Voretigene Neparvovec for Treating Inherited Retinal Dystrophies Caused by RPE65 Gene Mutations: An Evidence Review Group Perspective of a NICE Highly Specialised Technology Appraisal

C. Farmer*[1], A. Bullement[1,2], D. Packman[1], L. Long[1], S. Robinson[1], E. Nikram[1], A.J. Hatswell[1,2], G.J. Melendez-Torres[1], & L. Crathorne[1]

1 Peninsula Technology Assessment Group (PenTAG), University of Exeter, Exeter, UK

2 Delta Hat, Ltd. Nottingham, UK

*Corresponding author: C. Farmer, PenTAG, South Cloisters, St Luke's Campus, Heavitree Road, Exeter EX1 2LU. c.farmer@exeter.ac.uk

ORCID: C. Farmer 0000-0003-3553-8970

Running header: An ERG review of voretigene neparvovec for RPE65-mediated IRD

Declarations

Funding: The appraisal of voretigene neparvovec within the NICE HST programme was commissioned by the NIHR Systematic Reviews Programme as project number 17/56/05.

Conflicts of interest/competing interests: No competing interests

Availability of data and materials: Documents relevant to this appraisal (including the appraisal scope, company submission (CS), ERG report, stakeholder submissions, and NICE guidance issued) can be found on the NICE website at https://www.nice.org.uk/guidance/hst11

Authors' contributions: All authors commented on the submitted manuscripts and have given their approval for the final version to be published. CF, LL, AH, GMT, and LC critiqued the clinical burden and clinical effectiveness methods and evidence; SR critiqued the literature searches conducted by the company; AB, DP, EN, and AH, critiqued the health economic method and evidence, and the company's economic evaluation. LC supervised the appraisal and acted as guarantor.

This article has not been externally peer reviewed by PharmacoEconomics

Acknowledgements

The views expressed in this paper are those of the authors and not necessarily those of the NHS, the NIHR SER programme or the Department of Health and Social Care. Any errors are the responsibility of the authors.

S. Bello, S. Dodman, M Rezaei Hemami, A Churchill, and A Lotery all contributed to the ERG appraisal of voretigene neparvovec. A. Churchill and A. Lotery acted as clinical advisors to the evaluation and reviewed the epidemiology and clinical effectiveness sections of the report.

Voretigene Neparvovec for Treating Inherited Retinal Dystrophies Caused by *RPE65* Gene Mutations: An Evidence Review Group Perspective of a NICE Highly Specialised Technology Appraisal

Abstract

The National Institute for Health and Care Excellence (NICE) considered evidence for voretigene neparvovec (VN; Luxturna®) for the treatment of RPE65-mediated inherited retinal dystrophies (IRD) within its Highly Specialised Technology (HST) programme. This paper summarises the evidence provided by the company; the appraisal of the evidence by the Peninsula Technology Appraisal Group (PenTAG), who were commissioned to act as the independent Evidence Review Group (ERG); and the development of the NICE guidance by the appraisal committee. The evidence presented by the company highlighted the significant life-long burden of IRD for patients and carers. There was a paucity of evidence to support the effectiveness of VN; however, the available evidence showed a modest, sustained improvement across a variety of vision-related outcomes. While patients would remain visually impaired, the committee considered that VN would prevent further deterioration in vision. The modelling approach used by the company had a number of limitations, and relied heavily upon a large volume of clinical expert input in order to produce cost-effectiveness estimates with large uncertainty around long-term effectiveness. The ERG's main concerns revolved around these long-term outcomes, as well as the plausibility of utility values. The NICE committee were convinced that the clinical benefits of VN were important, and an appropriate use of NHS resources within a specialised service. The committee concluded that there was a high unmet need in patients with RPE65-mediated IRD, and that VN represents a step change in the management of this condition.

Key points for decision makers:

- This appraisal demonstrates how clinical judgement, patient experience and 'biological plausibility' may play an important role in addressing gaps in an evidence base.
- Attempts to develop complex economic models that provide a valid representation of patient experience need to also balance reliability when sample sizes are small
- Analyses to explore the potential impact of uncertainties in a limited evidence base, such as threshold analyses, may be informative for decision-making.

1. Introduction

Inherited retinal dystrophies (IRD) are a group of eye disorders caused by gene mutations that result in the gradual degeneration of photoreceptor cells on the retina[1]. In RPE65-mediated IRD, patients exhibit a mutation in the RPE65 gene, which (when functioning correctly) provides an instruction to make a protein that is essential for normal vision. RPE65-mediated IRD is a rare disorder usually diagnosed in childhood, and results in progressive loss of vision, affecting both eyes and ultimately leading to near total blindness[2, 3].

Voretigene neparvovec (VN; Luxturna®, Novartis) is a novel gene therapy treatment that introduces a healthy copy of the defective RPE65 gene into the retinal cells of patients with RPE65-mediated IRD.[4] VN represents the first licensed treatment for patients with vision loss due to IRD caused by confirmed biallelic RPE65 mutations. A course of treatment comprises a single injection into each eye with no further intervention required. In the UK context, VN would be administered within a specialist centre by a consultant surgeon experienced in performing macular surgery, with genetic testing and counselling also available to patients. Treatment with VN is expected to prevent further deterioration of vision for patients[5], yet is associated with a notably high acquisition cost (list price of £613,410 per treated patient).

The company (Novartis, UK) was invited to submit evidence concerning the clinical and cost effectiveness of VN (including information relating to the burden of disease) to the National Institute for Health and Care Excellence (NICE), which was then considered within its Highly Specialised Technology (HST) programme (Box 1).

Box 1 NICE Highly Specialised Technology (HST) programme

The HST programme considers evidence for interventions that meet the following criteria:

- The condition is very rare, chronic, and disabling
- The need for national commissioning of the intervention is significant
- Treatment would be expected to be delivered in a few highly specialised services
- Interventions have the potential for life-long use
- Interventions have a very high acquisition cost

In addition to evidence for the clinical and cost effectiveness of interventions, as is usual within NICE's health technology programme, HST submissions are also required to include evidence for the clinical and economic burden of the condition, the impact of the intervention beyond direct health benefits, and the delivery and budget impact of the specialised service.

The company presented a report containing clinical and cost effectiveness evidence concerning the potential use of VN in routine National Health Service (NHS) practice. The Evidence Review Group (ERG), Peninsula Technology Assessment Group (PenTAG), produced its own independent critique of the company submission and corresponding health economic model. A standing NICE appraisal committee, including clinicians and methodologists, and patient group representatives, considered the evidence presented by the company and the ERG report, and issued guidance on whether to recommend the technology by means of the Final Evaluation Document (FED).

This paper presents a summary of the ERG report and subsequent NICE guidance, from the perspective of the ERG. All documents relevant to this appraisal (including the appraisal scope, company submission (CS), ERG report, stakeholder submissions, and NICE guidance issued) can be found on the NICE website[6].

2. The Decision Problem

The remit of NICE's assessment of VN was to evaluate its costs and benefits within its marketing authorisation for treating IRD caused by RPE65 gene mutations for national commissioning by NHS England[7]. The CS defined the relevant population for VN as "adult and paediatric patients with vision loss due to IRD caused by confirmed biallelic RPE65 mutations and who have sufficient viable retinal cells". The requirement for patients to have sufficient viable retinal cells meant a narrower target population was considered in the appraisal versus the population specified in the NICE final scope[2].

The intervention (VN) and comparator (best supportive care [BSC], which in this population constitutes the use of visual aids and psychological therapy), were both aligned with the final scope.

Outcomes reported included visual acuity (VA), visual field (VF), photosensitivity, and safety. The company also presented data for a novel endpoint, the multi-luminance mobility test (MLMT)[8], developed in collaboration with the FDA, and a modified Visual Function Questionnaire (VFQ)[9, 10] - both of these measures are intended to evaluate 'functional vision' (which refers to the use of whatever vision a patient has). The outcomes were generally in line with the final scope, although no data were reported for health-related quality of life (HRQL) for patients or carers, contrast sensitivity, or the need for cataract surgery.

Evidence was provided by the company concerning the nature of the condition, although much of this evidence was based on a broader population of patients with IRD, and not specific to patients with RPE65-mediated IRD. There was also an absence of trial evidence for the impact of the condition on carers. The company presented evidence for the impact of the technology beyond direct health benefits and on the delivery of the specialised service in line with the final scope. The company provided additional colloquial evidence in the form of submissions from patients, patient organisations, clinicians and NHS England which presented expert views on the impact of the technology, including impact on quality of life for patients and carers.

3. Independent Evidence Review Group (ERG) Review

The ERG reviewed the CS and additional clarification provided by the company in response to queries raised by the ERG. Additional work was conducted by the ERG to further evaluate the clinical and cost effectiveness of VN, which primarily included the development of an alternative economic model using the ERG's preferred settings and assumptions. The following sections summarise the evidence presented by the CS, and the ERG's critique.

3.1. Clinical and Economic Burden

Informed via a systematic literature review (SLR) conducted by the company, the CS described how the disorder is associated with an inexorable and progressive loss of vision, culminating in near or total blindness, though the rate of deterioration varies considerably between patients. The impact of the condition begins in early life, with impacts on child social development arising from poor visual function[11, 12]. Adults may face decreased employment opportunities arising from challenges in accessing education[13]. The company estimated that there were approximately 78 existing patients in the UK who would be eligible for treatment with VN at the time of submission.

Evidence presented showed the impact of IRD on carers and household members through increased caring burden, and an increased risk of depression among patients and their family members[14]. No HRQL trial data for patients or carers was identified, although additional submissions provided by the company provided colloquial evidence for the impact of IRD on the quality of life for patients and carers. There is no available treatment for RPE65-mediated IRD - standard of care involves the use of vision aids and learning support.

3.2. Clinical Effectiveness Evidence submitted by the Company

The company presented the outputs of an additional SLR concerning the clinical effectiveness of VN for treating RPE65-mediated IRD. The SLR identified 22 publications (10 published, 12 unpublished) that reported data from 2 trials, comprising a total of 43 patients with RPE65-mediated IRD:

- **Study 101/102[15, 16]:** an open-label, phase 1, single-arm trial where 12 patients received either a 'low', 'medium', or 'high' dose of VN in a single (worse, non-preferred) eye [Study 101[15]]. After a minimum of 1 year, patients were invited to receive VN in the contralateral eye [Study 102[16]]. Follow-up data was available up to 7.5 years.
- **Study 301/302[17, 18]:** an open-label, multi-centre RCT involving 31 patients that compares a high' dose of VN in both eyes with BSC [Study 301[17]]. After one year, 9/10 (90%) patients from the BSC arm received VN [Study 302[18]]. Follow-up data was available up to 3-/4-years for the BSC and VN arms, respectively. At the time of the appraisal, Study 301/302 was ongoing, and evidence from Study 302 had only been published in conference abstracts/presentations.

Following the administration of VN in Study 301/302[17, 18], the evidence demonstrated a sharp and sustained improvement above the minimal clinically important difference (MCID) for MLMT performance and the VFQ. Improvements above the MCID were also reported for VF and photosensitivity; however, there was no clinically meaningful change in VA or contrast sensitivity. Evidence from Study 101/102[15, 16] also showed numerical improvements in visual outcomes, although this study was underpowered and full data was not consistently reported in the CS.

The most common adverse events (AEs) were related to the administration of VN, which occurred in approximately two-thirds of patients, and were generally mild and/or treatable (cataract, increased

intraocular pressure, nausea, and retinal tear). There was a small (2.44%) risk of serious AEs related to the administration procedure. Safety data for Study 101/102[15, 16] were comparable.

3.3. Critique of Clinical Effectiveness Evidence and Interpretation

The ERG agreed with the company's focus on the subgroup of patients with sufficient retinal cells, as the presence of sufficient retinal cells is required to facilitate the mechanism of action of VN. The description of VN provided by the company was consistent with its marketing authorisation, and the comparator (BSC) was consistent with clinical practice in the UK; though the ERG noted that BSC was provided to patients in both arms of the trials, and therefore the comparison was considered to be VN + BSC compared to BSC alone.

The availability of data from an RCT in such a rare disease area was notable, allowing more confidence that VN is efficacious, although the small sample size meant substantial uncertainty remained in the effect size and whether the trial was representative of the UK population. There is also generally poor understanding of characteristics that impact prognosis and treatment efficacy[5], and therefore the ERG was unable to fully appraise the risk of bias from imbalances between trial arms.

The ERG considered the evidence of improvement in visual function (MLMT and VFQ) to be substantive. However, the ERG was less convinced that the evidence supported long-term maintenance of effect due to the relatively short follow-up of the trials - follow-up in the pivotal Study 301/302[17, 18] was up until 4 years after treatment, at which time data were only available for 4 patients. Study 101/102[15, 16] was also subject to high attrition, with a trend towards vision worsening towards the end of the study, and as such the ERG did not consider the long-term clinical outcome data to be reliable. The ERG therefore considered that it remained unclear whether the benefits of VN demonstrated in the trials may be sustained over time, given the chronic, progressive nature of the condition.

Clinical advisors to the ERG advised that the availability of a treatment option for VN (were it to be approved) may lead to greater prevalence and incidence rates of RPE65-mediated IRD than estimated. This may have implications for the potential budget impact of VN.

3.4. Cost-Effectiveness Evidence Submitted by the Company

The company model adopted a state-transition cohort-level structure, comprising of five "alive" health states plus a sixth absorbing health state representing death (see Figure). A lifetime horizon (i.e. 85 years) was modelled, and discount rates of 3.5% for costs and outcomes were used in the company base case. Model health states were intended to capture the impact of VN on patients' HRQL, based on American Medical Association (AMA) visual impairment classes (see Table 1)[19]. Health state occupancy was determined by the worst of either VA or VF (as the company did not consider the novel MLMT outcome within the model structure).

Table 1 Health state descriptions included within the company model

Health state	Description	Worst of

		VA (LogMAR)	VF (degrees, ^{0})
HS1	Moderate VI	VA >1.0	240 < VF ≤ 360
HS2	Severe VI	1.0 ≤ VA < 1.4	144 < VF ≤ 240
HS3	Profound VI	1.4 ≤ VA < 1.8	48 < VF ≤ 144
HS4	CF	1.8 ≤ VA ≤ 3.0	0 < VF ≤ 48
HS5	HM, LP, NLP	VA < 3.0 or an indication of HM, LP, or NLP	-

Key: CF, counting fingers; HM, hand motion; HS, health state; LP, light perception; NLP, no light perception; VA, visual acuity; VF, visual field; VI, visual impairment.

Utility estimates were based on interviews with six clinicians, who were asked to complete proxy generic HRQL questionnaires for each health state based on summary descriptions and their experience with patients. Mean utility values in the company base case (using the HUI3) ranged from 0.52 (SD 0.16) for health state 1 to -0.04 (SD 0.07) for health state 5. A sensitivity analysis using EQ-5D values was also presented; however, the company noted limitations regarding the use of EQ-5D in measuring the impact of visual disorders on HRQL. Changes in health state occupancy estimated over the model time horizon were intended to reflect changes in HRQL following treatment with VN, as well as natural disease progression.

Patient transitions to 1 year were informed by observed proportions in Study 301[17], with both forward (worsening vision) and backward (improving vision) transitions permitted. After 1 year, only forward (i.e. worsening) transitions were permitted, informed by a combination of clinical expert opinion and a parametric multistate model (MSM) fitted to natural history data from the RPE65 NHx study[20]. The RPE65 NHx study is a retrospective chart review of 70 patients with RPE65-mediated IRD who would be eligible to receive VN. Patients had a mean age of 15 years at the start of data collection, and were followed up for a mean duration of 7.28 years. The treatment effect of VN was modelled across 3 phases: treatment maintenance, treatment waning, and residual treatment effect. The treatment effect of VN observed at 1 year was assumed to be maintained for 40 years following administration, following which a treatment waning of 75% was assumed to occur over the next 10 years, leading to a residual treatment effect of 25% for the remainder of the time horizon.

As no deaths were observed in the clinical trials, mortality in the model was based on adjusted general population life tables for England and Wales (adjusted for age, sex, and a health state-specific mortality effect from a study by Christ et al.[21]). Costs were based on published sources for the years 2017/2018, and included treatment acquisition, surgery, monitoring, resolution of AEs, and eligibility testing. Medical resource use was costed based on assumptions and input from clinical experts.

A cycle length of 1 year was applied with adjustments to 'twelfth-cycle' correct relevant costs and outcomes within the first year, after which half-cycle adjustments were made. The company base case takes the perspective of the National Health Service (NHS) and Personal Social Services (PSS), though within the CS no costs incurred by PSS are explicitly described. A confidential patient access scheme (PAS) discount to the NHS was incorporated in the company base case.

Results of the company's base-case analysis showed that the total cost of the VN strategy was much higher than that of BSC; however, the model suggested a difference of 7.06 quality-adjusted life-years (QALYs) between arms over the time horizon. The corresponding ICER for VN was reported to be

under the NICE HST threshold for approval (based on a weighted threshold of £200,000 per QALY gained) driven by the modelled lifelong improvement in vision, and with some cost offsets due to reduced resource use.

The company noted that VN may have cost implications beyond the NHS, including a reduction in costs to patients/carers, impact on work productivity, and on governmental spending on social security benefits. The company suggested that use of VN would have limited impact on healthcare staffing or infrastructure, due to the small number of patients eligible, and the case that VN is a 'one-off' treatment. As BSC is currently the only treatment option, the company assume that VN would receive a 100% market share, with all prevalent patients being treated in the first 5 years of VN availability.

Figure 1 Model schematic (re-drawn by the ERG)

--Figure 1 here--

Key: CF, counting fingers; HM, hand motion; LP, light perception; NLP, no light perception; VI, visual impairment.

3.5. Critique of Cost Effectiveness Evidence and Interpretation

The ERG considered that the cost-effectiveness evidence submitted by the company was consistent with the NICE reference case. Aside from the change in the target population, previously discussed, the evidence was also consistent with the NICE final scope[7]. The methods used in the company's SLR were also appropriate.

With regards to the company's base case model, the ERG agreed with the choice to adopt a cohort model, as this was consistent with the evidence available and the bilateral nature of the disease. The ERG also agreed with the approach of splitting the model into initial and long-term phases, and for decisions regarding forward and backward transitions. However, the ERG was concerned that there were insufficient data available to populate the model transitions in both phases. For example, in Year 1 there were 20 possible transitions for patients receiving BSC; the clinical evidence from Study 301[17] comprised of only 9 patients. After Year 1, due to the number of model health states, the MSM analysis required the estimation of 11 parameters, with the RPE65 NHx study[20] only able to inform each parameter with data from an average of 6.2 patients (n=68 patients for 11 parameters), equivalent to an average of 3.2 transitions (events) per parameter (n=35 transitions for 11 parameters). The ERG was therefore concerned that the model was overly complex for the available data, and likely 'over fitted' the data which had important implications for extrapolation. While reducing the sensitivity of the model to changes in vision over time, the ERG considered that a simpler model with fewer health states would have led to a more robust estimation of transitions. Related, the ERG noted that the incorporation of data from Study 302[18] would have increased the amount of data

available to inform transitions, and was unclear as to why these data were not leveraged by the company (acknowledging that Study 302 is an open-label extension to Study 301).

The ERG was particularly concerned with the validity of utility estimates used to represent health states. The specification of a negative utility value for Health State 5 (HS5) (analogous to a state worse than death) did not match with evidence of the burden of RPE65-mediated IRD, nor with clinical advice. In general, patients did not have other health problems, and continued to perform their usual activities, modifying these over time. Even with extremely poor vision, patients were described as leading meaningful lives with high levels of enjoyment. Moreover, the ERG identified that of 17 previous NICE appraisals of technologies for visual impairments[22-38], the lowest utility estimates were between 0.26 and 0.55, the majority of which included health states described as “blind”.

The ERG agreed that a lifetime horizon was appropriate in this condition, and that the RPE65 NHx study[20] was the best source of available evidence to inform long-term outcomes despite the limitations with the retrospective design. There were however some differences in baseline participant characteristics (including in the distribution of health state allocation), with the RPE65 NHx study representing a less severe group of patients than those included in Study 301/302[17, 18].

The ERG’s main concern was the validity of assumptions for the long-term benefits of VN. Extrapolations beyond the follow-up of the RPE65 NHx[20] did not appear to have been validated based on clinical plausibility, and were inconsistent with statements about long-term progression of the disease made by the company in its submission – for example, the CS stated “Patient burden is very high in this severe, progressive and extremely rare disease, with patients inexorably progressing to near-total blindness as early as the preschool years or as late as the third decade of life.” yet health state occupancy at 15 years (at which point the cohort would have a mean age of approximately 30 years) did not demonstrate this.

The ERG also did not consider there to be sufficient data to support the assumed maintenance and waning of treatment effect of VN. The company stated that assumptions on long-term outcomes were based on interviews with UK clinical experts (though no formal transcripts were recorded), while the 25% threshold for treatment waning was an arbitrary figure. Clinical experts to the ERG advised that assumptions of a general long-term treatment effect of VN were plausible based on the available data; however, whether the treatment effect would really persist over a patient’s lifetime remains unknown (and for the time being, unknowable). The ERG also noted evidence showing that the treatment effect for a different gene therapy was not necessarily lifelong for all patients[39].

Generally, the ERG considered that measures of functional vision would have been preferable for defining outputs in the model, as advice to the ERG was that VA and VF are not considered to best represent the nature of the condition. Furthermore, the ERG considered that the incorporation of patient-reported outcomes, including directly elicited HRQL, would have greatly reduced uncertainty in the results. The ERG were broadly in agreement with disutilities and costs applied to the company’s model.

The ERG considered that an annual discount rate of 3.5% for costs and outcomes was aligned with the NICE reference case, and was best supported by the available evidence. While VN is anticipated to lead to benefits extending beyond 30 years (potentially allowing for lower discount rates), this is as yet unproven, and VN comprises a technology that commits the NHS to significant, irrecoverable costs as VN is a 'one-off' gene therapy with uncertain long-term effectiveness. For this reason, lower discount rates were not used in the ERG's preferred base-case analysis as the criteria for doing so were not deemed to have been fully met.

The ERG acknowledged that patients with RPE65-mediated IRD will be in receipt of governmental spending and out of pocket expenses incurred by themselves and their care givers. However, the ERG considered that the expenditure estimated by the company may have been overstated. Moreover, a sensitivity analysis conducted by the ERG indicated that incorporating government spending had little impact on the ICER. The ERG were sceptical that patients eligible for VN would wait up to 5 years for treatment, and calculated that a higher number of patients treated earlier would cause expenditure on VN to exceed £20 million of sales in its first year of availability (at the list price of VN of £600,000, £20 million in sales would be reached if 34 patients were treated). Epidemiological estimates and input from clinical advisors to the ERG suggested that ultimately, given the rarity of the condition and requirements for treatment, there is substantial uncertainty in how many patients would be eligible for treatment with VN.

3.6. Additional Work Undertaken by the ERG

Based on the ERG's critique of the CS, the ERG made adjustments to the company's base-case model. This included corrections to minor computational and data errors, and the incorporation of the ERG's preferred assumptions. The ERG also performed additional exploratory and sensitivity analyses.

In the ERG's preferred base case, the largest impact on the ICER was due to the use of alternative utility estimates for the model health states. These estimates were derived from Rentz et al. (2014)[40], which was a time trade-off study with 607 members of the general public (in Australia, Canada, the UK, and US) that estimated utilities for 8 health states with varying degrees of vision problems. Mean estimates ranged from 0.956 (health state; no problems) to 0.343 (substantial limitations due to lack of vision). Additional changes included incorporating data from Study 302[18] and RPE65 NHx[20] to population transitions, the removal of the waning and residual effect phases, alterations to medical resource costs and costs for resolving AEs, the removal of mortality multipliers, the removal of carer disutility for adults, and a change in the carer disutility value for children. The ERG also conducted sensitivity analyses around uncertainties in the model, including a threshold analysis to determine the effect the assumed duration of treatment effect on the ICER.

The ERG's preferred base case reported a higher ICER than the company base case, which exceeded the NICE HST threshold for approval.

3.7. Conclusions of the ERG Report

The ERG accepted that evidence for disease burden in a broader population of patients with IRD was useful, although there was uncertainty in the accuracy of generalising this to patients with RPE65-mediated IRD. The ERG generally agreed with the proposed use of VN in the clinical pathway.

Evidence from the included studies appears to show that VN is associated with a modest and sustained improvement in visual outcomes in patients with RPE65-mediated visual impairment at 3/4 years. This includes a clinically meaningful benefit in patients' visual function; though no HRQL trial data were provided, leading to considerable uncertainty in the overall benefits of VN. VN demonstrated a good safety profile, with few serious AEs and no deaths observed in the trials. While limited evidence was available, this is common with rare diseases, and the availability of an RCT augmented the evidence base.

Nevertheless, the company's economic model was heavily informed by clinical expert opinion, with a high degree of uncertainty in multiple areas. In particular, the following three factors contributed to uncertainty in the model findings: (1) the assumed treatment effect of VN which is at present extremely unclear, (2) the modelling of long-term natural history outcomes, and (3) the utility estimates. The duration of treatment effect and estimation of utility values required extensive clinical expert input to inform the model base case and did not have face validity in the view of the ERG. The long-term natural history of RPE65-mediated IRD was modelled using a highly complex MSM which is subject to palpable uncertainty.

Consequently, the ERG was unconvinced that the assumptions relating to the long-term effects of VN were supported by the available evidence. The ERG also noted that should the treatment effect fail to remain at 100% for at least 35 years, the ICER rises precipitously; yet, due to discounting, should the duration of effect be longer than the company's base-case analysis, this made little difference to the ICER.

4. Key Methodological Issues

Due to the lack of evidence in this disease area, the company's economic model relied heavily on the input of clinicians to estimate model inputs; including parameters beyond their personal experience (such as the duration of effect of gene therapy). Scenario analyses found that these inputs impacted greatly on the ICER, and were therefore a major cause of uncertainty in the model. The process used by the company to elicit clinician input was not clearly documented, and therefore the rigor of the process used was unclear. Without an evidence base to corroborate clinical judgements, it was not possible to determine their reliability.

This scenario is likely to be common across rare diseases, where the evidence base will naturally be small, and therefore it would be useful for methodologists to consider the best way to incorporate the experience of clinicians, patients and carers in representing patient experience and assumptions of biological plausibility. More formal and explicit elicitation methods such as the SHeffield ELicitation Framework (SHELF)[41] may be especially useful in this context.

As is also common in the context of rare diseases, evidence for this appraisal was supported by small trials, which were unable to adequately populate all possible transitions within the company's economic model. Similar appraisals may therefore need to consider striking a balance between a model structure that has face validity for the patient experience, and simpler model structures that can be reliably parameterised from the available evidence. The conduct of an RCT by the company reduced the discussion to the magnitude of treatment effect (and not whether an effect exists), and should therefore be commended.

5. National Institute for Health and Care Excellence Guidance

5.1. Consideration of Clinical Effectiveness

The committee considered the evidence base to be applicable, though noted that the evidence was only in patients with leber congenital amaurosis (LCA) – one specific type of REP65-mediated IRD[42]. However, the therapy treats the underlying cause of the condition so, biologically, the clinical diagnosis is unlikely to impact treatment effect. The committee's clinical experts stated that the relevance of the study results to clinical practice was difficult to predict in patients with the less severe diagnoses, such as retinitis pigmentosa (RP).

While the committee were concerned with the reliability and validity of traditional visual performance markers (VA, VF, and contrast sensitivity) in this population, they considered the evidence to show that VN improved vision in the short term. Clinical experts to the committee explained that even small changes in vision would be important to patients and, as shown in MLMT and VFQ data, may have a substantial impact on mobility and functional vision. The committee further considered the evidence to show that VN may prevent the deterioration of vision, which patient experts explained would be important for quality of life. The committee considered that improvements seen in the modified VFQ were likely to be clinically meaningful, though noted that some improvement over time would be expected as patients adjust to their surroundings over time. The committee noted that having no direct measure of HRQL was a key limitation of the company's evidence base.

While there are no long-term data on the effectiveness of VN, the committee concluded that there was a biological rationale for the treatment effect of VN to be maintained. Clinical experts advised that, following successful delivery, the healthy RPE65 gene should continue to express indefinitely and will continue to restore vision. While the committee noted evidence that measures of VF and VA showed some deterioration between 3 and 4 years[18], this was in a small number of patients, and improvements remained above the MCID from baseline. Clinical experts explained that it is possible for vision to continue to deteriorate if some photoreceptor cells outside the area of injection die. They also noted that vision deteriorates as people age, both in the general public and for those with the condition, and that this is not a reflection of treatment failure.

Based on the evidence, the committee concluded that VN had an acceptable safety profile, with the majority of AEs occurring due to the administration of VN, which is a one-off event and may decrease with experience administering VN (in a small number of specialist centres) over time. The company

confirmed that, as part of the implementation period, all healthcare professionals (including pharmacists) have mandatory training. The committee judged that sufficient viable retinal cells was not fully defined in the marketing authorisation. The committee concluded that clinical judgement incorporating both structural assessment, based on that used in Study 301/302[17, 18], and functional assessment would be used in clinical practice to identify patients eligible for treatment.

5.2. Consideration of Cost Effectiveness

The committee considered that the structure of the company's model, including assumptions for transitions in the model, was appropriate for decision making. The committee also agreed with the specification of health states in the model, which they were advised would be able to represent a prevention in the deterioration of vision that would be of importance to patients. However, the committee noted concerns from the ERG of the low sample size used to populate the model, and agreed with the ERG's preference to incorporate data from Study 302[18] and RPE65 NHx[20] to inform model transitions. The committee noted a lack of face validity in utility estimates made for each of the health states in the company's model, having heard from patient experts that vision loss represented by HS5 is unlikely to be equivalent to a state worse than death. However, the committee was uncertain whether the ERG's preferred method for matching utilities from Rentz et al.[40] with AMA health states[19] would be better than the method used by the company that used clinician judgement. In the absence of further evidence, the committee considered that utility estimates for the model health states fell between the ERG's preferred estimations and the EQ-5D estimations provided by the company.

The committee noted the lack of data available to support the long-term extrapolations in the company's model, and considered that this led to considerable uncertainty in the corresponding results. However, clinical experts advised that patients generally experience deterioration in vision that could mirror the company's model. The committee agreed that the results of the analysis are likely to vary substantially depending on the expected duration of treatment effect. It concluded that, in the absence of any long-term evidence but given the biological plausibility for long-term treatment effect, a long-term treatment effect of 40 years' duration was uncertain but reasonable. However, the committee agreed with the ERG in respect of the removal of assumptions of treatment waning and residual treatment effect, which were not based on evidence or a biological rationale.

The committee did not consider there to be sufficient evidence to support the use of a 1.5% discount rate for curative technologies, as experts advised that treatment may be unlikely to fully resolve vision problems if photoreceptor cells have already been damaged, or if is not applied to all cells (and so patients who receive VN may still have lifelong visual impairment). The committee was highly uncertain about whether patients who had VN would be considered to have 'normal or near-normal health'. It also recognised that there were large uncertainties about whether the long-term benefits of treatment would be achieved because of the limited evidence. However, the committee decided to consider both a 3.5% and 1.5% discount rate in its decision-making, while giving greater weight to the 3.5% rate.

Overall, the committee acknowledged the uncertainties in utility estimates and treatment duration. However, it was convinced that clinical benefits of VN were important, and an appropriate use of NHS resources within the context of a specialised service. The committee concluded that there was a high unmet need in patients with RPE65-mediated IRD, and that VN is a step change in its treatment.

Based on the committee's preferred assumptions, and using the company list price and a discount rate of 3.5%, the ICER remained above the NICE threshold for approval. Considering both methods of utility estimation, the ICER ranged between £114,956 (EQ-5D) and £155,750 (Rentz et al.[40]) per QALY gained. The ICER fell below the NICE threshold when using a 1.5% discount rate. The committee considered that the submission met criteria for applying a QALY weighting of 1.2 to the results (the acceptable threshold in HST being based on the magnitude of health gain), based on VN being associated with a QALY gain between 12.1 and 17.7.The committee considered the potential impacts of VN beyond direct health benefits, and discussed the large emotional impact of VN on families and carers, and the substantial financial burden on families and carers. The committee also heard from patient experts that improvements in vision that could permit children to attend mainstream school could affect the course of their lives. The committee noted that VN could lead to reduced expenditure in non-NHS government departments that provide support for families affected by vision impairment, although considered that inclusion of these costs would be unlikely to have a meaningful impact on the magnitude of the ICER.

6. Conclusions

VN is a first-in-class gene therapy for patients with RPE65-mediated IRD where no other treatment exists. The evidence presented suggested that VN improves vision, and input from expert clinicians suggested that VN may have the potential to prevent the deterioration of vision over the long-term. This means that VN has the potential to meaningfully improve the quality of life of patients with RPE65-mediated IRD, as well as their families and carers. However, the paucity of evidence available means that there are substantial omissions in the company's evidence base demonstrating the clinical and cost effectiveness of VN. Where gaps were present, the subsequent uncertainties in the company's economic model have the potential to drastically increase the ICER. This appraisal highlights challenges in the appraisal and approval of interventions within the context of rare disease.

7. References

1. Daiger, S.P., L.S. Sullivan, and S.J. Bowne, *Genes and mutations causing retinitis pigmentosa.* Clin Genet, 2013. **84**(2): p. 132-41.
2. Hamel, C.P., *Gene discovery and prevalence in inherited retinal dystrophies.* C R Biol, 2014. **337**(3): p. 160-6.
3. Ripamonti, C., et al., *Nature of the Visual Loss in Observers With Leber's Congenital Amaurosis Caused by Specific Mutations in RPE65.* Investigative Ophthalmology & Visual Science, 2014. **55**(10): p. 6817-6828.
4. European Medicines Agency, *Luxturna: EPAR - Medicine overview*. 2019.
5. Rakoczy, E.P. and K. Narfström, *Gene therapy for eye as regenerative medicine? Lessons from RPE65 gene therapy for Leber's Congenital Amaurosis.* Int J Biochem Cell Biol, 2014. **56**: p. 153-7.
6. National Institute for Health and Care Excellence. *Voretigene neparvovec for treating inherited retinal dystrophies caused by RPE65 gene mutations Highly specialised technologies guidance [HST11].* Available from: https://www.nice.org.uk/guidance/hst11 (accessed February 2020).
7. National Institute for Health and Care Excellence. *Final scope for the evaluation of voretigene neparvovec for treating inherited retinal dystrophies caused by RPE65 gene mutations*. 2018.
8. Chung, D.C., et al., *Novel mobility test to assess functional vision in patients with inherited retinal dystrophies.* Clin Exp Ophthalmol, 2018. **46**(3): p. 247-259.
9. Mangione, C.M., et al., *Development of the 25-list-item National Eye Institute Visual Function Questionnaire.* Archives of Ophthalmology, 2001. **119**(7): p. 1050-1058.
10. Novartis, *(Draft) Psychometric evaluation of the modified Visual Function Questionnaire (VFQ-25) using data from AAV2-hRPE65v2-301 (voretigene neparvovec) phase III trial*. 2019.
11. Francis, P.J., *Genetics of inherited retinal disease.* J R Soc Med, 2006. **99**(4): p. 189-91.
12. Sonksen, P.M. and N. Dale, *Visual impairment in infancy: impact on neurodevelopmental and neurobiological processes.* Dev Med Child Neurol, 2002. **44**(11): p. 782-91.
13. RNIB, *Children and young people England : RNIB Evidence-based review*. 2016.
14. Chaumet-Riffaud, A.E., et al., *Impact of Retinitis Pigmentosa on Quality of Life, Mental Health, and Employment Among Young Adults.* American Journal of Ophthalmology, 2017. **177**: p. 169-174.
15. Maguire, A.M., et al., *Age-dependent effects of RPE65 gene therapy for Leber's congenital amaurosis: a phase 1 dose-escalation trial.* Lancet (London, England), 2009. **374**(9701): p. 1597-1605.
16. Bennett, J., et al., *Safety and durability of effect of contralateral-eye administration of AAV2 gene therapy in patients with childhood-onset blindness caused by RPE65 mutations: a follow-on phase 1 trial.* Lancet, 2016. **388**(10045): p. 661-72.

17. Russell, S., et al., *Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in patients with RPE65-mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial.* Lancet (London, England), 2017. **390**(10097): p. 849-860.
18. Russell, S., *Three-year update for the phase 3 voretigene neparvovec study in biallelic RPE65mutation–associated inherited retinal disease*, in *Association of Research in Vision and Ophthalmology (ARVO) annual meeting*. 2018.
19. Rondinelli, R.D., *AMA Guides to the Evaluation of Permanent Impairment*. 6th ed. 2007: American Medical Association.
20. Chung, D.C., et al., *The Natural History of Inherited Retinal Dystrophy Due to Biallelic Mutations in the RPE65 Gene.* Am J Ophthalmol, 2019. **199**: p. 58-70.
21. Christ, S.L., et al., *Longitudinal relationships among visual acuity, daily functional status, and mortality: the Salisbury Eye Evaluation Study.* JAMA Ophthalmol, 2014. **132**(12): p. 1400-6.
22. National Institute for Health and Care Excellence, *TA155: Ranibizumab and pegaptanib for the treatment of age-related macular degeneration*. 2008.
23. National Institute for Health and Care Excellence, *TA229: Dexamethasone intravitreal implant for the treatment of macular oedema secondary to retinal vein occlusion.* 2011.
24. National Institute for Health and Care Excellence, *TA274: Ranibizumab for treating diabetic macular oedema.* 2013.
25. National Institute for Health and Care Excellence, *TA283: Ranibizumab for treating visual impairment caused by macular oedema secondary to retinal vein occlusion.* 2013.
26. National Institute for Health and Care Excellence, *TA294: Aflibercept solution for injection for treating wet age-related macular degeneration*. 2013.
27. National Institute for Health and Care Excellence, *TA297: Ocriplasmin for treating vitreomacular traction.* 2013.
28. National Institute for Health and Care Excellence, *TA298: Ranibizumab for treating choroidal neovascularisation associated with pathological myopia.* 2013.
29. National Institute for Health and Care Excellence, *TA301: Fluocinolone acetonide intravitreal implant for treating chronic diabetic macular oedema after an inadequate response to prior therapy.* 2013.
30. National Institute for Health and Care Excellence, *TA305: Aflibercept for treating visual impairment caused by macular oedema secondary to central retinal vein occlusion.* 2014.
31. National Institute for Health and Care Excellence, *TA346: Aflibercept for treating diabetic macular oedema.* 2015.
32. National Institute for Health and Care Excellence, *TA349: Dexamethasone intravitreal implant for treating diabetic macular oedema.* 2015.
33. National Institute for Health and Care Excellence, *TA369: Ciclosporin for treating dry eye disease that has not improved despite treatment with artificial tears.* 2015.
34. National Institute for Health and Care Excellence, *TA409: Aflibercept for treating visual impairment caused by macular oedema after branch retinal vein occlusion.* 2016.

35. National Institute for Health and Care Excellence, *TA460: Adalimumab and dexamethasone for treating non-infectious uveitis.* 2017.
36. National Institute for Health and Care Excellence, *TA467: Holoclar for treating limbal stem cell deficiency after eye burns.* 2017.
37. National Institute for Health and Care Excellence, *TA486: Aflibercept for treating choroidal neovascularisation.* 2017.
38. National Institute for Health and Care Excellence, *TA532: Cenegermin for treating neurotrophic keratitis.* 2018.
39. Bainbridge, J.W.B., et al., *Long-Term Effect of Gene Therapy on Leber's Congenital Amaurosis.* New England Journal of Medicine, 2015. **372**(20): p. 1887-1897.
40. Rentz, A.M., et al., *Development of a preference-based index from the National Eye Institute Visual Function Questionnaire-25.* JAMA Ophthalmol, 2014. **132**(3): p. 310-8.
41. Oakley, J., *SHELF-package: Tools to Support the Sheffield Elicitation Framework*. 2017: https://cran.r-project.org/package=SHELF. p. R package.
42. Hanein, S., et al., *Leber congenital amaurosis: comprehensive survey of the genetic heterogeneity, refinement of the clinical definition, and genotype-phenotype correlations as a strategy for molecular diagnosis.* Hum Mutat, 2004. **23**(4): p. 306-17.