁻⁵ However, native production of FIX occurs exclusively within the hepatocytes, while FVIII is produced within sinusoidal endothelial cells of the liver.⁶ Thus, etranacogene dezaparvovec more closely approximates endogenous production of its clotting factor than valoctocogene roxaparvovec by targeting cells that naturally produce its transgene product. Furthermore, the comparative size of the FVIII gene in contrast to the FIX gene requires the use of shorter promoter elements to fit within the AAV5 cassette size, which necessitates higher vector doses to achieve similar response.⁶ Higher vector doses are associated with loss of transgene expression due to elevation of liver aminotransferases.⁷ Recently, Herzog and Pierce (2022) concluded that, while there are uncertainties around the comparativeness of vector doses/titers and infectivity, the most obvious hypothesis at first glance to explain the greater variability in durability in Hemophilia A gene therapy trials in contrast to Hemophilia B gene therapy trials is related to FVIII itself.⁶ We encourage ICER to add statements that contextualize the differences observed in durability between etranacogene dezaparvovec and valoctocogene roxaparvovec.</p>	<p>some evidence supporting declining levels over time. The long-term outcomes remain uncertain.</p>
<p>2.</p>	<p>In addition to the comments above, there is a more obvious objection. Published data from 3 years of the phase 2 study of etranacogene dezaparvovec and five years of a study of the AMT-060 precursor (which differs only in using wild-type rather than Padua FIX) clearly demonstrate lower levels of decline than trials of valoctocogene roxaparvovec.^{3; 5; 8} While we acknowledge that we cannot say for certain that etranacogene dezaparvovec will have <i>no</i> long-term decline of FIX production, any decline seen within the first three to five years will be more gradual than in valoctocogene roxaparvovec. While the second clause of the ICER quote does implicitly</p>	<p>We agree. This, in part, contributes to the B+ rating for etranacogene dezaparvovec and a C++ rating for valoctocogene Roxaparvovec.</p> <p>We have accepted your recommendation to change the phrase to read “more gradual” rather than “less rapid”.</p>

	<p>acknowledge this, we note that the phrase “<i>the same long-term decline</i>” indicates that is possible that etranacogene dezaparvovec and valoctocogene roxaparvovec could have similar patterns of gene expression decline. However, patterns of decline will differ in at least the first three years based on the best available data. Therefore, we suggest to rephrase this to “<i>it is not yet clear etranacogene dezaparvovec will have a long-term decline in factor levels as has been observed with valoctocogene roxaparvovec, though decline appears to be more gradual, if it occurs at all</i>”</p>	
<p>3.</p>	<p>In addition to the above comment, CSL Behring has concerns about the statement: “<i>Finally, the long-term impact of the therapy on liver function and the potential for oncogenesis remain a concern</i>”. First, we note that AAV5 mediates expression primarily from episomes and has a low probability of chromosomal DNA insertion, which reduces the risk of insertional oncogenesis.⁹ Long-term evaluations of patients with AAV gene therapies for over ten years have failed to demonstrate any gene therapy-induced oncogenesis.^{10; 11} We acknowledge that there was one patient in the HOPE-B trial that developed hepatocellular carcinoma (HCC). However, as ICER notes, this was considered not related to treatment. This assessment was made by independent molecular tumor characterization and vector integration analysis. As noted in a recent press release: “<i>AAV vector integration in the patient’s tissue sample was extremely rare and accounted for 0.027% of the cells in the sample. The integration events were distributed randomly across the genome, and there was no evidence of clonal expansion or any dominant integration event. Additionally, whole genome sequencing of the tumor confirmed that the patient had several genetic mutations that are characteristic of HCC and are independent of vector integration. Finally, gene expression analysis of the tumor and adjacent tissue suggested a precancerous state in the liver consistent with several risk factors that predispose this patient to HCC.</i>”¹² Taken together, the investigation strongly suggests that etranacogene dezaparvovec did not contribute to this particular case of HCC and it is unclear on what other basis the potential for oncogenesis remains a concern. Therefore, we recommend amending the statement to remove the clause around potential for oncogenesis or for</p>	<p>Thank you, but we respectfully disagree. Experts have told us that the large vector dose (6E13) will cause tens of millions of random integrations into liver cells and other tissues. Given the small number of treated patients and the short follow-up time, there remain concerns about oncogenesis, which may take 10 or more years from initial therapy to clinical detection of a cancer. It may turn out not to be an issue, but many experts remain concerned. For example, see Pierce et al. Hemophilia. 2020 Nov;26(6):931-933.</p>

	ICER to elaborate more clearly on their rationale for this statement.	
4.	<p>In their review framework, ICER cite etranacogene dezaparvovec as having a B+ rating. This indicates “... <i>that there is moderate certainty of a small or substantial health benefit with high certainty of at least a small net health benefit (B+) for etranacogene dezaparvovec compared with factor IX prophylaxis</i>”¹ In this review, ICER notes that causes for uncertainty include uncontrolled study design, small patient numbers and relatively short follow-up time. The HOPE-B trial design followed the Food and Drug Administration (FDA)’s Guidance for Industry on human gene therapy for hemophilia.¹³ The six-month lead-in data collection on annualized bleeding rate (ABR) on optimized prophylaxis allowed for within-subject comparison, which the FDA notes “<i>increas[es] the statistical power relative to a design with parallel control.</i>”¹³ We acknowledge that the HOPE-B trial did not have as many patients as phase 3 trials for non-rare conditions, but we note that the number enrolled is greater than a typical rare disease trial and note that the trial was sufficiently powered to draw conclusions.^{4; 14} We are also unclear on what relative criteria ICER concludes that HOPE-B had “<i>relatively short follow-up</i>”. Data provided to ICER academic-in-confidence had follow-up that met the FDA’s Guidance for Industry criteria for “Long-Term Monitoring”.¹³ The HOPE-B trial is planned with at least five years of follow-up which should serve to reduce extrapolation-related uncertainty in the future.¹⁵ However, waiting until the full five-year data cut rather than evaluating using the current data would require PWHB to wait for over an additional three years for a therapy that has already firmly established superior efficacy.</p>	<p>Thank you, but we respectfully disagree. Our rating agrees with your judgement that there is net benefit, but the trial design, small number of treated patients and only 18 months of follow-up for a therapy designed to last for a lifetime provide sufficient uncertainty to warrant a B+ rating rather than an A or B rating.</p>
5.	<p>More broadly, we disagree that there is only moderate certainty that etranacogene dezaparvovec will provide a substantial health benefit, as it is our position that the HOPE-B trial demonstrates a substantial health benefit in and of itself. As this data is empirically demonstrated, it provides high certainty of a substantial health benefit. As ICER noted, the majority of patients within HOPE-B who completed administration of etranacogene dezaparvovec were able to discontinue FIX prophylaxis for least eighteen months. When considering only patients that would match the expected label indication restrictions for baseline anti-AABV titers, 100% of patients that received a full dose etranacogene</p>	<p>Again, we respectfully disagree. Given the short follow-up we feel that there remains important uncertainty about whether or not there will be important declines in Factor IX levels over time. The use of the AAV vector with this therapy precludes its use in future gene therapy if this therapy fails. Finally, ICER typically would not rate the evidence from small, uncontrolled studies with follow-up that is short relative to its irreversible lifetime impact on patients as high certainty.</p>

	<p>dezaparvovec were able to discontinue FIX prophylaxis therapy. This is <i>per se</i> a net health benefit, by reducing the burden of self-infusions which directly contributes to quality of life of PWHB and their caregivers.^{16; 17} The benefit of treatment extended beyond prophylaxis discontinuation; patients had highly statistically significant improvement in ABR after receiving etranacogene dezaparvovec. Treated ABR was 3.65 during the six-month lead-in and 0.84 at months 7-18 post-dose (rate ratio [RR] = 0.23, $p < 0.0001$) and treated joint ABR was 2.13 during the six-month lead-in and 0.44 during months 7-18 post-dose (RR = 0.20, $p < 0.0001$). Over one year on average, this represents over 2.5 fewer treated bleeds per patient, and over 1.5 fewer treated joint bleeds per patient. Clinical benefits translated into improvements in quality of life as measured by the Haem-A-QoL Total, and Feelings, Treatment, Work/School, and Future subdomains. While a “substantial health benefit” is a subjective assessment, our position is that eliminating prophylaxis, reducing ABR to the magnitude seen within HOPE-B, and improving patient quality of life for over one year is sufficient to constitute a substantial health benefit. Additionally, the stability of the AMT-060 precursor for five years is strongly suggestive that clinical improvements would persist for at least five years.⁸ We would ask ICER to reconsider, or more clearly explain the criteria for determining the magnitude of health benefits for future reviews.</p>	
Genentech		
1.	<p>Revise the economic assessment to make full use of the available clinical trial evidence, or at minimum, state the results in context to the limitations of the data that was used.</p> <p>Recommendations: Update the efficacy inputs in the economic model to make full use of the GENE8-1 trial outcomes, including annualized bleed rates (ABR), among the defined population of PwHA previously receiving factor VIII (FVIII) prophylaxis.</p> <p>Rationale: A health technology assessment should be based on evidence of a product’s clinical efficacy. The cost-effectiveness model should be built upon the highest level of clinical evidence [10-13]. As the hierarchy of evidence dictates, level one evidence from clinical trials should preferentially be applied as parameter estimates over data from natural history</p>	<p>The GENE8-1 data are incorporated in model 2 in the early cycles as described in the methods along with adjustments for having 2% per year fail therapy and switch to emicizumab. However, we switch to the natural history numbers once those projected bleed rates are higher as described in the report. The model conclusions are robust to multiple sets of sensitivity analyses and scenario analyses around bleed rates, durability, and switching rates.</p>

	<p>studies [14]. For this assessment, ABR outcomes from the GENER8-1 trial should be used to model patient outcomes instead of using FVIII activity levels from a natural history study [2]. The approach that ICER chose based on the aforementioned natural history study, would lead to a clinically significant <i>underestimation</i> of bleeding events compared to simply applying GENER8-1 trial outcomes. For example, among the subgroup of 13 PwHA in the GENER8-1 trial who had <i>worse</i> outcomes 6 months after valoctocogene roxaparvovec [5], the actual mean ABR was approximately 5 bleeds per year. Whereas, the approach that ICER described would have resulted in an estimated ABR of 2 bleeds per year.</p>	
2.	<p>Another important problem to consider is the lack of clinical trial evidence comparing valoctocogene roxaparvovec to emicizumab. Clinical trial evidence from the GENER8-1 trial provides evidence to support a comparison between FVIII prophylaxis and valoctocogene roxaparvovec [5]. Unfortunately, ICER’s economic assessment does not make use of the available comparative clinical data. Without FVIII as a treatment option in this economic assessment, the common comparator between emicizumab and valoctocogene roxaparvovec is missing, precluding a rigorous indirect comparison. ICER’s own clinical assessment concluded the same, that there was insufficient evidence to make a comparison between emicizumab and valoctocogene roxaparvovec. This should be accounted for in the results of the economic assessment.</p> <p>Indirect treatment comparisons also require that populations across studies be reasonably comparable, with similar inclusion and exclusion criteria [15]. Populations treated in emicizumab trials [7,8,16,17] are substantially different from PwHA treated in the GENER8-1 trial [5]. These limitations should be explicit in ICER’s economic assessment.</p> <p>Implications: Failure to make full use of the GENER8-1 trial evidence leads to an economic assessment that relies too heavily on assumptions and places insufficient weight on the current evidence available, leading to an inaccurate conclusion about the relative value of valoctocogene roxaparvovec.</p>	<p>We’ve have incorporated the Gener8 data and we have added language acknowledging the indirect comparison in the limitations section.</p>
3.	<p>Quantify and describe uncertainty in the economic assessment, or at minimum, state results in context</p>	<p>We did this primarily through pessimistic and optimistic analyses as part of the “Single or Short-Term Transformative Therapies” (SST) framework.</p>

	<p>to assumptions about the treatment response and duration of effect.</p> <p>Recommendations: ICER should (A) describe the economic impact of uncertainty about treatment durability; (B) more accurately capture the heterogeneity of the patient journeys seen in the GENER8-1 trial through more robust sensitivity analyses, instead of relying on averages that fail to describe the range of observed experiences with valoctocogene roxaparvovec; (C) align the economic assessment results with those of the clinical assessment, or at least provide more transparency around the misalignment between the clinical and economic assessments.</p> <p>Rationale: Due to the limited durability of gene therapy over a lifetime horizon, it would not be a viable option to make an economic comparison between emicizumab versus valoctocogene roxaparvovec <i>alone</i>. At the core of the economic assessment, the draft evidence report presents “Results for the Base-Case for Valoctocogene Roxaparvovec Compared to Emicizumab” [1]. However, the economic assessment actually compares emicizumab alone to the implicit sequence of <i>valoctocogene roxaparvovec followed by emicizumab</i>. Based on model assumptions that ICER applied in their economic assessment, <i>100% of patients who received valoctocogene roxaparvovec would eventually initiate emicizumab</i> after 12 years (cycle 24) due to low FVIII activity (<1 IU/ mL). From a modeling perspective, the inclusion of subsequent therapy constitutes “sequential therapy” rather than “monotherapy”. When PwHA who are not “cured” after valoctocogene roxaparvovec then receive additional prophylaxis, over a lifetime they have experienced the costs and effects of <i>both</i> therapies.</p>	<p>We have also looked at durability thresholds for which we added language in the report.</p> <p>However, this is not an attempt to say that valrox on its own provides more QALYs, or is dominant to emicizumab alone. It is of Valrox followed by emicizumab when factor levels decline compared to emicizumab alone.</p> <p>We have added language in the economic summary section to stress that this is an attempt to look at the economic value of valrox in the context of there being emicizumab. It would be a mistake to view this as valrox in a world where emicizumab does not exist.</p>
4.	<p>Additionally, while the GENER8-1 trial data supports ICER’s prediction that all PwHA receiving valoctocogene roxaparvovec would eventually require re-initiation of prophylaxis, we disagree with the assumption of an “average” patient journey including re-initiation of prophylaxis after 12 years. Such an assumption does not account for the heterogeneity in the treatment effect of valoctocogene roxaparvovec, observed in the clinical trials [5,18]. Over 2 years of follow-up after treatment with valoctocogene roxaparvovec, 5% of trial participants re-initiated prophylaxis, with at least 24% of all trial patients having FVIII levels in the</p>	<p>We have 8% of the patients “fail” and switch to emicizumab in the first 8 cycles which is in line with the data from GENER8-1 trial.</p>

	<p>moderate to severe range [18]. Therefore, at the very least, sensitivity analyses should incorporate the full range of these heterogeneous patient experiences with valoctocogene roxaparvovec to better reflect the uncertainty related to outcomes.</p>	
5.	<p>Furthermore, we disagree with ICER’s approach to address uncertainty by applying base-case assumptions that align with a value-based contract proposed by BioMarin, which excluded data on patients with high ABRs after receiving valoctocogene roxaparvovec [1]. Since value-based contracts are confidential agreements between manufacturers and payers, there is <i>no transparency</i> about the net financial impact of these contracts, and they <i>do not</i> constitute evidence to inform an objective economic assessment. Applying an assumption that aligned with the aforementioned value-based contract could bias the economic results in favor of valoctocogene roxaparvovec, overstating the benefits of the new therapy.</p>	<p>ICER’s SST framework suggests modeling an outcomes-based arrangement if the manufacturer suggests one. Although the exact parameters of this contract are not made clear in this report as they were submitted in confidence, we are able to describe the failure rates and payments that would be expected under this plan.</p>
6.	<p>The draft clinical assessment accounted for uncertainty related to the heterogeneity in patient experiences, however the draft economic assessment did not [1]. Unless the economic assessment is updated to align with the clinical assessment, there remains a disconnect between the clinical and economic assessments, which should be clearly explained in the executive summary.</p>	<p>The model is not a comparison of valrox alone vs emicizumab, It is of Valrox followed by emicizumab when factor levels decline compared to emicizumab alone.. Further, it is not intended to be applied to any individual patient. The models look at the average value of therapy in a population.</p>
7.	<p>In the economic assessment, ICER reported valoctocogene roxaparvovec to be dominant versus emicizumab, even though the clinical comparative-effectiveness analysis determined that there was insufficient evidence to draw a conclusion [1]. The clinical assessment underscored the high uncertainty due to multiple factors (e.g., uncontrolled trial design with small numbers of patients in the studies, relatively short follow up periods, lack of long-term understanding of net benefits and impact on liver function and risk of hepatocellular carcinoma). At present, the excessive clinical uncertainty discussed in the clinical review is almost entirely absent from the economic assessment, leading to conclusions about dominance that can easily be misinterpreted by the healthcare decision-makers that are the end users of the evidence report. Furthermore, the face validity of the economic assessment appears misaligned with the uncertainty in the clinical assessment, as every probabilistic sensitivity analysis (PSA) ended in the same outcome. Such results suggest that the PSAs are not capturing the key</p>	<p>Please see our comment above. Further, as alluded to above, there are a wide variety of sensitivity analyses and scenario analyses applied to help address potential variance in the average estimates.</p>

	issues that need to be explored through more robust sensitivity and scenario analyses.	
8.	<p>Given the adeno-associated virus (AAV) mechanism of action of valoctocogene roxaparvovec, it is important to note that PwHA only have one opportunity to take an AAV gene therapy. As recently stated in a review by Drs. Samelson-Jones and George, “Potential recipients will need to fully understand that they can receive AAV gene therapy only once, and the loss of transgene expression of FVIII must be considered in their risk/benefit calculus” [19].</p> <p>Implications: As PwHA will likely only have one opportunity to receive a gene therapy, it is vital that ICER present results clearly, and with appropriate context about the uncertainty, to support informed decision-making that accounts for all relevant information. Without addressing the limited durability of valoctocogene roxaparvovec and the heterogeneity of patient experiences, readers of this assessment could easily misinterpret the economic results as depicting a lifetime cure that eliminates the need for future prophylaxis.</p> <p>Similarly, the large discrepancy between the clinical assessment and economic analysis may cause healthcare decision-makers to question the validity of ICER’s dominant economic result. Stakeholders may misinterpret these mixed results, which could lead to negative consequences with regard to patients’ access to therapies.</p>	Thank you for your comment. We have added language to address this into the controversies and uncertainties section.
9.	<p>Apply current clinical practice guidelines to the economic assessment by revising, or at least, testing assumptions that were made about the re-initiation of prophylaxis after valoctocogene roxaparvovec.</p> <p>Recommendations: (A) Define a cure versus a treatment failure after valoctocogene roxaparvovec based on guideline-concordant [20] need for prophylaxis, using FVIII activity levels and ABRs reported in the GENE8-1 trial. (B) At the very least, ICER should state the economic results with respect to the assumptions that were made about re-initiation of prophylaxis.</p> <p>Rationale: The economic assessment involved assumptions related to re-initiation of prophylaxis that do not reflect the current standard of care. According to the World Hemophilia Federation (WFH) [20], recommendations on prophylaxis include the following:</p>	We are doing this in the first 8 cycles according to the GENE8 data. We are also keeping the switch rules and bleed calculations consistent with the prior report. We also conduct scenario analyses that change the switching and bleed rate assumptions. The basic conclusions of the model are robust across these scenarios.

	<p>“ For patients with hemophilia A or B with a severe phenotype (note that this may include patients with moderate hemophilia with a severe phenotype), the WFH strongly recommends that such patients be on prophylaxis sufficient to prevent bleeds at all times, but that prophylaxis should be individualized, taking into consideration patient bleeding phenotype, joint status, individual pharmacokinetics, and patient self-assessment and preference.”</p> <p>Despite the guidelines recommending sufficient prophylaxis to prevent all bleeds, the ICER economic assessment assumes that re-initiation of prophylaxis occurs after a projected 7 bleeding events annually [1], which is excessive, unrealistic, and unethical. Additionally, the guidelines state that PwHA with a severe phenotype (which could include low FVIII or high ABR criteria) should be offered prophylaxis to target a trough FVIII level of 3-5% [20]. However, ICER seems to require a FVIII level <1% to model re-initiation of prophylaxis [1]. Notably, emicizumab prophylaxis is reported to meet FVIII level equivalence of >10% in those receiving therapy [21,22]. The gap between ICER’s assumptions and the current clinical practice warrants testing different assumptions about re-initiation of prophylaxis after valoctocogene roxaparvovec (e.g., under different thresholds of FVIII levels and ABRs), to better reflect current practice and the individualized approach recommended by the clinical guidelines.</p>	
10.	<p>Another important problem with the current assumptions underlying the draft economic assessment relate to the current options for hemophilia A prophylaxis, which include FVIII products as well as emicizumab. Given the real-world prophylaxis utilization rates, ICER should include FVIII replacement products as an option for prophylaxis after failure of valoctocogene roxaparvovec, rather than limiting prophylaxis only to emicizumab. In the American Thrombosis and Hemostasis Network (ATHN) registry, 46% of the participating PwHA without inhibitors used emicizumab, while 28% infused standard half-life (SHL) FVIII products, and 18% utilized an extended half-life (EHL) recombinant factor concentrate for prophylaxis [23]. The ATHN data demonstrates that the WFH guidelines have been incorporated into current practice, showing how “prophylaxis should</p>	<p>The model purposely examines valoctocogene roxaparvovec followed by emicizumab relative to emicizumab from a value perspective as emicizumab was found to be dominant to factor VIII in the prior model. The purpose is to show the relative value of valoctocogene roxaparvovec compared to the next most valuable treatment which is emicizumab.</p>

	<p>be individualized” [5]. It is imperative that the real-world options for hemophilia A prophylaxis are included in the assessment to reflect the current treatment landscape and clinical guideline recommendations.</p> <p>Implications: If ICER does not apply the current guidelines and model real-world scenarios for initiation of prophylaxis, the economic assessment will overestimate the efficacy and durability of valoctocogene roxaparvovec, and substantially underestimate the lifetime costs. Additionally, if ICER does not account for the current standard of care and treatment options for prophylaxis, the resulting economic assessment will not reflect the reality that PwHA and healthcare decision-makers face. Ultimately, this oversight could result in misinformed decisions about access to therapies and pricing negotiations that could negatively impact patients. ICER has an obligation to prevent such misunderstandings and ensure fair access to treatment for PwHA.</p>	
<p>Novo Nordisk</p>		
<p>1.</p>	<p>The market basket derived from the HOPE-B clinical trial is not representative of real-world practice.</p> <p>The pivotal phase 3 HOPE-B clinical trial evaluated the efficacy and safety of etranacogene dezaparvovec (EtranaDez), an investigational adeno-associated virus five (AAV5)-based gene therapy, among adults with severe or moderately severe hemophilia B⁵⁻⁸. The trial was designed as an open-label, multi-center, multinational, single arm study. The reference therapy was described as prophylactic factor IX replacement therapy used during the 6-month lead-in period prior to treatment with EtranaDez. The primary outcome measure, annualized bleeding rate (ABR), was compared between the lead-in period with prophylactic factor IX replacement therapy and following administration of EtranaDez.</p> <p>In its cost-effectiveness analysis, ICER determined that clinical effectiveness data for the market basket comparator was derived from the 6-month lead-in period from the HOPE-B trial. The market basket comparator is comprised of Alprolix, BeneFIX, Idelvion, and Rebinyn. There are several concerns with this approach:</p>	<p>We received the market basket from CSL Behring and we believe that this was the best available data for a consistent comparison to factor IX. Patient adherence, patients not considering gene therapy, and other factors can impact the numbers seen in the trial relative to other populations. Hence, using the trial data for both arms reduces selection bias when comparing the same patients to themselves. We also include language in the limitations section acknowledging these are trial data and may have limited generalizability, but still maintain it is best for the relative comparison of the treatments in terms of value.</p>

	<ul style="list-style-type: none"> Key model input data for the comparator’s effectiveness was extracted from a small, multinational trial of 56 participants recruited through 35 sites throughout the world, which likely underrepresents a US-only population of heterogeneous patients with hemophilia B. 	
2.	<ul style="list-style-type: none"> The ABR for the 6-month lead-in period from the HOPE-B trial was 4.19, which was substantially (numerically) higher than the ABRs from the pivotal trials of each therapy in the market basket^{3,9-11}. The substantially higher ABR for the 6-month lead-in period from the HOPE-B trial indicates sample selection bias when applied to ICER’s economic model, namely that patients eligible for the HOPE-B trial did not respond well to, or were non-compliant with, prophylactic factor IX products and thus elected to participate in a gene therapy clinical trial. In its draft evidence report, ICER acknowledged that sample selection bias may be present; however, ICER stated that the ABRs observed during the run-in phase of the HOPE-B trial were comparable to those reported in a recent systematic review of ABRs for patients with hemophilia B on prophylactic therapy (Davis 2019)¹². There are two primary issues with ICER’s use of Davis 2019 as supportive evidence: <ul style="list-style-type: none"> a. Davis 2019 includes a comparison of Idelvion to Alprolix, BeneFIX, Ixinity, Refixia (Rebinyn), and Rixubis. It is important to note that ICER’s market basket comparator does <i>not</i> include Ixinity or Rixubis, limiting the relevance of this publication for comparative purposes. b. Furthermore, Ixinity (mean [SD] ABR: 3.55 [7.15]) and Rixubis (4.26 [5.80]) were associated with two of the highest trial reported mean ABRs among the interventions under study in Davis 2019. A simple average across the remaining interventions in the publication yields a mean ABR of 2.3, which is substantially lower than the ABR of 4.19 used for the market basket comparator in ICER’s model, 	<p>Same as above.</p> <p>We don’t have adequate evidence to separate out bleeding rates for any factor IX replacements used in the run-in period. If high quality evidence becomes available, we believe that providers and patients will take this into account.</p>

	further refuting the relevance of this publication for comparative purposes.	
3.	Since Rebinyn was recently approved by the US Food and Drug Administration (FDA) for the routine prophylactic treatment of adult and pediatric patients with hemophilia B ¹³ , it is unlikely that many patients were treated prophylactically with Rebinyn during the lead-in period of the HOPE-B trial.	See above. Note also that Rebinyn made up just slightly more than 2% of the market basket. We added a footnote to clarify this in the supplement.
4.	<p>ICER's incorporation of HOPE-B trial data for the model comparator has profound implications on its model findings. Notably, in ICER's one-way sensitivity analysis, the model was highly sensitive to bleeding rates associated with the factor IX comparator (i.e., treated target joint bleeds; treated nontarget joint bleeds; joint bleeds); these three parameters were among the top 10 variables that had the largest impact on model results. As currently constructed, a higher ABR for the model comparator inappropriately yields a lower CE ratio in favor of EtranaDez, and specifically a dominant strategy in ICER's current analysis.</p> <p>Recommendation: ICER should consider conducting scenario analyses using the average ABR across the individual trials of the four therapies comprising the model's market basket of factor IX therapies. Doing so will better reflect the benefits of the current standard of care in the United States today.</p>	See above. We have probabilistic sensitivity analyses and scenario analyses around different bleed rates.
5.	<p>Rebinyn should not be grouped with other factor IX products in a single comparator given its differential, and more beneficial, pharmacokinetic (PK) and efficacy profiles</p> <p>Rebinyn (Nonacog beta pegol) is a glycopegylated recombinant factor IX protein with an extended terminal half-life. In July 2022, the US FDA extended Rebinyn's approval to routine prophylactic treatment to prevent or reduce the frequency of bleeding events in adult and pediatric patients with hemophilia B¹³. The approval was based on data from the main phase of the adult and pediatric previously treated patients (PTP) trials, which demonstrated median ABRs of 1.04, 2.00, and 0 in patients 13 years of age and older, 7 to 12 years of age, and ≤6 years of age, respectively¹³.</p> <p>Additional clinical trials have been conducted comparing Rebinyn to other factor IX prophylactic therapies. For instance, a phase 1 study comparing 25, 50, and 100 IU/kg doses of Rebinyn to BeneFIX</p>	See comments 2 and 3 above.

	<p>found an approximate five-fold increase in half-life associated with Rebinyn (93 hours for Rebinyn vs. 19 hours for BeneFIX)¹⁴. Moreover, the study found a two-fold increase in recovery rate associated with Rebinyn compared to BeneFIX. A graphical representation of these outcomes follows (Appendix – Figure 1).</p>	
6.	<p>Similarly, a second phase 1 study comparing a single dose of Rebinyn 50 IU/kg to a single dose of Alprolix 50 IU/kg found a four-fold increase in factor coverage associated with Rebinyn (96.6 IU x h/mL for Rebinyn vs. 22 IU x h/mL for Alprolix)¹⁵. Moreover, the study found that Rebinyn was associated with a two-fold increase in the recovery rate and a six-fold increase in factor levels at seven days compared to Alprolix. A graphical representation of these outcomes follows (Appendix – Figure 2).</p> <p>These phase 1 studies demonstrate superior clinical properties for Rebinyn compared to interventions even within the same therapeutic class. ICER's aggregation, therefore, of factor IX products in a single market basket comparator ignore the clinical benefits of Rebinyn.</p> <p>Recommendation: ICER should include separate discussion regarding the benefits of each individual factor IX product, notably Rebinyn, in the summary section of its final evidence report.</p>	See above.
7.	<p>Conservative estimates should be applied to the durability of gene therapy in the absence of long-term clinical evidence</p> <p>Gene therapies in hemophilia B have the potential to reduce or eliminate the need for further prophylactic therapy with factor IX products. In the HOPE-B trial of EtranaDez, nearly all patients discontinued factor IX prophylaxis during the duration of the clinical trial period (i.e., months 7-18); however, evidence of the long-term duration of benefit of this gene therapy beyond the clinical trial duration does not exist. Contrary to this uncertainty, ICER's model assumes a much longer-term durability of treatment associated with EtranaDez before resulting in an increase in ABR (i.e., >20 years). As currently constructed, a lower ABR for EtranaDez maintained for a long period of time prevents (1) increases in pharmacy and direct medical costs and (2) decreases in QALYs, inappropriately yielding a lower CE ratio in favor of</p>	<p>We have a conservative durability scenario in the SST scenarios. We also included language on this in the uncertainties and controversies section. And we have done durability threshold analyses for which we added language in the report.</p>

	EtranaDez, and specifically a dominant strategy in ICER's current analysis.	
8.	<p>Instead of applying a liberal assumption to its base case analysis, ICER should take a more conservative approach to modeling the long-term treatment effect of EtranaDez given the lack of evidence in this regard. In fact, in ICER's 2022 evaluation of betibeglogene autotemcel (beti-cel; Zynteglo) for the treatment of beta thalassemia, ICER assumed that a small percentage of patients transitioned from a health state of transfusion independence to a state of transfusion dependence after seven years, citing the uncertainty of long-term durability of beti-cel treatment effect over time¹⁶. More specifically, concerns regarding the limited number of patients (N=32) in the phase 3 trial of beti-cel and the limited duration of data (seven years of data for three patients) led to ICER's conservative assumption. For comparison, the HOPE-B trial of EtranaDez only evaluated 56 patients over 18 months, including a 6-month run-in period; yet ICER does not maintain consistency in its methods in the current hemophilia assessment. Moreover, in published models of gene therapies in hemophilia A, 10% of patients were assumed to fail gene therapies after a short time period (i.e., 4 weeks) and resume prophylactic therapy with factor VIII products^{17,18}.</p> <p>Recommendation: ICER should incorporate a consistent and more conservative assumption regarding the durability of EtranaDez treatment effect over time given its deep impact on model findings. This more conservative assumption should be considered in the base case analysis, or at the least in scenario analyses, to ensure a more balanced and evidence-based assessment.</p>	ICER is clear that it evaluates all gene therapies separately. Base case assumptions around durability of specific gene therapies have been based on available evidence and expert input on that evidence. ICER has not assumed that all gene therapies have similar durability.
Sanofi		
1.	The report's executive summary mentioned the observed risk of elevated liver enzymes requiring prolonged immunosuppressive therapy with its unknown long-term impact, and the hepatocellular carcinoma risk of gene therapy. While these risks and costs were not integrated into the model or widely discussed subsequently, they are a significant impact and consideration for treatment decision making. The National Hemophilia Foundation (NHF) recently submitted a citizen's petition to the FDA to this effect recommending that the FDA place a risk evaluation and mitigation strategy (REMS) as a condition of approval to address potential safety concerns. ³	We agree with this concern and now include a disutility of -0.03 for the first cycle after Etranadez and the first two cycles after Valrox.

2.	<p>Recommendation: We encourage ICER to further articulate the limitations arising from long-term uncertainty on its findings. This includes the inherent uncertainty of projecting efficacy from existing studies to a lifetime in the context of a one-time treatment with observed waning effect and unknown safety consequences of observed adverse events and the need for steroid use in the long term.</p> <p>We also recommend that ICER highlight the applicability of model findings in the context of individual patient and doctor risk-benefit assessments specifically within the context of unknown long-term impact. It is vitally important that treatment be determined based on individual patient needs and risk-benefit preferences.</p>	<p>See comments above. The model is meant to examine average group level value not to inform individual decisions. That said, we have several sensitivity and scenario analyses to explore potential variation in the average results including around durability. We also have included language around the limitations of the model inputs.</p>
3.	<p>Assumptions on prophylactic restart after gene therapy effect wane: In both hemophilia A and B models, an assumption of treatment re-initiation of prophylaxis only began at projected factor levels between 1 and 5 IU/mL (moderate severity) whereby 5% of gene therapy patients were assumed to switch to prophylaxis treatment. The National Hemophilia Foundation recommends that prophylactic therapy may also be considered for people with hemophilia with moderate and mild hemophilia with a severe phenotype to prevent bleeds. They also recommend that prophylactic therapy should be instituted early (prior to the onset of frequent bleeding). In Hemophilia A, this is supported by observations that higher endogenous FVIII activity levels (well above 5 %) is associated with a trend of lower bleeds than factor activity levels between 1 and 3%⁴⁻⁶</p> <p>Recommendation: Given the long-term uncertainty with factor level projections, need for treatment re-initiation, and data showing factor levels above 5% are associated with reduced bleed -a guideline recommended treatment goal, we recommend that ICER keep ABR as an endpoint (and that patients treated with gene therapy re-start on prophylactic treatment after 5 years as a base case for hemophilia A). Furthermore, we recommend that ICER conduct sensitivity analysis assuming 100% of the gene-therapy-treated patients require prophylactic treatment after 1, 2, 3, 4 and 5 years (for hemophilia A).</p>	<p>See above.</p> <p>GENEr8 data indicate that Valrox fails for at most 5% of patients after two years and we include a similar assumption in our model.</p> <p>Also as stated above, we have modeled optimistic and pessimistic scenarios for time to failure of gene therapy but do not believe the data suggest a five-year time window as a base case.</p>
4.	<p>Use of Petterson Score (PS) scores in model structure: The advances in factor and non-factor therapies in the recent years have significantly</p>	<p>We believe the Petterson score is the best evidence-based mechanism for incorporating the long run impact of bleeds. Patients start with a</p>

	<p>altered the disease outcomes for patients who are now extensively treated with prophylaxis, and with a greatly reduced need for surgeries. The PS structure widely applied for classification of arthropathy with increasing scores culminating in the need for joint surgery may not be as relevant today. For example, following the draft model assumptions, let us define an annualized joint bleed rate of 1 in a prophylactic-treated hemophilia A population (for simplicity, the life expectancy of each individual is equal to 100 years exactly). Let us also consider the link between joint bleeds and PS (36.52 or 6.52 joint bleeds result in a one-point PS increase in patients aged less or more than 25 years respectively) as well as the link between PS and surgery (a PS of 28 leads to surgery). In this specific population, the average newborn would undergo their first surgery after 392.56 years ($= 25/100 * 1 * 36.52 * 28 + 75/100 * 1 * 6.52 * 28$). This extremely long time to first surgery makes latest stage of PS unlikely on a lifetime prophylactic approach</p> <p>Recommendation: In patients treated prophylactically, a modelled PS stage is unlikely to result in surgery and hence may not influence the model results much. Nevertheless, it is important to take into account that the on-demand regimen is still widely used with poorer outcomes, higher bleed rates and higher potential for surgery in both Hemophilia A and B. Therefore, we recommend that ICER highlight the rarity of achieving surgical state with new prophylactic or gene therapy treatments.</p>	<p>score of 14 and transition through the model to higher states that include 28, where they receive surgery. As we acknowledge, it likely undercounts surgeries but it does not take 390 years for them to get there. We were not sure how you were coming to that number. Is there a specific citation you are referring to? It's also worth noting that the conclusions of the model are insensitive to the surgery rates.</p>
5.	<p>Budget impact: ICER concluded that the 5-year budget impact expectations did not warrant the creation of a dedicated budget impact model. However, typically payers consider shorter time horizons (often only 1-2 years) in line with their annual budgets.</p> <p>Recommendation: We would encourage ICER to still comment on the potential affordability and high upfront cost and budget impact implications of gene therapies in the short term. Specifically for smaller US payers, self-funded plans with smaller sizes, or payers who consider budgets on shorter timeframes such as 1-2 years, such budgetary implications and potential need for re-insurance could be of importance.</p>	<p>In line with ICER's Value Assessment Framework and Reference Case, a 5-year time horizon was chosen for the budget impact analyses in order to balance near-term estimates (e.g., 1-2 years) of financial impact and affordability with realistic durations over which uptake of new technologies may occur (e.g., 5+ years). We anticipate that nuanced discussions surrounding smaller plans, immediate upfront costs, and approaches for handling these costs may arise during our policy roundtable discussion during the corresponding public meeting on November 18th.</p>

#	Comment	ICER Response
Patients/Patient Groups		
National Hemophilia Foundation, Hemophilia Federation of America and Coalition for Hemophilia B		
1.	<p>Financial Toxicity / Health Benefit Price Benchmarks: Given the existing high life-time cost of hemophilia treatment, it is not surprising that gene therapies for both hemophilia A and B were deemed dominant and cost effective at a placeholder price of \$2.5 million. As ICER has noted and we have agreed during prior hemophilia reviews, the current cost of hemophilia treatment is “financially toxic” for PwH, their families, and the health care systems on which they depend. The finding that gene therapy is cost effective does not mean it is affordable, that it will be accessible within the marketplace post-approval, or that it is an optimal treatment for every eligible patient. We remain concerned that high target prices will impede access to these potentially transformative therapies.</p> <p>Considering these concerns, we call ICER’s attention to a recent Biomarin investor call held following the EU conditional approval of valoctocogene roxaparvovec for hemophilia A held on August 24, 2022. In this call, a “first year free pricing” launch price of 1.5 million euros net of all discounts and reserves was indicated for Germany (the expected initial market). Other EU pricing has not yet been announced but is anticipated this month. Taking into account the “financial toxicity” of currently available treatments, we caution against gene therapy pricing in the US being justified based on a multiple of the current treatment cost. While there are notable differences between US and European health systems and their respective current market prices for current therapies, the net resulting German price of 1.5 million euros is useful as a reference point for the Health Benefit Price Benchmarks discussion included in the final report.</p>	<p>We agree that there are major issues around financial toxicity and affordability among treatments for hemophilia including the new gene therapies. This Report focuses on shared savings models from ICER’s SST framework in the HBPBs that are presented.</p>
2.	<p>Risk Mitigation: In the last full paragraph on page 2 of the Background, to be clearer, we suggest the last sentence commenting on the approval of valoctocogene roxaparvovec in Europe (“Valoctocogene roxaparvovec was approved for the treatment of severe hemophilia A adults on August 24, 2022, by the European Commission.”) be moved to the end of the prior paragraph and not be part of the paragraph discussing etranacogene</p>	<p>Thank you for pointing out the inappropriate location of the sentence on the EU authorization of valoctocogene roxaparvovec. We have adopted your suggested changes.</p>

	<p>dezaparvovec. Further, we would suggest this sentence be revised to reflect the conditional nature of the approval as follows: “Valoctocogene roxaparvovec <i>received conditional market authorization with requirements for additional monitoring</i> for the treatment of severe hemophilia A adults on August 24, 2022, by the European Commission.”</p> <p>This EU authorization and mandatory risk management requirements are consistent with the recent National Hemophilia Foundation Citizen’s Petition requesting that, if the FDA approves a hemophilia gene therapy in the United States, the Agency should require a rigorous risk evaluation and mitigation strategy (REMS) to ensure the health and well-being of PwH who receive a gene therapy product.</p>	
3.	<p>Disability Paradox / Health Utility: Thank you for recognizing the “disability paradox” and resulting limitations of population-based health measures, and for responding to our request for the inclusion of a utility gain within the assessment. Although based on the generic EQ-5D-5L, the inclusion of this perspective provides a notable patient-centric improvement over prior hemophilia reviews and the earlier draft Modeling Analysis Plan for this review. We also agree, as noted by ICER on page 3, that the psychosocial impact of hemophilia for both PwH and their caregivers is enormous. The current model does not capture the potential transformation in mental health outlook when a “hemophilia-free mindset” is achieved. Similarly, we agree with the observation on page 5 that the currently available outcome data sets do not yet capture this transformation. We are working to resolve this limitation in outcomes research for PwH and look forward to its consideration in future value assessments. As we have consistently advocated, there are many important outcomes, including a change in psychosocial status, important to PwH beyond a low annual bleed rate or changed joint health score as used in the model. Inclusion of the health utility gain is a positive and welcome step, though we will continue our requests that ICER utilize a hemophilia-specific health utility measure fully incorporating outcomes important to PwH.</p>	<p>We thank the hemophilia community for their efforts in generating relevant data. We acknowledge this in the controversies section.</p>
4.	<p>Step Therapy / Return to Prophylaxis: As noted above, we remain concerned about our community’s ability to access innovative therapies due to the cost of treatment and the insurance barriers they face. In ICER’s prior hemophilia review</p>	<p>Thank you. We hope to cover these important topics at the upcoming public meeting on November 18th.</p>

	<p>(November 2020), it reported several policy recommendations that are important for the current review:</p> <ul style="list-style-type: none"> • Payers considering implementing formal step therapy, however, should recognize the heterogeneity of patient experience with factor VIII and its different delivery mechanism. • Payers should cover factor VIII prophylaxis at levels adequate to achieve higher troughs than the 1% level used in the past. <ul style="list-style-type: none"> - All payers should be aware of the widespread consensus among clinical experts and patient organizations that a trough factor VIII level of 3%-5% should be viewed as a minimum target for the vast majority of patients. <p>The decision whether or not to undergo gene therapy should remain with the PwH and their health care providers. We oppose barriers that threaten this shared decision-making, such as step therapy through and failure on emicizumab prior to accessing valoctocogene roxaparvovec. Besides impeding access for those for whom gene therapy is the preferred option, there are also several unresolved clinical challenges, including how best to transition from a non-factor replacement therapy with a 30+ day half-life (e.g., whether a wash out period is required), laboratory monitoring, and potential lifestyle modifications required for PwH during transition to gene therapy if they were required to first switch to emicizumab. For example, people taking clotting factor therapy designed to manage peak FVIII levels for periods of high activity could find themselves forced to switch to a suboptimal product before they could access gene therapy.</p>	
5.	<p>Equally, decisions on when a PwH may return to prophylaxis when factor activity levels decline should not be impeded by the construct of a performance-based agreement or requirement that a certain number of bleeding events occur. Thank you for taking note of this community concern on page 5 of the draft evidence report. Here, too, prior ICER recognition on the importance of higher trough levels (3-5% minimum target) remains relevant. This position is also consistent with a recent European consensus statement. We reiterate our comments submitted as part of the initial scoping for this review on May 25th:</p>	<p>We agree that the details of an outcomes agreement should not be a basis for clinical care. If there are significant concerns that this could occur, the issue will warrant discussion during the Policy Round Table.</p>

	<p>Today, factor activity level remains the driver of clinical decision making. Linking factor activity level to clinical outcomes is important for this and future evaluations of novel therapies. Although some have noted low bleeding rates even as factor activity levels taper post vector infusion, we are not aware of any published evidence to correlate this finding to patient important outcomes or long-term joint health. We continue to believe achieved factor activity level (e.g., the restoration of lacking clotting capacity) is the best indicator of anticipated long-term clinical outcome and discourage anything besides the most realistic assumptions about durability.</p> <p>Given the significant decline in factor VIII activity levels over time for recipients of valoctocogene roxaparvovec, it is evident that while the product may provide significant benefit for a period of years, it does not represent a long-term cure. Thus, anticipating an eventual return to prophylaxis is important for those considering this treatment. While we would have a similar concern for hemophilia B, data do not indicate a significant decline in efficacy at this time.</p>	
6.	<p>Contextual Considerations and Potential Other Benefits: Within Table 5.2 we offer the following comments:</p> <p>The first potential benefit relates to the ability to achieve major life goals. The information here should differentiate between hemophilia A and B. While highly relevant for hemophilia A, the comment about needing to pick the right period of life to undergo gene therapy appears less apt for hemophilia B given the durability data to date for etranacogene dezaparvovec.</p>	Thank you. We have added the potential distinction between the two therapies.
7.	<p>The second potential benefit relating to caregiver burden has not captured the significant burden on caregivers for adults. We would refer you to data from the US CHES Study, which quantified a significant economic impact on caregivers for adults with hemophilia. This estimate was largely derived from caregivers being part-time employed, unemployed or unable to work, due to their care duties preventing them from working more hours.</p>	Thank you. We have added this to Table 5.2.
8.	<p>The third bullet related to complexity accurately describes prophylaxis. We would also note the significant burden on participants during the first 6- to 12-month period post-gene therapy for</p>	Thank you. We have added the first-year burdens as an offset to the benefits of no longer requiring prophylaxis.

	laboratory testing, barrier contraception, avoidance of liver toxic substances (e.g., alcohol), and corticosteroid management. A high degree of compliance and adherence is required during this period.	
	The fourth bullet relates to health inequities. While it may not be immediately available to all, given the enormous variability in health care coverage in the United States and the significant financial impact (work, career, educational opportunities) hemophilia has on PwH and their families, a durable transformative therapy could have a generational impact that breaks the cycle of economic disadvantage experienced by many in our community.	Thank you. We think that some of this is captured in the section on the patient's ability to achieve major goals, but we have added these observations to the patient input section of the report.

#	Comment	ICER Response
Other		
Paul Langley (College of Pharmacy, University of Minnesota)[this grid includes only feedback relevant to the current review, please refer to public comment folio on our website for all comments provided]		
1.	The first question is whether or not you or your academic advisers undertook a systematic review of the use of preference instruments and quality of life scores in hemophilia A and B? There is a significant literature, but you only cite six papers. Indeed, you fail to point out that the Ballal et al EQ-5D-3L scores are simulated EQ-5D scores for a typical patient with and without knee surgery and the Fischer reference is for the SF-36. More discussion here would help the reader as to why you selected the EQ-5D-3L; apart from its application in your previous study report (and my critique) .	Please refer to our Research Protocol (https://osf.io/bmnz3) for more information on our Evidence Review Methods.
2.	A recent paper by Xu et al note that the EQ-5D may not be able to detect certain changes in HRQoL due to symptoms unique to hemophilia or side effects resulting from its treatment; hence they propose, unwisely, a bolt on methodology for condition specific items to capture better the needs of patients. Note that they utilize the EQ-5D-5L; is there any reason why you stay with the now historic EQ-5D-3L? The application of the EQ-5D-3L and its lack of sensitivity in patients with hemophilia has been noted by other authors. O’Hara et al note that across all five dimensions of the EQ-5D-3L the majority of patients reported no problems, with fewer than one in ten patients reporting extreme problems in any dimension 13. For the few extreme problems reported, pain/discomfort and anxiety/depression were the most common dimensions. It should be noted that as these 3 response levels in the EQ-5D-3L are ranked ordinal responses, the psychometric distance between the various pairs is unknown. The same issue arises with the EQ-5D-5L. Given the ordinal responses (note also that TTO scores are ordinal) it is impossible to see how these can result in an EQ-5D-3L algorithm to produce a ratio scale. Of course, if the majority of hemophilia patients report no problems, claims for any meaningful change in preference or utility would be minimal which means unit prices will dominate imaginary cost-per-QALY claims, supporting threshold criteria for price reductions. In your model you produce imaginary lifetime QALYs that only differ by 0.67 years (17.98 vs. 17.31; Table 4.3).	The bleed disutilities come from a 5L study. The baseline utility scores come from a study based on the 3L that maps well into the health states in our model. It is known that different instruments may produce different scores as well as be more or less sensitive in different situations. Part of the choice of utilities in the model is if they fit with our modeling strategy in terms of key health states, which in our case are primarily PS scores and bleeds. We conduct sensitivity analyses around changes in the utility scores and disutilities in the model. It is true that cost differences dominate the model, and they would do that with any type of utility score given the magnitude of the costs of these drugs.

3.	Perhaps the most comprehensive assessment of the variability in reported utility scores is the Carroll et al study for France and the UK . Applying four instruments to hemophilia populations in each country they found significant variations in the scores reported for the EQ-5D-3L, EQ-5D-5L, SF-6D and the SF36. The authors caution that while each instrument shows similar trends, each led to unique values: In general, the EQ-5D-5L gave higher values than the EQ-5D-3L, while the SF-6D was localized within a small range compared to these two instruments. It would have been useful for ICER to assess studies such as this one, pointing out that the SF-36 has been used widely in hemophilia populations. Unfortunately, none of the papers cited made any reference to the measurement properties of the instruments.	Assessing all available utilities is beyond the scope of this report. We do include sensitivity and scenario analyses that characterize potential variance along those lines.
4.	It is also worth noting that the disease specific, Likert-scale based, Haem-A-QoL (adult) and Haemo-QoL (children and adolescent) instruments for hemophilia produce only composite, ordinal scores. They cannot be applied to assess either health status or response to therapy. This means that efforts to map from the Haem-A-QoL to the EQ-5D-5L are a waste of time. As far as can be ascertained no attempt has been made to evaluate these instruments in terms of the Rasch measurement model; e.g., applying the Rasch Rating Scale Model to evaluate their approximation to an interval score 8. This failure in disease specific QoL in hemophilia mirrors the failure with the generic EQ-5D-3L/5L instruments. There is, of course, a solution: to develop disease specific instruments for adults, adolescents and caregivers in terms of the single needs-fulfillment attribute; to develop from first principles instruments that meet Rasch or modern measurement theory standards which provide empirically evaluable value claims for therapy responses, based on interval measures. Presumably this would be asking too much. Even so, it is important to note that there is an extensive literature and recommendations for outcomes measures in hemophilia which should be addressed if a comprehensive picture of empirically evaluable outcomes is to be presented instead of the limited generic instrument favored by ICER; these include, as you note in passing, recommendations specific to gene therapy.	We agree that disease-specific instruments should be developed for adults, adolescents, and caregivers.
5.	ICER's model oversimplifies the impact of this disease on patients' overall quality of life.	ICER Value Assessment Framework fully acknowledges that all too often what matters

	<p>ICER’s model is based solely on progression through Pettersson scores, which may be an oversimplification of the impact of hemophilia on patients’ quality of life. ICER acknowledges in its report that the burden of hemophilia is not limited to the number of bleeds. Patients also live with uncertainty and chronic pain, which can lead to anxiety, depression, and other mental health challenges. ICER’s simplified model is not able to capture these symptoms, particularly as they present in a very heterogeneous manner and are often exacerbated by the fact that access to specialist services in hematology varies wildly based on location, demographics, and insurance coverage.</p>	<p>most to patients is poorly captured in the available clinical trial data. We know that even when trials do capture the clinical outcomes that matter most to patients, there are other aspects of value related to the complexity of the treatment regimen or the impact of care options on the ability to return to work, on the negative impact of the condition on family and caregivers, on public health, or on other aspects of the health system or society. The ICER value framework identifies these “potential other benefits” as important elements of any overall judgment on long-term value for money, and all ICER reports have separate sections in which evidence and information pertaining to these elements are presented. Similarly, decisions about the value of care options do not happen in a vacuum. There may be broader contextual issues related to the severity of the condition. So even if these details are not captured in the model, the ICER value framework includes this element and it is explored in a separate section of each ICER report.</p>
6.	<p>The health state utilities used are problematic and they likely underestimate the net effect of treatments under review.</p> <p>There are several issues with the primary source for health utilities in the model that led to the data being oversimplified and unable to generate accurate utilities. These issues include that the three biggest drivers of utility in the Multivariate Poisson distribution are: being in the United Kingdom, age, and target joints (yes or no). The utility set is only varied on age, and a dichotomous proxy of target joints, which has only two states. The number of target joints was a significant driver of the results, so having just two categories is likely too simplistic and may undervalue the effect on quality of life of reducing bleeds.</p> <p>This data was questioned in a paper by some of the same authors one year later who put forward a strong case that the EQ5D results (and hence the health state utility values derived from them) are considerably over-estimated in the Hemophilia A population, as there is considerable evidence of disease-based hedonic adaption in the patient population. The</p>	<p>We welcome any new data around health state utilities, and we encourage manufacturers to collaborate with the research community to (a) develop these metrics and (b) use these in their clinical trials.</p>

	<p>authors go on to suggest that a preference elicitation study using multiple populations could be used to generate better HSUV for Hemophilia patients in the future.</p> <p>These findings suggest that the utility values used in the ICER model are underestimates and their use may dilute the full effect of the therapies being evaluated, showing them to be less cost-effective than they are.</p>	
7.	<p>The bleed-related disutilities are assumed to apply for just two days. This is an overly conservative assumption and does not reflect the sources being used for these inputs.</p> <p>The disutility inputs are based on one study involving 37 people, who were followed for 90 days. Of the patients who completed a diary, they reported 194 bleeding episodes in 90 days; that’s just over 5 per patient over 3 months and equates to 10 over the six-month cycle in the model. Yet the model uses a separate source for number of bleeds per cycle. Bleed rates used in models 1 are 4.19, for the standard of care arm, varying by factor level between 0.45 and 7.28 per cycle.</p> <p>The physical effect of a bleed is rarely limited to two days, and the mental manifestations can last for weeks, so the application for 2 days in the model is not an accurate representation of the patient experience. It also ignores the marginal impact on quality of life from the rate of bleeds over time. The difference between having ten bleeds a cycle and have 1 or 2 bleeds a cycle is considerable from a patient perspective.</p> <p>The conclusion of the paper from which these inputs are derived states that the results “indicated that frequent acute bleeds impair QOL beyond patient’s non-bleed day baseline.” This nuance is ignored in the model due to the linear function linking bleed rate and impact on quality of life, underestimating the burden of patients who have multiple bleeds in a short period.</p>	<p>We have included scenario analyses in the report that speak to potential variance around these assumptions.</p>
8.	<p>The model assumes no mortality effects.</p> <p>The model assumes no mortality effects, but presence, severity and treatment burden are all drivers of mortality in Hemophilia so ignoring mortality in a lifetime model does not present an accurate picture.</p>	<p>In our review of the evidence we did not see sufficient differences in mortality across treatments used in the model or between patients on emicizumab or factor IX relative to average US mortality rates to warrant inclusion in the model. Further the model results are robust to a wide variety of scenario and</p>

<p>ICER's assessment states that as prophylaxis for hemophilia A in patients without inhibitors has not been demonstrated to decrease mortality there should be no mortality effect because of these treatments being studied. This is an oversimplification, as the same paper suggests that severity of Hemophilia has a strong impact on relative mortality. It states that the mean life expectancy of someone with severe hemophilia is 63 compared to someone with mild or moderate hemophilia of 75. If the severe form of the disease can take 12 years off a patient's life expectancy, then it highly likely that treatment to alleviate the root cause of the disease and its consequences will result in a lower mortality.</p> <p>A recent Swedish cohort study based on a long-standing Hemophilia registry showed that the hazard ratio for all-cause mortality for those with Hemophilia A compared with controls was 1.7, ($P < 0.001$) when patients with HIV and/or viral hepatitis were excluded. The corresponding figures for the severe hemophilia subgroup were 6.6, ($P < 0.001$). This occurred even though those with Hemophilia were 57% less likely to die from ischemic heart disease than controls.</p> <p>There is also evidence of reduced inhibitors being a known risk factor for morbidity. Even though findings were mixed in smaller studies as to their role in mortality, recent larger studies suggest that they are a factor, increasing risk of death by up to 70%. A similar study suggested a mortality rate five times higher in Hemophilia A patients with inhibitors, than those without.</p> <p>A model that is assuming no mortality effects from treatments deemed to be 'highly effective' in a disease known to have higher rates of premature death, will underestimate the true value of these therapies.</p>	<p>sensitivity analyses whereby unless there were major mortality differences across the treatment arms the results would be robust.</p>
--	--