Erratum to:

Voretigene neparvovec for treating inherited retinal dystrophies caused by RPE65 gene mutations [ID1054]

Please note that a list of replacement pages with summary of and rationale for corrections is provided below.

Section, page (line) reference (vs original report and issues identified in the company fact check	Summary rationale
Summary, pages 15-21	Page 15 (line 29): Removed "Spark Therapeutics Inc." as this was noted by the company to be incorrect Page 15 (line 28): Clarified that a minority of patients would not be eligible for treatment due to participation in other clinical trials. Page 16 (lines 13-15, & 20-21): Updated wording to align with the marketing authorisation and to highlight that the decision problem was in line with the marketing authorisation. Page 17 (lines 18 & 20): Clarification of wording in respect of the effect of VN on contrast sensitivity; i.e. "change from baseline" and correction to state that the differences at 1 year were "above" the company's defined threshold for clinical significance. Page 17 (line 26): Replaced references to "lux" or "lux units" or "lux levels" with "light levels" as the wording use was inaccurate. Page 18 (lines 2-6): Text was adjusted to correct the adverse events associated with the administration of VN and the tmiepoint at which these were assessed. Page 18 (line 12): Text was adjusted to clarify that all patients were treated in the US Page 19 (line 19): The text was amended to mention the overall good safety profile of VN. Page 20 (lines 30-31): Costs were replaced to include the PAS discount Page 21 (lines 7-8): Text was aligned with the marketing authorisation
Section 2, p.26 (lines 9-10)	Clarified that a minority of patients would not be eligible for treatment due to participation in other clinical trials.

Section, page (line) reference (vs original report and issues identified in the company fact check	Summary rationale
Section 2, p.30 (lines 15-16)	Deleted “however this was not reported in the CS”. Nystagmus is discussed several times in the CS.
Section 2, p.32 (lines 12-13)	Sentence “Standard practice for the diagnosis of *RPE65*-mediated IRD was not reported in the company submission.” was deleted, and the following sentence was added: “The clinical pathway of care including details of diagnosis is provided in Section 8.2 of the CS.”
Section 2, p.36 (lines 6-8)	Added that the restriction in the population is in line with the marketing authorisation, which is the remit of the evaluation.
Section 3, p.39 (Population row)	Clarified that the population included in the submission is narrower than that specified in the NICE scope.
Section 3, p.43 (11-12)	The sentence was unclear and has been amended to provide clarification.
Section 4, p.54 (line 4) and p.55 (lines 24-28)	This sentence has been amended to clarify that all patients in the trial were treated in the US (p.54). In addition, text was amended to clarify that baseline VA was higher in patients treated in Study 301/302 than Study 101/102.
Section4, p.60 (line 6)	Corrected typographical error: “lose” to “low”
Section 4, p.64 (line 31) to	Added ***NEI*** VFQ. The full name of the questionnaire was needed to avoid confusion with the VFQ used in the trial.
Section , p.65 (line17 & line 24)	Corrected typographical error for clarity: “considered” to “consideration” Replaced references to “lux” or “lux units” or “lux levels” with “light
Section 4, p.69	Changed from ***MLMT*** to ***Mobility testing.*** Replaced references to “lux” or “lux units” or “lux levels” with “light levels” as the wording use was inaccurate.
Section 4, p.84 (lines 11-14)	Corrected to: “These changes were ***above*** the company’s defined threshold for clinical significance (≥***1*** log unit).” Corrected figure caption (Figure 15): Study 302 Full-Field Light Sensitivity Threshold at ***3*** year***s***
Section 4, p.86 (lines 10-19)	Revised summary of results and conclusions to clarify that FST results did exceed clinical significance thresholds.
Section 4, p.90 (row 15)	Replaced references to “lux” or “lux units” or “lux levels” with “light levels” as the wording use was inaccurate.
Section 4, p.91 (line 17)	Replaced “mean difference” (page 91) and “mean (SD) change from baseline” (page 94) with “mean difference in change from baseline”. Clarification was required that these numbers represent the mean difference in change from baseline between the arms, not the mean difference in scores between arms.

Section, page (line) reference (vs original report and issues identified in the company fact check	Summary rationale
Section 4, p.95 (lines 3-5)	Deleted the statement: "patients in the BSC arm could be considered to have experienced a clinically meaningful increase in VN scores in the 1st year of the trial (mean change 0.8 on both patient- and parent-reported scales)." The sentence was not accurate.
Section 4, page 95 (lines 9-12)	Corrected "[redacted]".
Section 4, p.99 (lines 8-9)	Replaced "It was not reported whether the TEAE was considered to be related to the administration of VN" with "which was related to the administration of VN" with "One SAE was reported, which was considered unlikely to be related to the study drug, but that resulted from treatment given for a previous TEAE (intraocular inflammation endophthalmitis), which was considered to be related to the administration procedure." This was also clarified in the ERG comment (p.103)
Section 4, p.103	See response to Section 4 (p99 (lines 8-9)) for context of change and rationale
Section 4, p.109 (line 13)	The CS repeatedly states that no TEAEs were associated with VN (p. 122, 126, 127) and the ERG consider this new information to be a significant omission. Following this further clarification from the company, the report has been adjusted to reflect this new information. This has required edits to multiple sections of the report where the ERG noted the apparent absence of TEAEs related to VN.
Section 5, p.113	Replaced "The population is broader" with "The population is narrower". The original wording was not correct.
Section 5, p.120 (line 7)	Added that the restriction in the population is in line with the marketing authorisation, which is the remit of the evaluation.
Section 5, p.129 (line 12)	Replaced "observed" with "reported". "The natural history study was a non-interventional retrospective chart review so data on deaths were not recorded."
Section 5, p.138 (lines 14-16)	The ERG notes their use of the phrase "will" may be unjust, and so has revised the report to say that this "may" introduce a framing effect, however overall the statement is not considered to be factually inaccurate.
Section 5, page 138-139 (line 32 page 138 to line 1 page 139)	Revised the wording in the statement "the ERG's clinical advisors stated that patients had restrictions imposed by their vision, but in general did not have other health problems" to contextualise the discussion.
Section 5 page 138 (lines 30-32)	The wording in the report has been revised to state "the absolute values derived via the proxy elicitation exercise..." and "the negative value for HS5 (HUI3 analysis)"

Section, page (line) reference (vs original report and issues identified in the company fact check	Summary rationale
Section 5, p.143 (lines 11–12)	The following sentence was deleted: "This study was not identified in the company's systematic literature review, as the review was targeted specifically at *RPE65* mediated vision loss." The Rentz study was captured in the company's systematic review but was excluded.
Section 5, p.178 (bottom two rows of Table 63)	The results presented for *Baseline characteristics derived from RPE65 NHx* were updated per the ERG model.
Section 8, p.185 (line 25)	"…patients with ***hypomorphic alleles giving a later less severe recessive phenotype who may have a different prognosis from the typical patient***.[100] ~~a dominant allele giving rise to a very different phenotype that may have a different prognosis from the typical patient.[101]~~ ~~Similarly, t~~There are ***also*** a subgroup of patients ***with a dominant allele giving rise to a very different phenotype that may have a different prognosis from the typical patient,[101] although these patients are not eligible for VN under its current marketing authorisation***. ~~hypomorphic alleles giving a later less severe recessive phenotype who may have a different prognosis from the typical patient.[100]~~
Section, p. 186 (line 12)	The clinician stated that impact of V~~A~~***N*** on delivery

Summary p. 15-21

1. SUMMARY

1.1 Background

Inherited retinal dystrophies (IRD) are a heterogeneous group of rare diseases caused by germline mutations in more than 260 genes, including the *RPE65* gene. The key outcome of *RPE65*-mediated IRD is inexorable and progressive loss of vision, culminating in near or total blindness, though the rate of deterioration varies considerably between patients. The pathophysiology underlying progressive loss of vision relates to the inability to complete the visual cycle because of deficiencies in the *RPE65* enzyme. Deficiencies in this enzyme arrest the molecular pathways that culminate in transmission of signals to the brain. In addition, the accumulation of toxic precursors in the visual cycle leads to apoptosis, or cell death, in photoreceptor cells. IRD is often diagnosed in infancy and adolescence. Night blindness is a common first symptom, but in infants, the 'oculo-digital sign', or eye poking, is a common presentation, though its association with *RPE65*-mediated IRDs is unclear. *RPE65*-mediated IRD is an autosomal recessive-transmitted disorder, including two related disorders; retinitis pigmentosa (RP) and Leber congenital amaurosis (LCA).

The impact of the condition begins in early life, with impacts on child social development arising from poor visual function. Adults may face decreased employment opportunities arising from challenges in accessing education. IRD also impacts carers and household members through increased caring burden, and is associated with an increased risk of depression among patients and their family members. The ERG noted that while evidence presented for these impacts drew from IRD generally, there was no evidence specific to *RPE65*-mediated IRD.

Diagnosis of *RPE65*-mediated IRD includes medical history and genetic testing. The company estimated that only 50% of people with the disease are currently diagnosed. Care for this condition is at present primarily supportive, and few national or expert guidelines exist. For children, visual aids and magnifiers are recommended, as well as supportive resources in school settings (e.g. specially qualified teachers).

While the ERG noted that the evidence related to incidence and prevalence of the condition is scant and thus any estimate is highly uncertain, the company estimated that the prevalence of IRD mediated by the *RPE65* gene would lead to a population of 86 patients in the UK although a minority of these patients would not be eligible for treatment due to participation in other clinical trials.

Voretigene neparvovec (VN; Luxturna®; Novartis Pharmaceuticals (UK).) is an adeno-associated virus (AAV) gene therapy treatment which introduces a healthy copy of the defective *RPE65* gene into the retinal cells of patients with *RPE65*-mediated IRD. VN is administered as two subretinal injections (no fewer than six days apart) once per lifetime. Prior to administration (approximately 3 day before), patients are required to receive an immunomodulatory regimen (such as prednisone), which is expected to be continued for a further 18 and 30 days, depending on the timing of the second administration (i.e. the other eye). The introduction of *RPE65* enables patients to produce functional *RPE65* protein. The subretinal injection of VN introduces a healthy copy of the defective *RPE65* gene into retinal cells. This enables patients to produce functional *RPE65*, resulting in improved functional vision (improved ability to perform vision-related daily activities) and visual function (improved performance of the eyes at the organ level). In order to derive benefit from VN treatment, the company states that patients must have confirmed biallelic (pertaining to both paternal and maternal alleles) *RPE65* mutations and have sufficient viable retinal cells into which healthy copies of the *RPE65* gene can be introduced.

VN is not currently used in the UK for any patient population. The European Medicines Agency (EMA) awarded VN marketing authorisation on 22 November 2018. VN is expected to be used in line with the marketing authorisation for the treatment of adult and paediatric patients with vision loss due to IRD caused by confirmed biallelic *RPE65* mutations and who have sufficient viable retinal cells.

1.2 Critique of the decision problem submitted by the company

The decision problem included in the company submission broadly adhered to the final NICE scope. The ERG noted that the company restricted the population of patients from those with *RPE65*-mediated IRD to include only those who additionally possessed sufficient viable retinal cells. The ERG regarded that this was clinically justified and was in line with the marketing authorisation for VN. The intervention as specified matched the NICE scope, but the ERG noted that comparators, broadly classes as best supportive care, were not defined in the company submission.

Outcomes presented by the company included the multi-luminance mobility test (MLMT), which was not in the scope but described by the company as a clinically relevant test of functional vision. The MLMT was the primary endpoint of the company's pivotal phase 3 trial. While most other scoped outcomes were reported in the CS, the ERG noted that health-related quality of life data were not collected as part of the phase 3 trial, nor were data reported relating to need for cataract surgery.

Finally, the company used an economic perspective in their evaluation in line with the NICE scope.

1.3 Summary of clinical effectiveness evidence submitted by the company

The company presented a systematic review that included evidence from two trials. The pivotal trial for the submission is Study 301/302; an open-label, multi-centre, phase 3 RCT involving 31 patients (Study 301), followed by an optional phase after one year where 9/10 (90%) patients from the control arm received VN (Study 302). Patients were recruited from multiple countries worldwide, and travelled to sites in the US for treatment administration and follow-up. Study 301/302 is ongoing: data up to and including a four-year follow-up was available for some, though not all, outcomes in this submission. Study 101/102 is an open-label, phase 1, single-arm trial. Study 101 employed a dose-ranging design; with patients receiving either a 'low', 'medium', or 'high' dose of VN in a single (worse, non-preferred) eye. Patients travelled to sites in the US for treatment administration, following which 7/12 (58.3%) were followed up in the US, and 5/12 (41.7%) were followed up in Italy. After a minimum of 1 year, patients from Study 101 were invited to receive VN in the contralateral eye: 11/12 (91.7%) patients from Study 101 were eligible for entry into Study 102. All patients in Study 102 received a 'high' dose of VN in their contralateral (better, preferred) eye.

Primary visual acuity (VA) outcomes in Study 301/302 did not demonstrate a significant difference in changes from baseline to 1 year between VN and BSC (0.16 LogMAR, 95% CI [-0.41, 0.08]; p=0.17). All changes in VA were under the company's definition of a clinically meaningful change (≥0.3 LogMAR). Study 101/102 had similar findings. In contrast, VF improved in VN patients as compared to BSC patients at 1 year in Study 301/302. Improvements in VF were demonstrated by 30 days in the VN arm, and these remained relatively stable until 1 year (assessed by Goldmann III4e, MD 378.7; 95% CI [145.5, 612.0]; post-hoc p = 0.0059). Despite numerical evidence of decline after the 2 year timepoint, clinical advice received by the ERG suggested changes from baseline were clinically meaningful. In Study 301/302, [redacted], but differences at 1 year in photosensitivity were significant and above the company's defined threshold for clinical significance (full-field light sensitivity MD -2.11 log units; 95% CI [-3.91, -1.04]; p=0.0004), which were sustained at 3 years following administration (2 years in the delayed treatment arm). The company also presented evidence for the MLMT, which suggested sharp and sustained improvement after administration in both the VN and BSC (delayed VN) arms through 3 years after administration (2 years in the delayed arm); at 1 year before the BSC arm patients received VN, the mean difference in light units was 2.0 (95% CI [1.14, 2.85]). Finally, patient-reported outcomes including a modified Visual Function Questionnaire (VFQ) were reported for Study 301/302. [redacted] No health-related quality of life nor cataract surgery data were reported.

With regard to common adverse events attributed by the company to administration procedure, in the short term (one year), [redacted]

1.4 Summary of the ERG's critique of clinical effectiveness evidence submitted

The ERG regarded that the quality of methods used to locate the evidence was reasonable, though the use of unconventional search methods meant that there was a small, albeit unlikely, chance that studies may have been missed.

The pivotal phase 3 trial submitted, Study 301/302, generally matched the decision problem. Though all patients in this study were treated in the US, the ERG considered that the setting would generalise to UK practice. Of note is that inclusion and exclusion criteria for Study 301/302 were narrower than included in the NICE scope, given the study's requirement for sufficient viable retinal cells. The ERG considered that this was a clinically relevant consideration. However, the ERG noted that this additional criterion means it likely that there will be some patients included in the population specified in the NICE scope who will be excluded for treatment with VN because they have no viable retina to treat.

The small sample size in Study 301/302 (n=29; following the exclusion of 2 patients following randomisation) introduced uncertainty in the estimation of treatment effect. [redacted] While age differences were noted between the randomised groups at baseline, clinical advice suggested that there is no clear relationship between outcomes and age within an *RPE65*-mediated IRD population. Retinal function at baseline was suggested to be a potentially stronger mediator of treatment response, which may be partially correlated with age. However, none of the differences at baseline were considered by the ERG to demonstrate a clear bias in any direction, although it was noted that only a small number of characteristics were reported at baseline.

The ERG regarded that outcome assessment was generally appropriate and clinically relevant in this population, and that statistical methods used to analyse outcome data were acceptable. However,

measurement of VA, VF and contrast sensitivity is widely considered to be unreliable, and some imprecision in their measurement should be expected. In addition, the ceiling effect inherent to the MLMT measure may underestimate the treatment effect reported for continuous data. The removal of HRQoL data from the VFQ suggested to the ERG that the VFQ was not an appropriate measure of HRQoL. No HRQoL data, or PRO data for the carers of patients with IRD, was reported in the CS, which the ERG considered to be an important omission. Finally, while the ERG noted that multiple years of follow-up were presented for multiple outcomes, the inconsistency of follow-up duration across outcomes and the small sample size present uncertainties in estimating duration of effect.

The quality of the submitted evidence was acceptable, though the ERG noted Study 301/302 may be at high risk of bias. The ERG agreed substantially with the company's risk of bias assessment for Study 301/302. Study 301/302 did not include blinding of patients and providers given that the use of sham injections was considered unethical. However, quality of methods used for randomisation and the evaluation of the primary endpoint, MLMT, by a blinded rater were strengths of the trial. The ERG did note, however, that the company did not report co-interventions in sufficient detail. The company did not provide quality assessment for Study 101/102, which the ERG undertook. The ERG concluded that the small sample size of the study was a key limitation. Ambiguities in the trial inclusion criteria relating to LCA vs RP meant that the ERG could not draw a conclusion about the applicability of the evidence base across diagnoses.

Overall the evidence indicates a good safety profile. More serious risks associated with subretinal administration of VN and concomitant oral corticosteroid use include endophthalmitis, permanent decline in visual acuity, increased intraocular pressure, retinal abnormalities (e.g., retinal tears or breaks), and cataract development and/or progression. The ERG highlight that these might have long term consequences, especially if they were left untreated. With concomitant use of oral corticosteroid (prednisone) at the time of subretinal injection of VN, the ERG agree that the immune response to AAV capsid and *RPE65* appears mild.

Due to the small patient population included in the trials and indeed the small population with the condition, the representativeness of patients with respect to the UK population of patients with inherited retinal dystrophies is difficult to assess. The ERG regarded that no important groups, by age, ethnicity or sex, were unduly excluded from the relevant trials. The small evidence base presented in the submission is reflective of the rare nature of this condition, but does limit the generalisability of the evidence base beyond the included trials. As there is poor understanding of the characteristics that may impact on disease prognosis and treatment efficacy, it is not possible for the ERG to determine whether the populations of the included trials is consistent with the UK population.

1.5 Summary of value for money evidence submitted by the company

The company submission comprised of a *de novo* cost-effectiveness model constructed to assess the cost-effectiveness of voretigene neparvovec versus best supportive care. The model adopted a Markovian state-transition cohort structure, and comprises of five "alive" health states plus a sixthabsorbing health state representing death. The cost-effectiveness model was constructed in line with the anticipated use of voretigene neparvovec in clinical practice. A lifetime horizon was modelled, and annual discount rates of 3.5% for costs and outcomes were used in the company base case.

The cohort model structure was developed primarily to capture the impact of voretigene neparvovec treatment on health-related quality of life outcomes. Five "alive" health states (based on differing degrees of vision impairment) were used such that different utility values could be assigned to these states. The use of these health states was considered necessary in order to reflect clinically-meaningful differences in health-related quality of life following treatment with VN, and as patients experience progression as part of the natural history of the condition.

Patient transitions from baseline to 1 year were informed by the pivotal Study 301/302, whereas long-term transitions were informed by a combination of clinical expert opinion regarding the long-term effect of voretigene neparvovec and a multistate model fitted to natural history data from the *RPE65* NHx study. Outcomes within the model were based on a combination of visual acuity (VA, clarity of vision) and visual field (VF, range of vision), though the primary endpoint of Study 301/302 was the improvement in the multi-luminance mobility test (MLMT).

Health-state utility values were derived through interviews held with clinicians to complete proxy generic health related quality of life questionnaires for each of the health states in the economic model, based on summary descriptions and their experience with patients. Costs were based on published sources, and were inflated where necessary to reflect the 2018 cost year. The included cost categories considered treatment acquisition, surgery, monitoring, medical resource use, resolution of adverse events, and eligibility testing. Medical resource use utilisation was informed through a combination of assumptions made by the company and input from clinical experts. The company also presented additional analyses to ascertain the impact of treatment beyond costs borne by the NHS and PSS.

In the company's base case analysis, voretigene neparvovec was associated with an incremental cost of [redacted] and a QALY gain of 7.06, leading to an incremental cost-effectiveness ratio of [redacted] (including the proposed simple PAS discount for voretigene neparvovec). The company also presented a range of one-way deterministic and multi-way probabilistic sensitivity analyses, which

illustrated that the key drivers of cost-effectiveness for voretigene neparvovec are the expected long-term outcomes and the quantification of patient health-related quality of life.

1.6 Summary of the ERG's critique of the value for money evidence submitted

The company's submission has been generally developed in accordance with the requirements stipulated within the NICE reference case, and is broadly aligned with the final scope issues by NICE. The company deviated slightly from the final scope to exclude patients without sufficient retinal cells from the economic analysis, which the ERG agreed was appropriate and aligned with the marketing authorisation for voretigene neparvovec and its anticipated use in clinical practice. While the ERG is generally satisfied that the company's model provides a sufficient basis for decision making, the ERG is concerned with a number of assumptions and settings incorporated within the company's submission which have the capacity to lead to substantially different cost-effectiveness estimates.

The cost-effectiveness model structure makes use of a multistate modelling component which the ERG considered to have been unnecessary to inform the estimation of cost-effectiveness within the context of a rare disease. Furthermore, the company's assumed duration of treatment effect for voretigene neparvovec is not considered by the ERG to be robustly supported by the available data from Studies 101/102 and 301/302. The ERG feels that the combined effect of these two features of the company's modelling approach means that the estimation of the long-term effect of voretigene neparvovec is highly uncertain.

Outside of the quantification of longer-term outcomes for patients with *RPE65*-mediated IRD, the estimation of utility values is an incredibly important aspect of the cost-effectiveness model which has the potential to greatly influence cost-effectiveness estimates. A number of methodological issues were identified with the values produced as part of the elicitation exercise, and so the ERG does not consider these utility values to constitute an appropriate basis for decision making.

The ERG also identified a number of other assumptions made in the model that were not clearly supported by the evidence presented. The company assumed vision impairment was associated with increased mortality, though this was based on the findings of a study conducted in elderly patients without *RPE65*-mediated IRD. Medical resource utilisation estimates were also primarily taken from a non-*RPE65*-mediated IRD population, and adjusted based on a number of assumptions relating to relative use between patients with differing extents of vision impairment, and across age groups.

1.7 ERG commentary on the robustness of evidence submitted by the company

1.7.1 Strengths

The company identified what is likely the only other published cost-effectiveness analysis of voretigene neparvovec, conducted by the Institute for Clinical and Economic Review in the United States. The ERG noted some limitations in the company's systematic review that led to the

Section 2, p.26

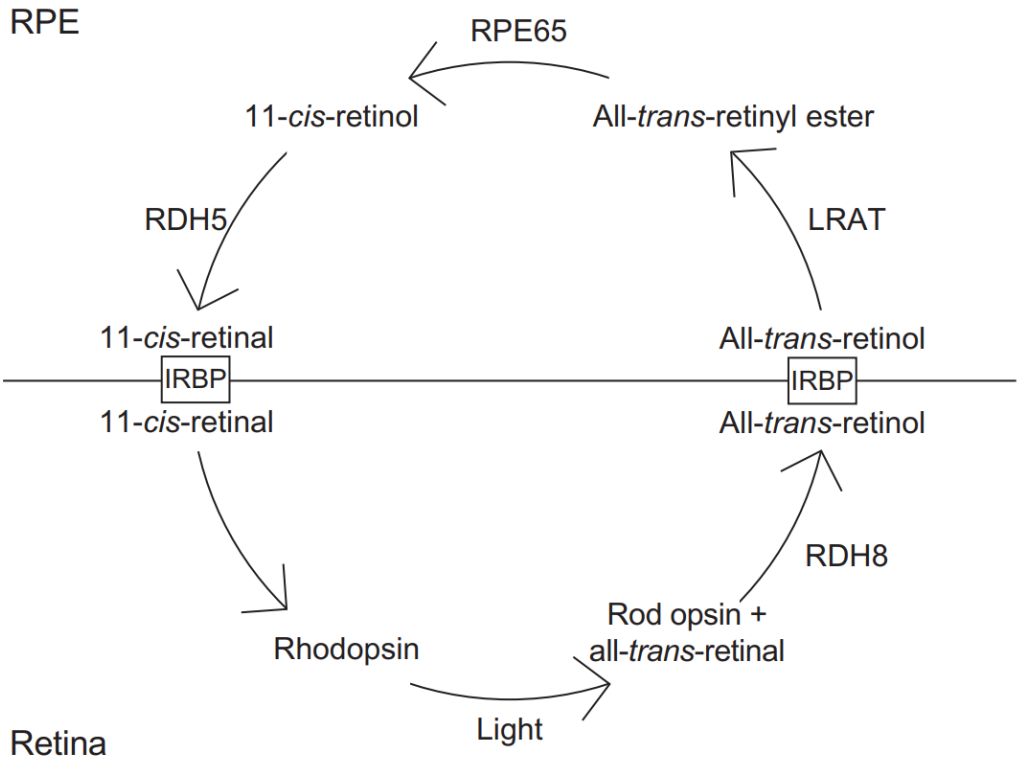
range of values documented in the literature. However, the methods used to arrive at the median values, were unclear.

The CS estimated the prevalence of RPE65-mediated IRD in England to be between 57-564 patients. No references were cited for this data, and the ERG could not find evidence to support these numbers. The incidence of RP was estimated in the CS to be between 0.6 – 1.6 per 100,000 people per year. This evidence was derived from Danish, South Korean and American populations.[7-9] No data was found for the incidence of LCA. The incidence data reported in the CS is consistent with evidence identified by the ERG.

The company estimates that the target patient population for VN in the UK is 86 patients, although it is anticipated that a minority of patients would not be eligible for treatment due to participation in other clinical trials. Their calculations, alongside comments from the ERG, are reported in Table 1.

Section 2, p. 30

Figure 2: The biochemistry of the visual cycle



Source: CS (page 38); original source Wright 2015[21]

Abbreviations: IRBP, interphotoreceptor retinoid-binding protein; LRAT, lecithin retinol acyltransferase; RDH5, retinol dehydrogenase 5; RDH8, retinol dehydrogenase 8; RPE, retinal pigment epithelium; *RPE65*, retinal pigment epithelium 65kDa protein.

2.1.4 Clinical features

The CS reports that individuals with *RPE65*-mediated disease can present at a range of ages between infancy and adolescence. The submission states that nyctalopia (night-blindness) is the first symptom of this disease. The ERG agreed that nyctalopia is typically considered the first symptoms of *RPE65*-mediated IRD,[22] however notes that not all affected patients experience this symptom.[19] The CS reports that infants frequently present with the 'oculo-digital sign' or eye poking. This symptom is a common feature of LCA;[23] however, based on the literature it is unclear how frequently this symptom presents in those with *RPE65*-mediated IRD. Evidence suggests that involuntary eye movement, termed nystagmus, is often observed within this population.[3,19,24]

The CS describes the degenerative nature of the condition and reports that both VF and VA deteriorate over time, accompanied by a loss of retinal sensitivity. It is also stated that there is no Section 2, p. 32

deletion/duplication analyses and/ or other non-sequence based analyses, and is therefore able to detect several of the different types of variants implicated in *RPE65*-mediated IRD.[23]

The CS reports that RPE65 IRD is currently under-diagnosed, with only 50% of people with the disease expected to be diagnosed (CS, p. 42). The company suggest this may be due to the lack of available treatment options undermining the needs for a diagnosis (CS, p.15).

The company note that differentiation of LCA and RP IRD is unreliable, with a minority of patients having received both diagnoses. LCA and RP are typically differentiated by clinical presentation and family history, with LCA presenting earlier and having a more aggressive prognosis (CS, p.36). Clinical experts to the ERG advised that LCA is typically diagnosed shortly after birth, while RP is typically diagnosed in late childhood or early adulthood.

ERG comment:

As discussed in Section 2.1.1, it's unclear whether a 50% diagnosis rate is representative of current practice in the UK; however, the ERG agreed with the company that it is likely that diagnosis rates will increase following the availability of a suitable treatment. The ERG also recognised that diagnosis of the subtypes of LCA and RP IRD may be unreliable. The clinical pathway of care including details of diagnosis is provided in Section 8.2 of the company submission.

2.1.6 Prognosis

The CS discussed the degenerative nature of the disease, which eventually culminates in complete/ near-total blindness.[22] Furthermore, the CS states that there is no evidence of spontaneous sustained improvements in either VA or VF.

The rate at which vision deteriorates in patients with this disease varies considerably, this is briefly acknowledged in the CS. The ERG found evidence which suggests that in some patients vision deteriorates rapidly, while some individuals retain some vision into the second and third decades of life, and others maintain central vision until the end of life.[25-27] Conversely, a cohort study of 70 individuals diagnosed with biallelic *RPE65*-mediated IRD reported that more than half of the cohort were blind by age 18, defined as VA<20/200.[3] This study reported that VA was impaired but stable up until age 15, rapid deterioration was reported between the ages of 15-20, followed by more accelerated deterioration after the age of 20.[3] Overall, this evidence suggests that the prognosis for individuals with *RPE65*-mediated IRD is heterogenous.

Section 2, p. 36

before), patients are required to receive an immunomodulatory regimen (such as prednisone), which is expected to be continued for a further 18 and 30 days, depending on the timing of the second administration (i.e. the other eye).

The European Medicines Agency (EMA) awarded VN marketing authorisation on 22 November 2018.[45] VN is expected to be used in line with the marketing authorisation for the treatment of adult and paediatric patients with vision loss due to inherited retinal dystrophy caused by confirmed biallelic *RPE65* mutations and who have sufficient viable retinal cells. Orphan status was maintained at the time of marketing authorisation:[46] the two previous orphan designations for the "treatment of LCA" and "treatment of RP" were merged to "treatment of IRDs".

The introduction of *RPE65* enables patients to produce functional *RPE65* protein. The subretinal injection of VN introduces a healthy copy of the defective *RPE65* gene into retinal cells.[44] This enables patients to produce functional *RPE65*, resulting in improved functional vision (improved ability to perform vision-related daily activities) and visual function (improved performance of the eyes at the organ level). In order to derive benefit from VN treatment, the company states that patients must have confirmed biallelic (pertaining to both paternal and maternal alleles) *RPE65* mutations and have sufficient viable retinal cells into which healthy copies of the *RPE65* gene can be introduced.

ERG comment:

The CS provides a relatively short description of VN. *RPE65* was noted by the clinical experts as crucial in the visual (retinoid) cycle, and is located in the retinal pigment epithelial cells (discussed further in Section 2.1). Successful introduction of a healthy copy of the *RPE65* gene is expected to lead to long-term improvements in visual function (and consequently, functional vision), though it was noted by the ERG's clinical experts that there is currently no evidence to suggest that introduction may stop degeneration entirely or cause regeneration. The ERG's clinical experts also noted the importance of having sufficient retinal cells in order to benefit from VN – some patients with *RPE65*-mediated IRD may have irreversible retinal deterioration and therefore would be highly unlikely to be able to benefit from treatment.

2.4 Current usage in the NHS

Voretigene neparvovec (VN) is not currently used in the UK for any patient population. VN is the first gene therapy to be approved for a retinal disease.

In the CS, the company proposed that treatment is offered to patients with confirmed biallelic *RPE65* mutations with sufficient viable retinal cells (Figure 3). Genetic testing will therefore be required to determine eligibility for treatment. In the clinical trials of VN, patients were deemed to have sufficient

Section 3, p. 39

3 CRITIQUE OF COMPANY'S DEFINITION OF THE DECISION PROBLEM

3.1 Introduction

The objective of this section is to critique to what extent the CS adheres to the final NICE scope. The scope aimed to evaluate the benefits and costs of VN within its marketing authorisation for treating inherited retinal dystrophies caused by *RPE65* gene mutation. The critique will consider the intervention, population, comparators, outcomes, nature of the condition, impact of the new technology and the cost to the NHS and Personal Social Services (PSS) addressed in the CS.

3.2 Adherence to the decision problem

Table 2 presents a summary of the decision problem as set out in the NICE and some comments from the ERG considering the CS.

Table 2: Adherence of the CS to the decision problem

	Final Scope	**Deviation of CS from final scope**
Population	People with inherited retinal dystrophies caused by *RPE65* mutations	The population is narrower than specified in the scope, but is in line with the licensed indication; i.e. Adult and paediatric patients with vision loss due to inherited retinal dystrophy caused by confirmed biallelic *RPE65* mutations and who have sufficient viable retinal cells
Intervention	Voretigene neparvovec	The intervention is in line with scope
Comparator(s)	Best supportive care	The comparator is in line with scope
Outcomes	• Best corrected visual acuity (both eyes) • Visual field • Contrast sensitivity • Photosensitivity • Need for cataract surgery • Adverse effects of treatment • Health-related quality of life (for patients and carers)	The outcomes assessed are broadly in line with the scope. Of note, the multi-luminance mobility test (MLMT) is also considered as an outcome measure in the CS. This outcome is the primary measure considered in the pivotal clinical trial. The ERG also noted that data for contrast sensitivity and the need of cataract surgery were not reported in the CS; and no health-related quality of life data was presented. No data for the impact of treatment on carers was presented.

	Final Scope	Deviation of CS from final scope
Subgroups to be considered	None specified	Not applicable
Nature of the condition	• Disease morbidity and patient clinical disability with current standard of	The nature of the condition is broadly in line with scope. However, the ERG noted the

Section 3, p. 43

ERG Comment:

In summary, given the population for which evidence has been submitted, the ERG and its clinical advisors agreed with the company that BSC is the most relevant comparator in the setting of IRDs caused by *RPE65* gene mutations.

3.6 Outcomes

The company state that no treatments are currently available for *RPE65*-mediated IRD, and therefore no precedents exist for endpoints to assess the therapeutic benefits of products for this unique group of diseases. The measurement of visual acuity (VA), VF and contrast sensitivity are generally well accepted as the best visual predictors of mobility performance. For people with low vision, orientation and mobility are more affected by spatial contrast sensitivity and VF than by VA, although these parameters vary widely. The measurement of VA, VF and contrast sensitivity is clinically relevant in the population for this assessment, and is consistent with the evaluation of visual impairment across other populations. However, these endpoints are challenging to measure in the population considered in this assessment because baseline visual function is poor, and they do not capture characteristic features of the condition; e.g., night blindness, reduced light sensitivity, and nystagmus. These measures are also difficult to use in paediatric populations.

In context of these condition-specific features the company designed and validated the multi-luminance mobility test (MLMT).[47] The MLMT measures changes in functional vision, as assessed by the ability of a subject to navigate a course accurately at a reasonable pace at different levels of environmental illumination. Change in MLMT from baseline to one year was the primary endpoint of the company's pivotal Phase 3 clinical trial (Study 301/302). Although the ERG noted that these data are not used in the economic model.

The ERG noted that no data was reported for the need of cataract surgery following treatment. Safety data indicate that patients receiving VN are at a higher risk of cataracts, and the proportion of patients who would require cataract surgery was estimated in the company's economic model, although the basis for this estimation is unknown.

Finally, no health-related quality of life (HRQoL) was reported in the CS. Rather, the company present the impact of treatment with VN on visual function using a patient-reported outcome (PRO). However, this evidence does not capture the possible impact of treatment on the broader HRQoL of

patients. Further, no evidence was presented on the impact of treatment on the carers of patients with *RPE65* IRD.

impact on the sight of patients with biallelic *RPE65* IRD. This can evidently not be demonstrated from the current treatment follow-up; however the ERG judged that a four year follow-up is acceptable for determining whether VN may result in some clinical benefit for patients.

Treatment was administered at centres in the US, however feedback from clinical experts for the ERG was that the settings of the evidence base can be generalised to UK practice.

4.2.2.2 Population

The key inclusion and exclusion criteria reported in the CS for both studies are summarised below in Table 9.

The ERG noted that the population characteristics used in the included trials for the technology of interest (VN) and best supportive care (BSC) were consistent with licensing authorisation; i.e. adult and paediatric patients with vision loss due to inherited retinal dystrophy caused by confirmed biallelic *RPE65* mutations who have sufficient viable retinal cells. The ERG noted that the population characteristics included in all three studies were narrower than those specified in the NICE scope for this appraisal; however, the ERG judged the change to be appropriate. Expert advisors to the ERG suggested that the requirement for patients to have a sufficient number of viable retinal cells is necessary to facilitate the treatment mechanism of VN. The ERG noted that patients are excluded from the included trials if they have a retina less than 100 microns (equivalent to more than half of a normal retina's thickness). Expert advisors to the ERG acknowledged that while 100 microns seems to be an arbitrary number (and apparently being used as a proxy for the health of the photoreceptors), if VN is injected into a retina with thickness of less than 100 microns, it seems reasonable to assume that there would be fewer viable retinal cells and hence improvements would be less likely. Given the localised action of gene therapy, and the need for safe administration of VN to sufficient retinal cells to ensure there are grounds for improvement, the ERG agreed that it seems reasonable to limit the trial population to people with retina thickness of more than 100 microns at the site of injection. However, the ERG noted that this additional criterion would mean it likely that some patients included in the population specified in the NICE scope would be excluded for treatment with VN because they have no viable retina to treat. In practice, it's unclear whether this threshold of retinal thickness would be strictly used: the company state (CS, p.54) that they expect OCT tests in practice to be more qualitative, and to be supplemented by tests of VA and VF. Clinical advisors to the ERG suggested that this may result in a similar population identified for treatment, as patients who demonstrate visual function using VA and/or VF tests may be assumed to have sufficient retinal cells to experience some treatment benefit.

The ERG also noted that population inclusion criteria for Studies 101/102 and 301/302, as described in the CS and trial CSRs, specify the inclusion of patients with a specific subtype of *RPE65* related IRD, Leber's congenital amaurosis (LCA). However a footnote to the inclusion criteria (CS Table 9, p. 71-74; CS Table 11, p. 84-87) adds that patients were eligible if they had a "molecular diagnosis (or confirmation of diagnosis) of biallelic *RPE65* mutations… regardless of clinical diagnosis". This presumably permits the inclusion of patients with RP IRD. However if this is the case, the ERG are unclear why trial inclusion criteria primarily specify patients with LCA only, and whether this means that patients with LCA were favoured in recruitment strategies for the trials, or constituted a higher proportion of patients in the trial samples. The CS did not provide a breakdown of the proportion of patients diagnosed with LCA vs. RP IRD, and the ERG were unable to find this information in the respective CSRs. While the ERG acknowledge some overlap in the diagnostic criteria for RP and LCA, typically patients with LCA are rarer and exhibit a more aggressive prognosis.[65] Clinical advisors to the ERG were unaware of evidence that would prevent generalising evidence from patients with LCA to those with RP, and suggested that the treatment effect is likely to be unaffected by diagnosis. However, the ERG noted that absolute data (such as the speed of visual deterioration) may not be comparable between LCA and RP patients. Nevertheless, as it is not clear from the CS whether trial samples involved a greater proportion of LCA patients, it is not possible to draw a conclusion about whether this could affect the applicability of the evidence base.

Patient populations and eligibility criteria were broadly similar between the Phase 1 and Phase 3 trials, although three changes in inclusion criteria for Study 301/302 are notable. Firstly, trial inclusion criteria for Study 301/302 was extended to include younger children between the ages of 3 and 7 years. Age is thought to influence the potential treatment effect of VN, due to the potential benefits of administering VN prior to further retinal degeneration. Criteria for Study 301 were further restricted to include those with less severe deficits in VA (from VA of 20/160 in Study 101/102 to 20/60 in Study 301/302). Baseline VA was not reported in the CS for Study 101/102, although at clarification the company provided mean baseline VA for Study 101 (no variability data was provided), which suggested that baseline VA was better for patients in Study 301/302. Clinical experts to the ERG advised that both age and baseline VA may have an impact on treatment outcome, and therefore differences may be expected in the treatment outcome between Study 101/102 and Study 301/302; although the direction and magnitude of any difference is not yet understood. Ultimately as Study 101/102 is under-powered to evaluate clinical effectiveness of VN and is non-comparative in design, emphasis on clinical efficacy outcomes should be given to data from Study 301/302.

Changes in eligibility criteria were included for patients in Study 102 following their participation in Study 101; these were intended to ensure that patients had VA equal to or greater than light perception

Section 4, p. 60

Study 101/102

Population characteristics for patients in Study 101/102, as reported in the CS, are summarised in Table 12.

As expected for the low sample size within each dosing arm,

[redacted]. There is an absence of evidence for the role of gender in treatment prognosis for this patient group, however as noted above, clinical advisors to the ERG advised that age at baseline may impact on the likely treatment effect, with treatment at a younger age being potentially more beneficial. Baseline visual performance was not reported in the CS, although consistent with procedures for favouring the worst, non-preferred eye for injection in Study 101,

[redacted] (Study 101 CSR, p. 59).

Table 12: Study 101/102 Patient Demographics (all patients)

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	7 (58%)	6 (55%)
	[redacted]	[redacted]	[redacted]	[redacted]	5 (42%)	5 (46%)
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	11 (92%)	10 (91%)
	[redacted]	[redacted]	[redacted]	[redacted]	1 (8%)	1 (9%)
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	20.8 (11.2)	22.8 (10.26)
	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: CSR, Appendix 6, Table 1

ERG Comment:

There are several differences in population characteristics between the VN and BSC arms in Study 301. Given the small size of the trial, the ERG considered a number of differences between arms to be inevitable and to not necessarily represent a violation in randomisation. None of the differences at

baseline were considered by the ERG to demonstrate a clear bias in any direction, although it was noted that only a small number of characteristics were reported at baseline.

Section 4, p. 64-65

The ITT population (all patients randomised) was stated to be prioritised for clinical outcomes, while the mITT/safety population (excluding 2 patients who dropped out of the study prior to knowing treatment allocation) is reported for AE data and for some outcomes, which was judged by the ERG to be appropriate.

Several limitations in outcome assessment were noted as important. Firstly, while randomisation was stratified by age (</≥ 10 years), it was not feasible for the company to adjust outcome data for baseline characteristics, due to the small sample size of both trials. It is unclear how this limitation may impact on the treatment outcome; based on the limited data provided and the evidence known about prognostic markers in this population, there is no consistent pattern in either amplifying or reducing the potential treatment effect.

Secondly, as noted in the CS, scoring for the MLMT exhibits a ceiling effect inherent to the design of the task. As the test does not allow for testing at light levels lower than 1 lux (equivalent to a moonless summer night or an indoor night light; CS p. 78), change scores will be capped at this light setting. The ERG agreed with the company's assertion that this may underestimate the mean change in patient scores on the test, which may result in an underestimation of the treatment effect. This will be applicable to continuous data only (mean final/change scores), but will not impact on the proportion of patients who achieved a change greater than 1 light level, which is also reported in the CS, as all patients were at least 1 light level away from the ceiling at baseline.

Thirdly, while VA and VF are the only two outcome measures that have been used successfully to approve new drugs for retinal application, there are known limitations with the reliability of their measurement. Natural variability in VA between assessments means that obtaining a representative estimate may require multiple tests. In Study 301/302, VA was assessed as the average BCVA of each eye (rather than bilaterally). The company state that this may underestimate the clinically useful vision that is achieved with both eyes open (CS, p. 136). Further, many patients with IRD have such poor vision or fixation that VF testing cannot be performed reliably; while VF testing is clinically relevant as a loss of visual field is a key and early symptom of the condition, this very feature can lead to indeterminate test results (CS, p. 82), and is likely only possible in children over 7 or 8 years of age. Further, it should be noted that available measures of contrast sensitivity rely on knowledge of the alphabet, and are therefore not suitable for use in children unable to recognise letters.

Fourthly, the ERG do not consider the VFQ to be an appropriate replacement for a measure of HRQoL. The NEI VFQ, which has been used extensively to evaluate vision-related functioning in patients with age-related macular degeneration, and demonstrates good reliability and construct

validity,[66] was modified for use in Study 301/302. The CS does not report details about the way in which the measure was modified, however a report of the psychometric properties of the measure provided by the company describes the modifications as 'substantial' (p.10).[67] These modifications are stated to have been made to better assess functional vision in patients with *RPE65* IRD, and clinical advisors to the ERG advised that the modifications were appropriate. Psychometric data for the tool also indicates that it demonstrates good reliability and validity. However, the ERG noted that in this process items related to HRQoL were removed from the tool, and therefore this outcome is considered by the ERG to be appropriate for evaluating visual function in this patient population, but cannot be used to evaluate HRQoL.

Finally, it should be noted that the objective of Study 101/102 was to evaluate the safety of VN, and while clinical efficacy endpoints were evaluated (including VA, VF, FST, contrast sensitivity, and mobility assessment), the study was not powered to evaluate change in these outcomes.

ERG Comment:

The measurement of VA, VF, and contrast sensitivity was clinically relevant in this patient population, and is consistent with evaluation of visual impairment across other populations. However, their measurement is widely considered to be unreliable, due to inter-test variability in this population requiring greater improvements from baseline to demonstrate a treatment benefit. MIDs for these outcomes are derived in consideration of inter-test variability.

The ERG agreed that the ceiling effect inherent to the MLMT measure may underestimate the treatment effect reported for continuous data. The ERG considered this to be an important outcome for evaluating the impact of visual impairment on functioning; however a clinical advisor to the ERG suggested that the current scoring (change in the light level under which patients could complete the course) may be less sensitive to assessing functional vision than a change in the time it takes patients to complete. The ERG also considered there to be uncertainty in the validity of the company's threshold for a clinically meaningful change (1 light level).

The modified VFQ should be considered an appropriate measure of functional vision in these patients, and has acceptable psychometric properties. However, items related to HRQoL from the original tool were removed, and the ERG did not consider this measure to measure HRQoL following treatment with VN. No HRQoL data, or PRO data to evaluate the burden of *RPE65*-mediated IRD for carers, was reported in the CS, which the ERG considered to be an important omission.

Section 4, p. 69

Endpoint		**Study 101**	**Study 102**	Study 301/302
	Statistical methods	Change in full-field light sensitivity before and after injection FST data were not available for all patients/timepoints as the equipment was not available at the start of the trial (CS, p. 116). Missing values were treated as missing without any imputation	Change in FST following injection to the contralateral eye evaluated using pre-injection, follow-on baseline evaluations as a control.	Change in white light FST averaged over both eyes at year 1 relative to baseline
	Analysis population	PP	ITT	ITT and mITT
Mobility testing	**Definition**	Subject's ability to navigate a short obstacle course with both eyes open (except for some cases where either the injected eye or the uninjected eye was occluded) and varying light levels. Lower scores = better performance Change ≥1 light levels indicates a clinically meaningful improvement	Subject's ability to navigate a short obstacle course with both eyes open and varying light levels. Lower scores = better performance Change ≥1 light levels indicates a clinically meaningful improvement	Subject's ability to navigate a short obstacle course with both eyes open. Lower scores = better performance Change ≥1 light levels indicates a clinically meaningful improvement
	Time-points outcome reported	N/A	Baseline, d60, d90, yr1, yr2, yr3 and yr4	Baseline, d30, d90, d180, yr1, yr2, yr3 and yr4
	Statistical methods	ITT population Monocular assessment: evaluated in first treated eye.	ITT population. Monocular and bilateral assessment. Change in MLMT following injection to the contralateral eye evaluated using	ITT [primary] and mITT [secondary] Monocular and bilateral assessment. Change in bilateral mobility test performance relative to baseline.

			pre-injection, follow-on baseline evaluations as a control	Bilateral performance on the MT as measured by a change score.

Section, p. 84

No data were reported in the CS with regards to contrast sensitivity for patients in Study 101/102.

ERG comment:

[redacted]

4.2.3.1.4 Photosensitivity

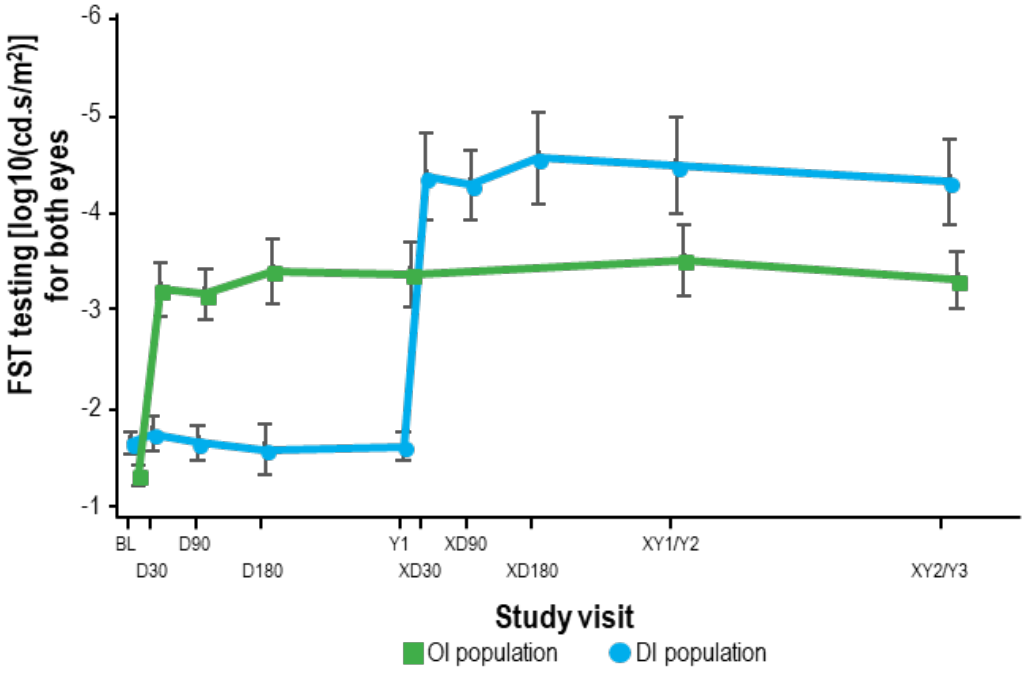
Details of the measurement of photosensitivity in the included trials is summarised in Section 4.2.2.4.

Study 301/302

A statistically significant difference in full-field light sensitivity (FST) threshold was reported at 1 year (MD -2.11 log units; 95%CI -3.91, -1.04; p=0.0004; ITT population). Patients in the VN arm exhibited a mean improvement in FST of -2.08 (SE 0.29), while no change was exhibited by patients receiving BSC (mean change 0.04; SE 0.44).

At 3-year follow- up, the effect of VN on FST was maintained in the original intervention arm (mean change -2.04; SD 1.43; N=19), as well as in those who crossed over from the BSC arm (mean change -2.69; SD 4.41; N=9; see Figure 15). These changes were above the company's defined threshold for clinical significance (≥1 log unit).

Figure 15: Study 302 Full-Field Light Sensitivity Threshold at 3 years



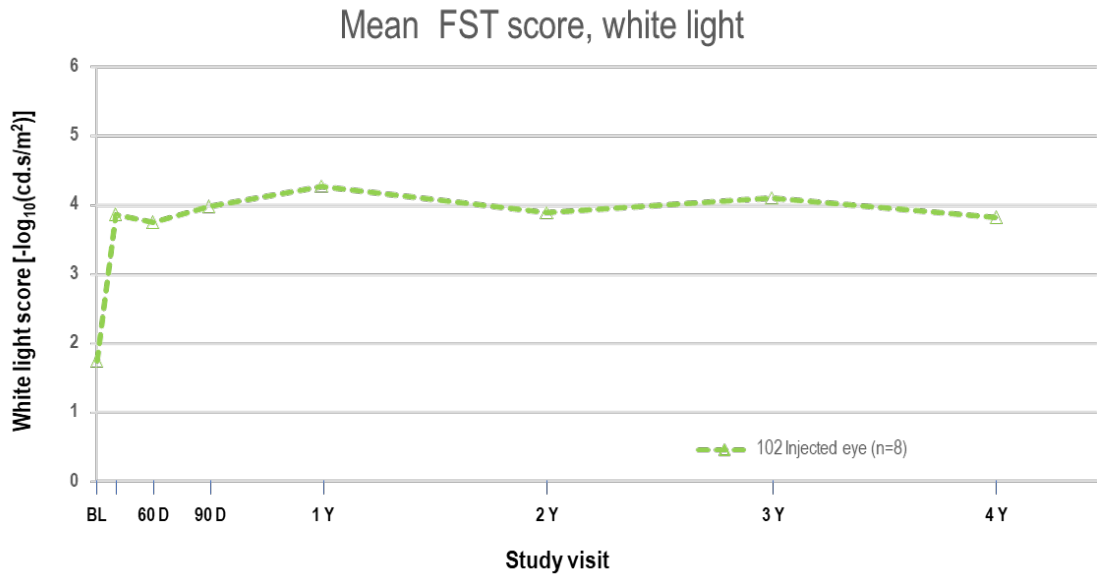
Abbreviations: DI, delayed intervention; OI, original intervention

Error bars represent standard errors

Section 4, p. 86

Eyes in Study 102 were better functioning at baseline, and all received a high dose of VN. The company provide a graph (Figure 17), which appears to show an improvement in FST from baseline, which is then maintained at four years (N=8).

Figure 17: Study 102 FST Mean Score for Eyes injected at 4 years



Abbreviations: FST, full-field light sensitivity threshold

Source: Maguire 2017[49]

ERG Comment:

The evidence from Study 301 suggests that VN has a small, statistically significant effect on FST at 1 year, which was above the company's threshold for a clinically meaningful difference (3.90 dB; Roman et al, 2005).[72] While the effect was seen consistently across follow-up, wide error bars around the effect were noted. No further data for FST is reported for study 301/302, and therefore it is not possible for the ERG to determine if the effect was maintained, or altered, after 1 year.

Evidence from Study 101 indicates a possible numerical improvement in FST following VN, which was shown consistently across follow-up, but again below the threshold for a clinically meaningful difference. A large effect on FST was reported in Study 102, however only 8 patients were included, and no variation data was reported.

Section 4, p. 90-91

[redacted]

	[redacted]	[redacted]
[redacted]		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

[redacted]

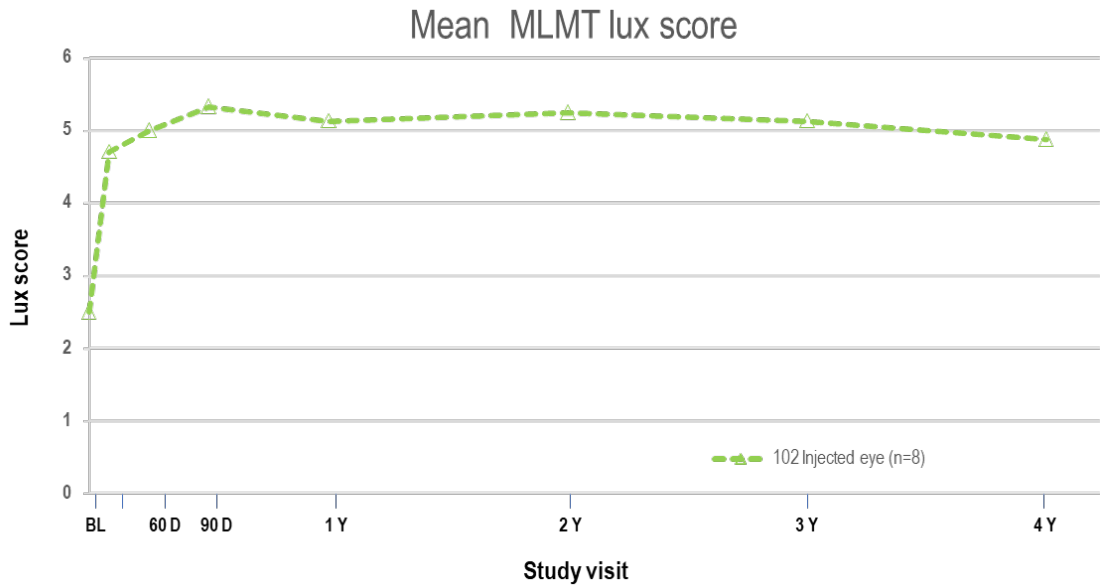
Source: Trial CSR,[59] p. 27

At clarification, the ERG requested if the company had found a difference in treatment effect between children (<18 years) and adult (≥18 years) patients. [redacted]

Study 101/102

In Study 102, 8/11 (72.7%) patients were evaluated using a mobility test (which subsequently became the MLMT). The CS reports that all 8 patients demonstrated a clinically significant improvement of ≥1 light level with their second (better, preferred) eye, and 5/8 (63%) patients passed the MLMT at the lowest level (1 lux). This data is presented in Figure 20 below. This figure demonstrates a sharp improvement in mean MLMT following administration of VN, which is maintained until follow-up at 4 years. Mean change in MLMT score was 2.6 (SD 0.56) at 1 year follow-up, and 2.4 (SD 0.46). These 8 patients were all stated to meet inclusion criteria for Study 301/302.

Figure 20: Study 102 MLMT Mean Score at 4 years



Abbreviations: MLMT, multi-luminance mobility test

Source: Maguire 2017[49]

ERG Comment:

The evidence from Study 301/302 indicates that treatment with VN was associated with a statistically significant improvement in MLMT, which is clinically significant according to the company's chosen clinically meaningful threshold (change ≥1 light level). Based on this threshold, all patients who received VN in the included trials exhibited a clinically meaningful change in MLMT score. This improvement was also shown to be maintained until follow-up at 4 years (3 years in delayed arm).

4.2.3.2 Patient-Reported Outcomes/Health-Related Quality of Life

Details of the measurement of visual function in the included trials is summarised in Section 4.2.2.4.

Study 301/302

Mean scores for the modified Visual Function Questionnaire (VFQ) at 1 year are presented in [redacted]

Section 4, p.95

It is interesting that – according to the distribution method of deriving MIDs[67] – [redacted] Clinical advisors to the ERG advised that patients are likely to adapt to their surroundings over time, which may explain a proportion of the change in HRQoL in both arms. [redacted] .However [redacted]

As with several of the other outcomes included here, evidence for the impact of VN on VFQ scores is based on one small RCT only, with no follow-up data. [redacted]

Finally, the ERG noted that the absence of HRQoL data in the trial adds an additional uncertainty to the economic evaluation. This is explored in depth in Section 5.2.7 below.

4.2.3.3 Safety data

The CS reports that no deaths were reported in any of the included trials. Safety data was reported as treatment-emergent AEs (TEAEs; Section 4.2.3.3.1); serious AEs (SAEs; Section 4.2.3.3.2); drug-related AEs (Section 4.2.3.3.3) and administration-related AEs (Section 4.2.3.3.4).

Details of the measurement of adverse events in the included trials is available here.

4.2.3.3.1 Treatment-emergent adverse events

The company did not report their definition of TEAE in the CS; however the ERG assumed that a general definition of TEAE was used, i.e. any AE occurring following administration of treatment, irrespective of the frequency or whether this was deemed to be related to the study drug. A breakdown of TEAEs according to whether these were deemed to be SAEs, drug- or administration-related is provided in Sections 4.2.3.3.2 and 4.2.3.3.4.

Section 4, p. 99

[redacted] was recorded in Study 101, [redacted]

Study 102

[redacted]

4.2.3.3.3 Drug-related adverse events

The CS reports that [redacted]

4.2.3.3.4 Administration procedure-related adverse events

Study 301/302

The proportion of administration-related AEs were not reported separately for patients in Study 302; i.e. the first year after treatment for patients in the Original VN arm of Study 301. These AEs are incorporated into follow-up data for Study 102.

The company provides a summary table of administration procedure-related TEAEs reported by patients in Study 301/302 from baseline to final follow-up ([redacted]). In total, [redacted] patients receiving VN exhibited a total of [redacted] that were considered by the company to be related to the administration procedure: [redacted] patients in the Original arm and [redacted] in the Delayed arm. In total, [redacted] patients experienced an eye disorder related to administration: [redacted]

[redacted] although the company’s criteria for determining this was not reported.

Section 4, p. 103

[redacted]	[redacted]
[redacted]	[redacted]

[redacted].

Source: CS Table 23, p. 126

ERG Comment:

Overall, the evidence indicates that VN is associated with an acceptable safety profile. No deaths were recorded during the trials, and no serious AEs were thought to be related to VN itself. The company reported that 7.3% of patients experienced a non-serious, transient AE related to VN, which did not require treatment. The administration of VN, however, is associated [redacted]

4.2.4 Meta-analysis

Only 1 comparative study (Study 301/302)[61] has been conducted to evaluate the relative effectiveness of VN to treat IRD. As such, no meta-analysis of clinical effectiveness was provided, or expected.

4.2.5 Quality assessment of the included evidence

The company conducted quality assessment of Study 301/302; quality assessment judgements reported by the company are reported in Table 25, alongside ERG comments. No quality assessment was reported for Study 101/102, but was conducted by the ERG (Table 26).

Table 25: Study 301/302 Quality Appraisal

Study question	Company response (yes/no/not clear/ N/A)	Company description of how the question is addressed in the study	ERG response (yes/no/not clear/N/A)	ERG comments
Was randomisation carried out appropriately?	Yes	A randomisation list was generated under the direction of the independent party biostatistician using a	Yes	The ERG agree that there is a low risk of selection bias associated with the randomisation procedure.

Study question	Company response (yes/no/not clear/ N/A)	Company description of how the question is addressed in the study	ERG response (yes/no/not clear/N/A)	ERG comments
		permuted block design, stratified by age (<10 years and ≥ 10 years) and baseline mobility testing passing level.		Randomisation was determined by order of enrolment, verification of study eligibility, and the participants' randomisation stratum.

Section 4, p.109

unreliable due to natural variations in visual function between tests. Nevertheless, there is some uncertainty over the validity of MIDs for both the MLMT and the modified-VFQ; both are new outcomes with limited validation. Furthermore, as no HRQoL data was reported, it is not possible for the ERG to conclude on whether improvements in visual function translate to broader improvements in patients' HRQoL.

The ERG noted that numerical improvements in visual function were exhibited by patients receiving VN; including VF, VA, FST, and [redacted]. These improvements exceeded MIDs for VF and FST. While improvements in VA and [redacted], these were nevertheless demonstrated consistently across follow-up timepoints, suggesting a potential minor effect of VN on these outcomes, beyond the natural variation that would be expected in these outcomes.

The evidence suggests that VN demonstrates an acceptable safety profile. No SAEs were considered to be due to VN, and no deaths were recorded in the included trials. The administration of VN is associated with [redacted]; [redacted]. However, as per the current license for VN, these risks would be limited to a single administration.

4.4.1 Key areas of uncertainty

The ERG noted that a small sample of patients available at later follow-up for Study 301 exhibit [redacted]; however the potential of VN for longer-term gains in visual performance and function remains unclear until longer follow-up data is available.

The small evidence base presented in the submission is reflective of the rare nature of this condition, but does limit the generalisability of the evidence base beyond the included trials. As there is poor understanding of the characteristics that may impact on disease prognosis and treatment efficacy, it is not possible for the ERG to determine whether the populations of the included trials are consistent with the UK population.

[redacted]

Section 5, p.113

The CS does not contain a clear summary of the findings of the review (including how the ICER study may have helped inform the cost-effectiveness model submitted to inform this appraisal). In clarification, the company provided the table of excluded studies for this systematic review. This is clearly presented with most studies being excluded on publication type, population or outcome. The company also provided the tables of excluded studies for the resources and health utilities reviews (almost all were excluded on outcome).

While not necessarily a summary of the findings of the review, the CS provides a comparison of outcomes between the company model and the study identified by the literature review, as well as where assumptions and/or analytical methods differed. Discussion of the identified cost-effectiveness study is presented in Section 11.2 of the CS (p. 158-159).

5.2 Summary and critique of company's submitted economic evaluation by the ERG

5.2.1 NICE reference case checklist (TABLE ONLY)

Table 28: NICE reference case checklist

Element of health technology assessment	Comments with reference to the scope	Issues arising	Section providing details
Defining the decision problem	The company's description of the decision problem builds on the scope definition. The population is narrower than specified in the scope, but is in line with the licensed indication.	None.	**3.2 & 5.2.3**
Comparator(s)	The comparator described in the CS is BSC, which is in accordance with the final scope.	A formal definition for BSC is not provided. VN in the cost-effectiveness analysis might be equivalent to VN+BSC.	**5.2.4**
Perspective on outcomes	The list of the outcomes in the CS includes all those listed in the final scope, as well as MLMT, the primary measure in the pivotal clinical trial. Some of these, including MLMT, are not used in the economic evaluation due to a lack of related cost and utility data. Health states in the economic evaluation are defined by VA and VF.	It is written in section 9.4.1.1.1 of the CS that VA and VF do not capture all of the features of the condition, and hence some direct health effects may not be accounted for in the economic evaluation.	**3.6 & 5.2.2**
Perspective on costs	The company consider costs from the perspective of the NHS and PSS.	None.	**5.2.5**
Type of economic evaluation	A cost-utility analysis with outcomes reported as ICERs in cost per QALY gained.	None.	**5.2.2**

Element of health technology assessment	Comments with reference to the scope	Issues arising	Section providing details
Time horizon	A lifetime horizon has been adopted, which means that patients have been followed until maximum age of 100 years.	None.	**5.2.5**

Section 5, p.120

ERG Comment:

The ERG was generally satisfied that the cost-effectiveness model reflects the patient population specified within the final NICE scope, which is aligned with the 301/302 study and the European Marketing Authorisation. The ERG acknowledged that studies of rare diseases are often fraught with issues relating to sample sizes, generalisability and non-standard clinical study design. The decision to deviate from the scope in regards to the population of patients with insufficient retinal cells is consistent with the marketing authorisation for VN, and is aligned with the expected use of VN in clinical practice.

Clinical expert opinion sought by the ERG confirmed that it was appropriate for the two conditions (RP and LCA) to be grouped for the purpose of assessing the clinical- and cost-effectiveness of VN. However, it should be noted that only patients with LCA were enrolled within the clinical studies of VN, and therefore there is no clinical evidence pertaining to the use of VN in an RP-specific population.

Within the company's cost-effectiveness model, the distribution of patients at baseline by health state is based upon the pooled estimate across both treatment arms of Study 301/302. Due to the small sample size, the proportions of patients within each treatment arm differ to the pooled estimate (as shown in Table 31). Furthermore, the natural history study (*RPE65* NHx) comprises of a less severe population (87% of patients reside within HS1 or HS2 at baseline, versus approximately 55% of the ITT population within Study 301/302 [based on Table 31 and Table 32]).

The ERG noted that a total of n=70 patients were considered "eligible" in the *RPE65* NHx study. However, in Table 32 the total number of patients sums to 68. Further to this, within Section 12.1.8.3.3 of the CS, it is stated that *"67 patients were included in the analysis"*. The ERG requested clarity from the company regarding the baseline characteristics of patients in the *RPE65* NHx study, and were referred to the original study report which unfortunately does not provide information regarding health state allocation, or specific reasons why some patients may have been excluded.

For the purpose of the ERG report, a total of n=68 patients are assumed to be relevant to the analysis (based on the outputted log file from the statistical analysis discussed in Section 5.2.6).

The differences in characteristics between treatment arms extends to the average age of the cohort. The mean age for patients treated with VN is 14.8 years, versus 15.9 years for patients receiving BSC.

Clinical advice provided to the ERG suggested that treatment may be given at any age, and that there is no clear relationship between outcomes and age within an *RPE65*-mediated IRD population.

Within NICE HST7 (strimvelis for treating severe combined immunodeficiency caused by adenosine deaminase deficiency, a different gene therapy),[79] it was stated that *"age is a factor that may*

Section 5, p.129

> *"A multistate survival model allows for the risk of moving between health states to vary over time, as may be expected in clinical practice. Multiple alternative survival distributions can be tested to determine the most plausible extrapolation of observed data, including the assumption of constant risk (i.e. the exponential distribution). In addition, by parameterising the risks of moving between health states, this approach allows for parameters determining the long-term health state distribution to be tested in univariate and probabilistic sensitivity analysis."* (CS, Section 12.1.8.3.2, page 185)

In line with clinical expert opinion, the company specified the MSM as *"progressive only"*, such that the only permitted transitions were those to "poorer" health states (i.e. it was not possible for patients to experience an improvement in health state beyond Year 1). The company also highlighted that the implementation of a progressive only MSM is less complex to implement within the cost-effectiveness model (versus an unrestricted MSM). Transitions to the "dead" health state were not captured by the MSM, as no death events were reported within the *RPE65* NHx study.

A parametric multistate (five state) Markov MSM was fitted by the company. Within the context of the MSM approach, the Markov assumption implies that the probability of movement to another state is independent of the time spent in the current state, instead the probability of movement to another state is dependent on the time since model entry. The ERG noted that it is important to flag that the Markov assumption within the context of an MSM differs to the traditional definition used to describe a Markov cost-effectiveness model wherein transitions may be considered Markovian (memoryless) if they are independent of time, such as in the case of an exponential distribution.

The company fitted the MSM using the Stata software package *MULTISTATE*.[82] The company successfully fitted a total of 5 MSMs, based on the following statistical distributions: exponential, Weibull, Gompertz, log-logistic, and log-normal. A generalised gamma MSM was also attempted, though the company noted that this model did not converge. The MSM fits were specified assuming proportionality between baseline hazard functions and the transition intensities within the same distributional model.

The statistical fits of the models were compared using Akaike and Bayesian information criteria (AIC and BIC, respectively), in addition to an analysis of Cox-Snell residual plots. The ERG requested further information regarding the Cox-Snell residual plots provided within the CS at clarification stage. The company provided some data used to inform the Cox-Snell residual plots, but did not provide as explanation as to what exactly they were intended to illustrate.

The Weibull MSM was selected to inform the company's base case, and was selected according to both statistical fit (lowest AIC and BIC) and *"visual inspection"*. To illustrate the base-case projections of the MSM component of the model, a plot is presented in Figure 25 which shows

Section 5, p.138-139

ERG Comment:

The lack of patient-reported values for patients treated with VN is a key limitation of the evidence base provided by the company, and introduces considerable uncertainty to the economic evaluation. This uncertainty relates to both the ERG's assessment of the clinical-effectiveness of VN, as the impact of treatment on patient HRQoL is unclear due to the lack of a validated patient-reported outcome measure, and in terms of the economic evaluation, as it is unclear which utility values are the most appropriate for use.

The proxy elicitation exercise that was conducted by the company suffers from severe methodological and face validity issues, as well as being subject to a number of biases. These include the use of proxies (clinicians in this case) for patient values, which have been seen in multiple instances to be a poor surrogate of patient values, and the questions being asked over the telephone by researchers, as opposed to completed by the clinicians without interaction. Methodologically, the ERG is concerned that as clinicians will be focused primarily on vision-related issues faced by patients (the health state descriptions are vivid in their descriptions of limitations), and that this may introduce a 'framing' effect wherein clinicians are unlikely to take into account the broad range of activities patients can perform that are unrelated to vision loss. The use of only 6 respondents (not taken from the general public), also limits the generalisability of the results and is not aligned with the NICE reference case.

At clarification stage, the ERG asked the company to confirm which order questions were asked in, as this may also influence the responses provided. The company provided a further report detailing the design of the elicitation exercise, but unfortunately this did not explicitly state the order the questions were asked. However, given the order of the report, it appears that clinicians were asked to first complete the questionnaires for the 'best' health state, and then subsequently the questionnaires for deteriorating health states. This ordering is likely to have impacted results by 'capping' the utilities of each state by the previous one. Were the order of the health states reversed and HS5 (hand motion, light perception, no light perception) valued first, the results may have been substantially different. A clear example of the effect of ordering can be seen in the Czoski-Murray et al (2009) study referenced by the company.[86] In the study members of the public were given vision altering contact lenses to simulate different levels of vision impairment - their valuation of the states however varied depending on the order in which they received the contact lenses (Table 2 of the paper).

The lack of face validity is due to two related issues: firstly, the absolute values derived via the proxy elicitation exercise not appearing to match with the patient experience described by the ERG's clinical advisors, and secondly, the negative value for HS5 (HUI3 analysis). When asked to describe the

HRQoL of patients, the ERG's clinical advisors stated that unlike many other vision disorders, patients had restrictions imposed by their vision, but in general did not have other physical health problems. As the patients had always experienced vision problems, they did not experience a sense of 'loss' from otherwise average vision, and continued to perform their usual activities, modifying these over time – for instance taking up disability sport (possibly to high levels). Even with extremely poor vision, patients were described as leading meaningful lives with high levels of enjoyment. The description given of patient's lives did not correspond to the utility values provided by the company. When asked specifically about the value for HS5 (for which a negative value is indicative of a health state *"worse than death"*), this was not recognised by clinicians for patients in this indication, and did not appear to be representative of the patient population residing in this health state.

To investigate the apparent lack of face validity, the ERG reviewed all previous NICE submissions involving vision loss to gain a broader understanding of the utility values used to inform previous appraisals. While there have been no specific submissions in *RPE65*-mediated IRD, nearly all appraisals incorporated health states based on vision loss. The results of this review are reported in Table 37.

Table 37: Summary of range of utility values in previous NICE TAs

Number	Category	Lowest and highest utilities
TA155	Macular degeneration	0.40 and 0.89
TA229	Macular oedema and retinal vein occlusion	0.548 and 0.750
TA274	Macular oedema and retinal vein occlusion	0.353 and 0.869
TA283	Macular oedema and retinal vein occlusion	0.314 and 0.869
TA294	Macular degeneration	0.31 and 0.920
TA297	Eye conditions: general and other	0.314 and 0.8280
TA298	Refractive errors including astigmatism, myopia and presbyopia	0.353 and 0.991
TA301	Macular oedema and retinal vein occlusion	0.245 and 0.920
TA305	Macular oedema and retinal vein occlusion	0.469 and 0.828
TA346	Macular oedema and retinal vein occlusion	0.26 and 0.86
TA349	Macular oedema and retinal vein occlusion	Not reported
TA369	Eye conditions: general and other	Not relevant
TA409	Macular oedema and retinal vein occlusion	0.29 and 0.83
TA460	Eye conditions: general and other	0.353 and 0.869
TA467	Corneal conditions	Not relevant
TA486	Refractive errors including astigmatism, myopia and presbyopia	Not reported
TA532	Corneal conditions	Not relevant

Abbreviations**:** TA, Technology Appraisal.

Section 5, p.143

presented in Rentz et al. (433354). The resulting values used in the ERG's analysis are presented in Table 40.

Table 40: ERG analysis utility values

Health state	Values based on value from Rentz et al.	Values based on value from Rentz et al. (UK only)	Values using health state 433354 for Health State 5
HS1 (Moderate VI)	0.717	0.687	0.717
HS2 (Severe VI)	*0.624*	*0.581*	*0.638*
HS3 (Profound VI)	*0.530*	*0.476*	*0.560*
HS4 (CF)	*0.437*	*0.370*	*0.481*
HS5 (HM, LP, NLP)	0.343	0.264	0.402

Abbreviations**:** CF, counting fingers; HM, hand motion; LP, light perception; NLP, no light perception; SD, standard deviation; VI, visual impairment.
Note**:** Interpolated values shown in *italics*.

The values produced in the analysis based on the study by Rentz *et al*.[89] are clearly imperfect, however a strength of the study is that the descriptions (shown above for 333322) are described via the functional impact of vision problems, as opposed to being linked to VA alone as in many other conditions. Importantly however when valued by 600 members of the general public, the results indicated a poor but plausible utility for blindness (0.343 for all patients, 0.2644 for UK patients), as opposed to a 'worse than death' health state.

5.2.7.2 Adverse event disutilities

The company submission includes disutilities for three adverse events; cataract (-0.14 for 1 month), eye inflammation (-0.30 for 3.6 months), and increased intraocular pressure (-0.10 for 1 month). Both cataract and eye inflammation were referenced to previous macular degeneration submissions, with a reference to the literature for increased intraocular pressure.

ERG Comment:

The company's approach to accounting for the impact of adverse events on HRQoL appears broadly acceptable, though the disutility for eye inflammation appears to be particularly large, especially when patients already have relatively low health-state utilities (versus the general population). Nevertheless, the ERG maintains this assumption in the preferred base case, given the lack of an alternative value that may instead be used.

Section 5, p.178

Table 63: Summary of the ERG's exploratory and sensitivity analyses (including PAS)

Arm	Costs	QALYs	Inc. Costs	Inc. QALYs	ICER
ERG's preferred base case (all changes combined)					
BSC	£35,731	12.9			
VN	[redacted]	16.9	[redacted]	4.0	[redacted]
Duration of treatment effect per Institute for Clinical and Economic Review analysis					
BSC	£35,731	12.9			
VN	[redacted]	15.0	[redacted]	2.1	[redacted]
Remove all healthcare resource use costs					
BSC	£0	12.9			
VN	[redacted]	16.9	[redacted]	4.0	[redacted]
Use company-preferred healthcare resource use costs					
BSC	£48,254	12.9			
VN	[redacted]	16.9	[redacted]	4.0	[redacted]
UK utility values (based on Rentz et al.)					
BSC	£35,731	11.4			
VN	[redacted]	15.9	[redacted]	4.5	[redacted]
Alternative (higher) utility values (based on Rentz et al.)					
BSC	£35,731	13.8			
VN	[redacted]	17.1	[redacted]	3.3	[redacted]
Baseline characteristics derived from Study 301/302					
BSC	£35,667	12.4			
VN	[redacted]	16.5	[redacted]	4.1	[redacted]
Baseline characteristics derived from RPE65 NHx					
BSC	£35,773	13.1			
VN	[redacted]	17.0	[redacted]	3.9	[redacted]

Abbreviations**:** BSC, best supportive care; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; VN, voretigene neparvovec.

Section 8, p.185-186

8.1.2 Current management of RP

NHS England reported that because there are no specific genetic treatments available in England, current management for affected patients is supportive and involves ensuring good liaison between clinical and educational care together with low vision aids as appropriate for children. For affected adults, treatment is also supportive between clinical care, employers and social services. Low visual aids are provided for adults. Genetic counselling is provided via medical genetic services to affected families.

8.1.3 VN

NHS England stated that treatment with VN would provide the first treatment option for patients with the aim of stabilising vision and preventing further visual loss. The impact would be to improve mobility and independence for those patients very poor vision. In addition if treatment with VN is given earlier in the course of the disease, NHS England stated there is the potential to preserve central vision. A clinician from the Royal College of Ophthalmologists expressed a view that the most important outcome in the assessment of VN is gain of navigation, which will likely have a significant effect on the independence of affected patients. Expert advisors to the ERG are in agreement with this view. A clinician from the Royal College of Ophthalmologists expressed the view that side-effects were unlikely to be a barrier to adoption of the treatment, again a view endorsed by clinical advisors to the ERG. Both the clinician from Royal College of Ophthalmologists and clinical advisors to the ERG consider that adverse reactions caused by a short course of steroids administered post-operatively (e.g. red eye, transient blurred vision, etc.) would be no more than expected or from a similar eye operation.

8.1.3.1 Subgroups

A clinician from the Royal College of Ophthalmologists stated that while all patients with RP and some retained retinal structure might benefit from treatment with VN to some extent irrespective of age, there are a subgroup of patients with hypomorphic alleles giving a later less severe recessive phenotype who may have a different prognosis from the typical patient.[100] There are also a subgroup of patients with a dominant allele giving rise to a very different phenotype that may have a different prognosis from the typical patient,[101] although these patients are not eligible for VN under its current marketing authorisation.

8.1.4 Changes to service delivery and resources required if VN is recommended

NHS England stated that because genetic networks are in place across England, patients with known molecular diagnoses who could benefit from treatment can be identified. A clinician from the Royal College of Ophthalmologists reported that diagnosis and monitoring uses technology that is standard in specialist clinics (imaging, psychophysics, and electrophysiology).

NHS England currently directly commissions specialised ophthalmology services including the treatment of ocular genetic disorders. NHS England state that these are best managed by specialist networks which provide multidisciplinary services including diagnosis, testing, counselling and imaging as well as treatment. NHS England anticipate that the treatment with VN can be implemented using the current clinical services available for ophthalmic medical genetic services and vitreoretinal services. This view is endorsed by a clinician from the Royal College of Ophthalmologists who reported that the surgery is standard i.e. not significantly different to present clinical vitrectomies and is within the capabilities of specialist units. The clinician stated that impact of VN on delivery will be limited as the number of patients affected is small and the treatment is relatively quick; i.e. it is a single treatment given to each eye in an operation that takes about one hour.

8.1.5 Conclusion

There are no specific treatments currently available in England for this small patient group and current management for affected patients is supportive. Treatment with VN would provide the first treatment option for patients with this condition with the aim of stabilising vision and preventing further visual loss and with the potential to preserve central vision if given early.

Clinical experts, both from the Royal College of Ophthalmologists and advisors to the ERG, agree that the most important outcome in the assessment of VN is gain of navigation, which will likely have a significant effect on the independence of affected patients. Furthermore, all the clinical experts agree that side-effects are unlikely to be a barrier to adoption of the treatment and that adverse reactions caused by a short course of steroids administered post-operatively (e.g. red eye, transient blurred vision, etc.) would be no more than expected or from a similar eye operation.

NHS England anticipate that the treatment with VN can be implemented using the current clinical services available for ophthalmic medical genetic services and vitreoretinal services. This view is endorsed by a clinician from the Royal College of Ophthalmologists who reported that the surgery is standard i.e. not significantly different to present clinical vitrectomies and is within the capabilities of specialist units.

8.2 Patient support group and patient submissions

Submissions were received from the Fight for Sight charity and a patient expert with the condition nominated by the Fight for Sight charity. The patient expert's statement was in keeping with the