CADTH

CADTH COMMON DRUG REVIEW

Clinical Review Report

VORETIGENE NEPARVOVEC (LUXTURNA)

(Novartis Pharmaceuticals Canada Inc.)

Indication: Vision loss, inherited retinal dystrophy

Service Line:	CADTH Common Drug Review
Version:	Final (with redactions)
Publication Date:	January 2021
Report Length:	108 Pages

CADTH

Disclaimer: The information in this document is intended to help Canadian health care decision-makers, health care professionals, health systems leaders, and policy-makers make well-informed decisions and thereby improve the quality of health care services. While patients and others may access this document, the document is made available for informational purposes only and no representations or warranties are made with respect to its fitness for any particular purpose. The information in this document should not be used as a substitute for professional medical advice or as a substitute for the application of clinical judgment in respect of the care of a particular patient or other professional judgment in any decision-making process. The Canadian Agency for Drugs and Technologies in Health (CADTH) does not endorse any information, drugs, therapies, treatments, products, processes, or services.

While care has been taken to ensure that the information prepared by CADTH in this document is accurate, complete, and up-to-date as at the applicable date the material was first published by CADTH, CADTH does not make any guarantees to that effect. CADTH does not guarantee and is not responsible for the quality, currency, propriety, accuracy, or reasonableness of any statements, information, or conclusions contained in any third-party materials used in preparing this document. The views and opinions of third parties published in this document do not necessarily state or reflect those of CADTH.

CADTH is not responsible for any errors, omissions, injury, loss, or damage arising from or relating to the use (or misuse) of any information, statements, or conclusions contained in or implied by the contents of this document or any of the source materials.

This document may contain links to third-party websites. CADTH does not have control over the content of such sites. Use of third-party sites is governed by the third-party website owners' own terms and conditions set out for such sites. CADTH does not make any guarantee with respect to any information contained on such third-party sites and CADTH is not responsible for any injury, loss, or damage suffered as a result of using such third-party sites. CADTH has no responsibility for the collection, use, and disclosure of personal information by third-party sites.

Subject to the aforementioned limitations, the views expressed herein are those of CADTH and do not necessarily represent the views of Canada's federal, provincial, or territorial governments or any third-party supplier of information.

This document is prepared and intended for use in the context of the Canadian health care system. The use of this document outside of Canada is done so at the user's own risk.

This disclaimer and any questions or matters of any nature arising from or relating to the content or use (or misuse) of this document will be governed by and interpreted in accordance with the laws of the Province of Ontario and the laws of Canada applicable therein, and all proceedings shall be subject to the exclusive jurisdiction of the courts of the Province of Ontario, Canada.

The copyright and other intellectual property rights in this document are owned by CADTH and its licensors. These rights are protected by the Canadian *Copyright Act* and other national and international laws and agreements. Users are permitted to make copies of this document for non-commercial purposes only, provided it is not modified when reproduced and appropriate credit is given to CADTH and its licensors.

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

CADTH

Table of Contents

Tables

Figures

Abbreviations

ADL	activities of daily living
CI	confidence interval
cd.s/m^2	candela second per square metre
CLIA	Clinical Laboratory Improvement Amendments of 1988
dB	decibel
EMA	European Medicines Agency
ETDRS	Early Treatment of Diabetic Retinopathy Study
FBC	Fighting Blindness Canada
FST	full-field sensitivity threshold
HRQoL	health-related quality of life
ITT	intention to treat
IRD	inherited retinal dystrophy
LCA	Leber congenital amaurosis
LogMAR	logarithm of the minimum angle of resolution
mITT	modified intention-to-treat
MLMT	multi-luminance mobility testing
OCT	optical coherence tomography
RP	retinitis pigmentosa
RPE65	retinal pigment epithelium 65 kDa protein
SAE	serious adverse event
SD	standard deviation
SE	standard error
TEAE	treatment-emergent adverse event
VA	visual acuity
VF	visual field
VFQ	Visual Function Questionnaire
VFQ-25	25-item Visual Function Questionnaire
vg	vector genome
WDAE	withdrawal due to adverse event

Executive Summary

An overview of the submission details for the drug under review is provided in Table 1.

Table 1: Submitted for Review

Item	Description
Drug product	Voretigene neparvovec (Luxturna), 5×10^{12} vector genomes per mL concentrate for solution for subretinal injection.
Indication	For the treatment of adult and pediatric patients with vision loss due to inherited retinal dystrophy caused by confirmed biallelic *RPE65* mutations and who have sufficient viable retinal cells. Disease-causing biallelic *RPE65* mutations should be confirmed by an accredited laboratory using validated assay methods.
Reimbursement request	As per indication
Health Canada approval status	NOC
Health Canada review pathway	Standard review
NOC date	October 13, 2020
Sponsor	Novartis Pharmaceuticals Canada Inc.

NOC = Notice of Compliance; RPE65 = retinal pigment epithelium 65 kDa protein.

Introduction

Inherited retinal dystrophy (IRD) consists of a broad group of genetic retinal disorders that are associated with progressive visual dysfunction. These conditions are caused by mutations in more than 260 different genes, including the retinal pigment epithelium 65 kDa protein (RPE65) gene.[1] For patients with IRD caused by biallelic mutations in the *RPE65* gene, common clinical diagnoses include Leber congenital amaurosis (LCA) and retinitis pigmentosa (RP).[2] As mutations in a gene are more directly linked to the underlying molecular pathogenesis, it is now considered more appropriate to categorize this group of disorders by the individual disease-causing gene (genotype).[3] Symptoms of LCA can become evident from as young as 2 to 3 months of age.[2,4] There is progressive, profound reduction of visual acuity (VA), concentric reduction of visual fields (VFs), night blindness (nyctalopia), and nystagmus.[2,4-6] Patients with a clinical diagnosis of RP tend to have a more variable onset and slower progression of disease than those with LCA. Symptoms of RP usually begin around 10 years of age and these patients tend to have better preservation of visual function compared with those with a clinical diagnosis of LCA.[7,8] Visual impairment in individuals with *RPE65*-mediated IRD can present at a range of ages, from infancy to adolescence, with the unavoidable progression to complete blindness occurring as early as the preschool years or as late as the third decade of life.[4,9] Due to the complexity of signs and symptoms, diagnosis based solely on clinical exam and electroretinogram is insufficient. A comprehensive assessment of medical and ocular history, clinical eye exams, optical coherence tomography (OCT) imaging, assessment of VFs, and an analysis of family history with genetic testing are essential to accurately diagnosing and classifying IRD.[10,11] Genetic testing in IRD plays an important role in improving the accuracy of diagnosis and prognosis and is critical for guiding treatment decisions and identifying appropriate patients for clinical trials.[10] The exact prevalence and incidence of *RPE65*-mediated IRD is uncertain. CADTH reviewers estimated that the number of individuals with *RPE65*-mediated IRD in Canada ranges from 129 to 378, as

noted in the CADTH Pharmacoeconomic Review Report. The sponsor estimated that the number of individuals with *RPE65*-mediated IRD in Canada ranges from 42 to 277.[12]

In Canada, there are currently no pharmacological treatment options for patients with IRDs. The clinical experts consulted by CADTH for this review indicated that the standard of care for patients with IRD is supportive in nature and focuses on monitoring, psychological support, mobility training, and visual rehabilitation to maintain the patients' ability to perform activities of daily living (ADL) and improve health-related quality of life (HRQoL).

The indication for voretigene neparvovec is for the treatment of adult and pediatric patients with vision loss due to IRD caused by confirmed biallelic *RPE65* mutations and who have sufficient viable retinal cells. Disease-causing biallelic *RPE65* mutations should be confirmed by an accredited laboratory using validated assay methods.[13] Voretigene neparvovec is a 5×10^{12} vector genomes (vg) per millilitre concentrate for solution for subretinal injection. The dose of voretigene neparvovec for each eye is 1.5×10^{11} vg.

The objective of this report was to perform a systematic review of the beneficial and harmful effects of voretigene neparvovec administered as a single dose of 1.5×10^{11} vg in each eye via subretinal injection for the treatment of adult and pediatric patients with vision loss due to IRD caused by confirmed biallelic *RPE65* mutations and who have sufficient viable retinal cells.

Stakeholder Engagement

The information in this section is a summary of input provided by the patient groups who responded to CADTH's call for patient input and from clinical expert(s) consulted by CADTH for the purpose of this review.

Patient Input

One patient group submission was received for this review, which was authored jointly by Canada's largest blindness organizations: Fighting Blindness Canada (FBC), the Canadian Council of the Blind (CCB), the CNIB Foundation, and Vision Loss Rehabilitation Canada (VLRC).

The submission is mainly oriented around a survey that was prepared in consultation with an independent researcher; the analysis was done by FBC and the submission was developed collaboratively by all of the submitting organizations.

Approximately 71% of the survey group respondents reported having RP, 4% reported having LCA, and the remaining reported a broader spectrum of IRDs. Seven individuals had the *RPE65* mutation treated by voretigene neparvovec, which in each case was confirmed by genetic testing. Even though VA was maintained for some participants, 67% of the overall group were diagnosed as legally blind, over 70% had moderate-to-severe low vision, and 12% had near or total blindness.

Among the survey participants, 54% responded that IRD has negatively affected their employment or school status, with 22% rating the impact of their disease on their ability to perform job or school responsibilities as very severe. Survey participants reported struggling through school, requiring assistive technologies, or having to quit schooling as a result of the daily challenges being too difficult to overcome. Maintaining a stable career was flagged as being equally challenging.

Patients indicated that their eyesight interferes to some degree with most daily activities, including mobility and getting around, hobbies and leisure, socializing and interacting with others, looking after their appearance, reading a book or newspaper, and using a phone or iPad. Patients worry about their condition getting worse; struggle with challenges presented by daily activities, including parenting; experience long wait times for appointments; feel anxiety and uncertainty about the future and the impact of their diseases on their families; and, in some cases, are impacted by a lack of meaningful work, education, or social life.

Patients expressed a desire for a cure for the condition entirely, improved night vision and mobility at night, or facilitation of regular day-to-day activities such as social interactions, maintaining personal relationships, work, and study. Even if the treatments were to only enhance vision and mobility at night, most survey participants reported this would improve their overall quality of life. Many respondents also indicated that a stabilization of vision would be valuable, something that would at least halt progression if retinal damage cannot be reversed.

Clinician Input

The information in this section is based on input received from a panel of 6 clinical specialists consulted by CADTH for the purpose of this review.

Currently, in the absence of treatment for inherited retinal disease, the ultimate objective is to help each patient achieve their self-described life goals despite the limitation in their visual functioning. The most important outcome for patients is to maintain or improve their mobility, independence, and ability to achieve their life goals. The ideal treatment would be one that delays or stops disease progression and maintains the vision that patients had before they initiated therapy.

The clinical experts indicated there is currently no pharmacologic treatment available for the treatment of patients with vision loss due to IRD caused by confirmed biallelic *RPE65* mutations. The current treatment is supportive care where treatment of patients with progressive and inevitable vision loss revolves mainly around counselling the patient. Given there is currently no pharmacologic treatment available, clinical experts anticipated that voretigene neparvovec would cause a shift in the current treatment paradigm.

Patients best suited for treatment with voretigene neparvovec are those who are at least 4 years of age with sufficient viable retinal cells. Patients who are younger than 4 years of age may benefit from treatment with voretigene neparvovec; however, the surgery and administration of treatment in that age group is difficult and the assessment of treatment response is influenced by the inability to measure and monitor visual function, as the testing is not designed for young children under 4 years of age.

The most vulnerable patients with *RPE65* mutations appear to be those 10 years to 20 years of age, as patients in that age group lose considerable VF and start to lose their cone and rod functions. Therefore, it would be better to intervene before patients reach that age group, but age on its own should not be the only criteria for treating patients.

The clinical experts indicated that age on its own should not be the criterion for treating patients but rather the presence of sufficient viable retinal cells. The panellists also indicated that patients who might benefit from treatment should not be distinguished as having LCA or RP, as patients diagnosed with either LCA or RP caused by *RPE65* mutations are both suited for treatment with voretigene neparvovec.

The clinical experts indicated that measuring viable retinal cells is not a straightforward procedure. The panellists also indicated that there is no benchmark or threshold of viable retinal cells that can be used to objectively define the sufficient viable retinal cells criterion. The clinical experts also indicated that a numerical cut-off could not be applied universally throughout all OCT technologies and generations. In clinical practice, the presence of sufficient viable retinal cells would be determined by the treating physician using OCT examinations to measure the area of remaining viable photoreceptors, which would be supplemented by VA and visual function tests.

The clinical experts indicated that it is expected that treatment response will wane over time; however, it is uncertain how long the treatment effect will last and when the treatment effect will start waning.

The clinical experts indicated that a medical geneticist would be required to complete the genetic testing, and that the genetic diagnosis, patient selection, and pre- and post-surgical evaluations (safety and efficacy) should be confirmed by an inherited retinal disease specialist, and that treatment with voretigene neparvovec should be initiated and administered by a retinal surgeon experienced in performing sub-macular injection and managing its complications.

Clinical Evidence

Pivotal Studies and Protocol Selected Studies

Description of Studies

Study 301 (N = 31) was a phase III, open-label, randomized controlled trial designed to evaluate the efficacy and safety of sequential subretinal injection of voretigene neparvovec into each eye in patients diagnosed with LCA due to *RPE65* mutations. Randomization, which followed screening and confirmation of study eligibility, occurred in a 2:1 ratio of intervention (voretigene neparvovec) to control and used a block design stratified by age (≥ 10 years versus < 10 years) and mobility testing passing level (pass at ≥ 125 lux versus < 125 lux), as determined at screening. A total of 31 patients in 2 study sites in the US (the study enrolled international patients, including 1 patient from Canada) were randomized to either the treatment group (n = 21) or the control group (n = 10). Patients randomized to the voretigene neparvovec group received a dose of 1.5×10^{11} vg of voretigene neparvovec in each eye; the non-simultaneous subretinal injections were to occur within an 18-day period (12 days ± 6 days). Patients randomized to the control group did not receive voretigene neparvovec, sham injection, or corticosteroids for a period of at least 1 year from baseline evaluations. Following repeated retinal and visual function analysis, including mobility testing at 1 month, 3 months, 6 months, and 1 year, patients in the control group were crossed over to receive non-simultaneous injections of 1.5×10^{11} vg of voretigene neparvovec in each eye (within 18 days) after 1 year of randomization, provided they still met all study eligibility criteria. The primary end point was change in bilateral multi-luminance mobility testing (MLMT) performance at year 1 relative to baseline. Secondary end points were change in full-field sensitivity threshold (FST) averaged over both eyes at year 1 relative to baseline, change in assigned first eye MLMT performance at year 1 relative to baseline, and change in VA averaged over both eyes at year 1 relative to baseline.

In Study 301, there were more females (n = 18, 58%) than males (n = 13, 42%) and patients were primarily White (n = 21, 68%). At the time of randomization, the mean patient

age was 15.1 (standard deviation [SD] = 10.9) years, with a range of 4 years to 44 years. Patients randomized to the voretigene neparvovec group were slightly younger (mean = 14.7 years; SD = 11.8 years) than those randomized to the control group (mean = 15.9 years; SD = 9.5 years). More patients with better baseline MLMT performance (< 125 lux) were enrolled in the voretigene neparvovec group (57%) than in the control group (40%).

Efficacy Results

Treatment with voretigene neparvovec resulted in statistically significant improvements in navigational ability in low-to-moderate light conditions (measured by MLMT score). The mean bilateral MLMT change score after 1 year was 1.8 (SD = 1.1) for the voretigene neparvovec group and 0.2 (SD = 1.0) for the control group. The difference in change from baseline in bilateral MLMT between the voretigene neparvovec and control treatment groups at 1 year was 1.6 (95% confidence interval [CI], 0.72 to 2.41; P = 0.001) which was statistically significant in favour of voretigene neparvovec; however, the difference between the treatment groups did not exceed 2 points. Eleven patients (52%) in the voretigene neparvovec group had an MLMT score change of 2 or more (the difference considered meaningful by the FDA and the European Medicines Agency [EMA]). In contrast, only 1 patient (10%) in the control group had a score change of 2 and none of the patients in the control group had a score change greater than 2. Although 62% of patients in the voretigene neparvovec group achieved a score of 6 on the MLMT (the maximum possible score) following administration of voretigene neparvovec, compared with no patients in the control group achieving a score of 6 on the MLMT, the observed mean increase in MLMT score of 1.8 observed in the voretigene neparvovec group could be an underestimate of the within-groups magnitude of the change due to the potential ceiling effect. Improvements in the MLMT score observed at 1 year seemed to be maintained until the 4-year follow-up.

Patients treated with voretigene neparvovec experienced a mean improvement in FST greater than 2 log units, whereas mean FST did not change in the control group; the mean change from baseline to year 1 was −2.08 (standard error [SE] = 0.29) $\log_{10}$(candela second per square metre [cd.s/m^2]) for the voretigene neparvovec group and 0.04 (SE = 0.44) $\log_{10}$(cd.s/m^2) for the control group. There were statistically significant improvements in full-field light sensitivity with voretigene neparvovec (mean difference versus control = −2.11 log units; 95% CI, −3.19 to −1.04; P < 0.001) at 1 year. This between-group difference exceeded the suggested clinical significance threshold of 10 decibels (dB) or 1 log unit for clinical significance.[14] The improvements were sustained for 4 years after the second-eye injection, where for all patients who received treatment with voretigene neparvovec, the mean change from injection baseline at 4 years after the second-eye injection was −2.00 (SE = 1.35) $\log_{10}$(cd.s/m^2). The clinical experts explained that the changes seen would be clinically meaningful in terms of improving visual function.

A longitudinal repeated measures analysis of VA using the Holladay scale averaged over both eyes showed a mean change from baseline to 1 year of −0.16 (SE = 0.07) logarithm of the minimum angle of resolution (LogMAR) for the voretigene neparvovec group and 0.01 (SE = 0.10) LogMAR for the control group, resulting in a mean treatment-effect difference of −0.16 LogMAR (95% CI, −0.41 to 0.08; P = 0.17), an 8-letter improvement. This difference was neither statistically significant nor clinically meaningful. Using the Lange scale for off-chart VA results, the mean between-group treatment difference was −0.16 LogMAR (95% CI, −0.31 to −0.01; nominal P = 0.035) corresponding to a 7.5-letter improvement on the eye chart for people who had voretigene neparvovec. All changes were smaller than the accepted clinically meaningful change (≥ 0.30 LogMAR). By year 3, little further change was seen in VA for either arm after treatment. The clinical experts consulted by CADTH for this

review explained that even a small change would be important for patients. The clinical experts also noted that even if there were no improvement, preventing vision deterioration would be important for the patient's quality of life.

For the patient-completed Visual Function Questionnaire (VFQ), the mean change from baseline to year 1 was 2.6 (SD = 1.8) for the voretigene neparvovec group and 0.1 (SD = 1.4) for the control group, for a mean between-group treatment difference of 2.4 (95% CI, 1.0 to 3.8; nominal P = 0.001). For the parent-completed surveys, the mean change from baseline to year 1 was 3.9 (SD = 1.9) for the voretigene neparvovec group and −0.2 (SD = 1.3) for the control group, for a mean between-group treatment difference of 4.0 (95% CI, 2.1 to 6.0; nominal P = 0.002). Although the VFQ assessed the patients' ability to perform ADL for those patients who received voretigene neparvovec, the questionnaire did not contain any items to specifically assess HRQoL for patients. In addition, the VFQ used in Study 301 was not assessed psychometrically and, given the modifications made to the original National Eye Institute VFQ, the minimal important differences (MIDs) identified in the literature for that measure were not considered directly generalizable to the version used in Study 301.

Long-term data for patients in the voretigene neparvovec group suggest durable improvements in visual performance across multiple end points for at least 4 years following voretigene neparvovec administration. Similarly, after crossing over to voretigene neparvovec, patients in the control/voretigene neparvovec group exhibited visual performance improvements comparable to those observed in the voretigene neparvovec group, effects that were maintained through at least 3 years following bilateral voretigene neparvovec administration. Study 301 aims to follow patients for up to 15 years after treatment. The current duration of follow-up is limited to 5 years and longer-term efficacy and safety data for voretigene neparvovec are awaited. The duration of treatment effect is unclear. Also, given the small sample size (30 patients), there is uncertainty around the generalizability of the results observed for the long-term treatment effect. The clinical experts indicated that voretigene neparvovec would likely provide long-term benefits, although this was associated with substantial uncertainty.

Subgroup analyses of interest to this review were based on age, clinical phenotype, and number of viable retinal cells. No preplanned subgroup analyses were conducted in Study 301. The sponsor conducted a post hoc subgroup analysis by age (younger than 18 years at first injection versus 18 years or older at first injection); however, this analysis should be considered hypothesis-generating. Also, it is not clear if this subgroup analysis was conducted at the request of a regulator, and no rationale was provided for the age cut-off. The clinical experts consulted by CADTH indicated that greater clinical benefit from treatment is expected earlier in the condition when there are more viable retinal cells for the gene replacement therapy to restore. While patient age can be used as a proxy to estimate how advanced the condition is, due to the heterogeneous nature of *RPE65* mutation–associated retinal dystrophy and the differences in the age of onset and disease progression, retinal cell viability should be assessed on an individual patient basis, regardless of patient's age (as long as the patient is at least 4 years of age), to determine if the specific individual is likely to respond to voretigene neparvovec.

Harms Results

All patients in Study 301 experienced at least 1 treatment-emergent adverse event (TEAE). Most adverse events were mild in severity and no patient had any adverse events that led to study discontinuation or death.

The most frequently reported TEAEs in the voretigene neparvovec group were leukocytosis (in 45% of patients); vomiting (in 40% of patients); nasopharyngitis, headache, and pyrexia (in 35% of patients for each); oropharyngeal pain, cough, and nausea (in 30% of patients for each); increased intraocular pressure (in 20% of patients); and cataract and hematuria in 15% of patients for each.

Overall, 13 (65%) patients in the voretigene neparvovec group had at least 1 TEAE considered to be related to the study drug administration procedure. The TEAEs most often considered to be probably related to the administration procedure were cataract and increased intraocular pressure (n = 3 [15%] patients for each).

During the control period, 2 (10%) patients in the voretigene neparvovec group experienced 3 serious adverse events (SAEs) at time points distant from vector administration. One patient experienced a possible seizure requiring hospitalization and 1 patient experienced an adverse reaction to medications administered during oral surgery, which required hospitalization.

At the time of the data cut-off for the Clinical Study Report (CSR), which provided updated results of safety data through July 2, 2018, including follow-up for up to 5 years after the second injection for some patients, 6 SAEs occurred in 5 patients, including convulsion (1 event), adverse drug reactions (2 events) and retinal disorder (1 event of foveal thinning and loss of vision), retinal detachment (1 event), pneumonia (1 event), and menorrhagia (1 event).

One ocular SAE occurred in Study 301, where a patient with pre-existing atrophy of the retina who received voretigene neparvovec experienced a retinal disorder (foveal thinning and a loss of central vision) related to the subretinal injection of the treatment.

In the voretigene neparvovec group during the first year after the randomization period, 3 patients experienced a cataract, 2 patients experienced a retinal tear, and 1 patient developed an asymptomatic full-thickness macular hole. During the follow-up, 1 patient who was originally in the control group and crossed over to voretigene neparvovec experienced a retinal disorder (foveal thinning and a loss of central vision). Another patient experienced 1 SAE of retinal detachment.

Table 2: Summary of Key Results From Study 301

	Voretigene neparvovec N = 21	Control (N = 10)
Bilateral MLMT performance at 1 year compared with baseline (ITT)		
Number of patients contributing to the analysis	21	10
Change from baseline, mean (SD)	1.8 (1.1)	0.2 (1.0)
Treatment group difference versus control (95% CI)	1.6 (0.72 to 2.41)	Reference
P value	0.001[a]	
Monocular MLMT performance at 1 year compared with baseline for the first assigned eye (ITT)		
Number of patients contributing to the analysis	21	10
Change from baseline, mean (SD)	1.9 (1.2)	0.2 (0.6)
Treatment group difference versus control (95% CI)	1.7 (0.89 to 2.52)	Reference
P value	0.001[a]	
Full-field light sensitivity testing: white light (log_{10}[cd.s/m^2]) at 1 year (ITT)[b,c]		
Number of patients contributing to the analysis	20	9
Baseline, mean (SE)	−1.29 (0.09)	−1.65 (0.14)

	Voretigene neparvovec N = 21	Control (N = 10)
At 1 year, mean (SE)	−3.36 (0.28)	−1.61 (0.42)
Change from baseline, mean (SE)	−2.08 (0.29)	0.04 (0.44)
Treatment group difference versus control (95% CI)	−2.11 (−3.19 to −1.04)	Reference
P value	< 0.001	
Visual acuity (LogMAR) using Holladay scale for off-chart results at 1 year (ITT)[d]		
Number of patients contributing to the analysis	21	10
Baseline, mean (SE)	1.18 (0.14)	1.29 (0.21)
At 1 year, mean (SE)	1.03 (0.17)	1.30 (0.25)
Change from baseline, mean (SE)	−0.16 (0.07)	0.01 (0.10)
Treatment group difference versus control (95% CI)	−0.16 (−0.41 to 0.08)	Reference
P value	0.17	
VFQ average score (patient)		
Number of patients contributing to the analysis	21	9
Baseline, mean (SD)	4.4 (1.4)	4.9 (1.5)
At 1 year, mean (SD)	7.0 (1.9)	5.1 (1.8)
Change from baseline, mean (SD)	2.6 (1.8)	0.1 (1.4)
Treatment group difference versus control (95% CI)	2.4 (1.0 to 3.8)	Reference
P value	0.001[ef]	
VFQ average score (parent)		
Number of patients contributing to the analysis	15	5
Baseline, mean (SD)	3.6 (1.3)	3.3 (1.7)
At 1 year, mean (SD)	7.5 (1.5)	3.1 (1.8)
Change from baseline, mean (SD)	3.9 (1.9)	−0.2 (1.3)
Treatment group difference versus control (95% CI)	4.0 (2.1 to 6.0)	Reference
P value	0.002[e,f]	
Harms, n (%) (safety analysis set)		
AEs	20 (100)	9 (100)
SAEs	2 (10)	0
WDAE	0	0
Deaths	0	0
Notable harms, n (%)		
Retinal tear	2 (10)	0
Macular changes	1 (5)	0
Cataract	3 (15)	0

AE = adverse event; cd.s/m^2 = candela second per square metre; CI = confidence interval; ITT = intention to treat; LogMAR = logarithm of the minimum angle of resolution; MLMT = multi-luminance mobility testing; SAE = serious adverse event; SD = standard deviation; SE = standard error; VFQ = Visual Function Questionnaire; WDAE = withdrawal due to adverse event.

[a] The observed 2-sided P value is from a Wilcoxon rank-sum test using an exact method. The permutation test P value was computed from all possible permutations.

[b] All measures were averaged over both eyes and then analyzed.

[c] Changes, 95% CIs, and P values were estimated using a repeated measures model with time, treatment, and time by treatment interaction.

[d] All measures were averaged over both eyes and then analyzed. Changes, 95% CIs, and P values were estimated using a repeated measures model with time, treatment, and time by treatment interaction.

[e] The observed 2-sided P value is from a Wilcoxon rank-sum test.

[f] P value was not adjusted for multiple testing.

Source: Clinical Study Report for Study 301.[15]

Critical Appraisal

Study 301 used accepted methods to conceal allocation where a randomization list was generated by an independent biostatistician. However, because it was an open-label trial, patients were aware of the treatment allocation following randomization. Therefore, the evaluation of patient-reported outcomes (such as those measured by VFQ) and adverse events may be biased by treatment knowledge. The treatment effect on these subjective outcomes can potentially be overestimated as a consequence of the patient's expectation of the efficacy of a new drug. Certain steps have been taken to ensure appropriate blinding of the assessment of the primary outcome measure (MLMT) and it seems there is potential for an unbiased outcome assessment with the MLMT, despite the open-label design. On the other hand, while the trial states that orientation and mobility assessors were masked, there is insufficient detail provided in the CSR to judge if adequate blinding of the outcome assessment was performed for all secondary outcome measures (e.g., VF and VA).

Imbalances in the baseline patient characteristics between the voretigene neparvovec and control groups included age and visual performance. It is unclear whether differences in baseline age between the voretigene neparvovec and control groups may introduce a risk of bias; on the one hand, the clinical experts indicated that age on its own should not be a criterion for treating patients but rather the presence of sufficient viable retinal cells. The imbalance at baseline in MLMT performance following assignment to the voretigene neparvovec or control group may bias the observed treatment effect estimate; however, the direction of the bias is unclear. As differences in MLMT, VA, and VF were noted at baseline in Study 301 — and as these outcomes represent the key clinical data — there is uncertainty associated with the effect estimate for these outcomes given that patients in the voretigene neparvovec treatment group had better MLMT scores at baseline (a larger percentage of patients with a baseline MLMT performance < 125 lux were enrolled in the voretigene neparvovec group [57%] than in the control group [40%]). Also, at baseline, patients randomized to the voretigene neparvovec group had better acuity, on average (less VA loss), and better VF than patients randomized to the control group. However, there is no evidence to indicate how visual performance at baseline could affect the treatment effect and, if bias exists, it is not possible to judge the direction of the bias.

Scores on the MLMT may underestimate the treatment effect of voretigene neparvovec due to the potential ceiling effect, where patients who passed the test at the second-lowest light level at baseline were able to achieve a maximum increase of only 1 unit. Although 62% of patients in the voretigene neparvovec group achieved the maximum possible increase following administration, compared with none in the control group, the observed mean increase in MLMT score of 1.8 may be an underestimate of the treatment effect because of this ceiling effect. The MLMT also potentially has a floor effect, where 1 patient in the voretigene neparvovec group did not pass the MLMT at the highest light level at baseline and also failed at year 1. It is worth noting that of the 5 patients who did not pass screening, 2 (40%) were not eligible based on mobility test performance, 1 was excluded due to ceiling effect, and 1 was excluded due to floor effect. In addition, it is not clear whether the 12 unique MLMT course configurations were of equivalent difficulty.

While patients in Study 301 were required to have a diagnosis of LCA due to *RPE65* mutations to be enrolled in the pivotal trial of voretigene neparvovec, the clinical experts indicated the results reported for these patients should be generalizable to patients with RP as long as they have confirmed biallelic *RPE65* mutations. Voretigene neparvovec is designed to deliver a normal copy of the *RPE65* gene to cells of the retina in patients with a

reduced level of or no biologically active RPE65; it is therefore intended to treat the underlying mechanism of the disease, which is the same in patients with confirmed *RPE65* mutations regardless of clinical phenotype.[13]

The clinical experts indicated that measuring viable retinal cells is not a straightforward procedure, and that the methods used in the pivotal trial to determine whether a patient has sufficient viable retinal cells do not give a complete picture of the health of the retina or the number of viable photoreceptors. They also indicated that the technology to assess the structure and function of the retina has evolved since the trial. The clinical experts indicated that OCT can measure the thickness of the retina; however, it may not inform the treating physician whether there are viable retinal cells.

Indirect Comparisons

No indirect comparisons were submitted by the sponsor or identified by CADTH.

Other Relevant Evidence

Description of Studies

The sponsor submitted 2 supportive trials, Study 101 and Study 102. Study 101 (N = 12) was a phase I, open-label, single-arm, dose-escalation study that assessed the safety and tolerability of 3 different doses of voretigene neparvovec administered via subretinal administration to 1 eye (first-treated eye) of patients with LCA due to *RPE65* mutations. Study 101 also evaluated the clinical efficacy; however, no formal hypothesis testing was conducted. Study 102 was a follow-on to Study 101 in which patients received voretigene neparvovec treatment in the previously uninjected eye (second-treated eye). In Study 101, 12 patients, who were 8 years of age or older at the time of administration, received unilateral subretinal injection in the eye with the worse function (first-treated eye). Three doses of voretigene neparvovec were tested sequentially: 1.5×10^{10} vg, 4.8×10^{10} vg, and 1.5×10^{11} vg. The dose of voretigene neparvovec under review is 1.5×10^{11} vg. Study 102 was a follow-on study to Study 101. Eleven of the 12 treated patients in Study 101 received a subretinal injection in the contralateral eye (second-treated eye) consisting of 1 dose of 1.5×10^{11} vg of voretigene neparvovec in a total volume of 300 µL.

Efficacy Results

All efficacy data are descriptive, and no inferences can be made because there were no hypotheses being tested.

In Study 101, 4 patients were considered non-evaluable, given the inconsistent use of patching as well as the variability of lighting conditions and test-course difficulty. Following vector administration, follow-up mobility testing using the injected eye indicated that 4 of the 8 evaluable patients were able to complete the mobility test at a light level that was at least 1 level darker than baseline.

In Study 102, monocular mobility testing was assessed for change from baseline for the eye injected in Study 102. Of the 11 patients who received voretigene neparvovec, 8 were considered evaluable for mobility testing, while three were not and, therefore, were not included for presentations of this parameter. All 8 evaluable patients whose eyes had been injected in Study 102 completed the mobility test 1 year after injection at a light level that was at least 1 level darker than baseline. Five of the 8 (63%) evaluable patients received the maximum attainable score. Mobility testing results continued to show that, through year

4, all 8 evaluable patients completed the mobility test at a light level that was at least 1 level darker than baseline.

For the eye injected in Study 101 at year 1, 4 of 7 (57%) evaluable patients showed a decrease of 10 dB or more in FST compared with baseline, with a change from baseline in all 7 evaluable patients ranging from −8.1 dB to −33.7 dB. At year 2, 3 of 7 (43%) evaluable patients showed a decrease of 10 dB or more in FST, with a change from baseline ranging from −5.7 dB to −29.1 dB in 6 of the 7 evaluable patients.

At the time of database lock for the CSR for Study 102 (May 19, 2019), the available FST data indicated that the light sensitivity of the eye injected in Study 102 increased after injection for 8 of 11 patients; this increase was greater than the 10 dB cut-off, was considered clinically important in 7 of 11 patients, and was below this cut-off for 1 patient. Light sensitivity remained stable in the other 3 patients.

In Study 101 at year 1, for the eye injected in that study, 7 of the 12 (58%) evaluated patients had a change in LogMAR score of 0.3 or more compared with baseline (corresponding to an improvement of at least 3 lines [15 letters] on the eye chart). At year 2, 4 of 11 (36%) patients had a change in LogMAR score of 0.3 or more compared with baseline. At year 3, 5 of 9 (56%) patients had a change in LogMAR score of 0.3 or more compared with baseline.

In the Study 102 LogMAR scores, based on Holladay scale for off-chart results, 1 (9%) of the 11 patients assessed had a change in LogMAR score of 0.3 or more compared with baseline 1 year after receiving treatment.

Harms Results

All patients in Study 101 and Study 102 experienced at least 1 TEAE.

One SAE was reported in Study 101, where 1 patient experienced an anal fistula that required hospitalization. The SAE was mild in severity and considered related to an underlying diagnosis of inflammatory bowel disease. The event was recovered and/or resolved with no sequelae.

One SAE was reported in Study 102, which was elevated intraocular pressure (grade 4) in the patient's right eye resulting in hospitalization. The SAE, which was moderate in severity, was deemed related to the use of a depo-steroid injection for a known rare complication of vitrectomy (endophthalmitis).

No deaths or discontinuations from the study due to adverse events were reported in either study.

Critical Appraisal

Studies 101 and 102 are phase I, single-arm, open-label, non-randomized studies with a small sample size and are not considered to provide high-quality evidence to support the efficacy of voretigene neparvovec.

No conclusions can be drawn regarding the clinical efficacy of voretigene neparvovec. Study 101 was a dose-escalating study, and neither Study 101 nor Study 102 were designed or powered to assess the clinical efficacy of voretigene neparvovec. No formal hypothesis testing was conducted, and the within-group changes were not designed to be inferential. Only descriptive statistics were presented.

During the course of Study 101, the mobility test was further refined and standardized, which affected both the number of patients considered evaluable and the interpretation of the results.

The dosage of voretigene neparvovec used in 9 of the 12 patients enrolled in Study 101 was not the dosage approved by Health Canada; hence, the generalizability of the study results to the Canadian patient population is unclear.

Compared with Study 101, Study 301 included a broader population, including younger patients (4 years of age and older in Study 301 versus 8 years of age and older in Study 101) and patients with less advanced disease (VA no better than 20/60 in Study 301 versus VA no better than 20/160 in Study 101). The more stringent criteria introduced in Study 102 for determining the number of viable retinal cells (≥ 3 disc areas of retina without atrophy or pigmentary degeneration within the posterior pole) was also used in Study 301, whereas in Study 101, patients were eligible if they had one or more disc areas of retina that were not involved in complete retinal degeneration. Finally, periocular injection of the various corticosteroids used in Study 101 and Study 102 was discontinued in Study 301 to decrease the incidence and severity of elevated intraocular pressure and cataract formation or progression.

Conclusions

Currently, there is no pharmacologic treatment available for the treatment of patients with vision loss due to IRD caused by confirmed biallelic *RPE65* mutations; the current standard of care is supportive in nature and focuses on monitoring, psychological support, mobility training, and visual rehabilitation to maintain the patients' ability to perform ADL and improve HRQoL. Based on the results of 1 phase III study (Study 301), voretigene neparvovec, compared with the control group, demonstrated a statistically significant improvement in functional vision under dim light conditions, as measured by the MLMT one year post treatment. Voretigene neparvovec also resulted in a statistically significant improvement in FST one year post treatment. No improvement was observed in VA and the effect of voretigene neparvovec on HRQoL is unknown. The improvements observed with voretigene neparvovec after one year appear to be maintained for up to 4 years; however, further data are required and there is uncertainty about the duration of treatment effect. Harms, although present, were related to the surgical aspects of administration; the most common ocular adverse events related to the administration procedure were conjunctival hyperemia, cataract, increased intraocular pressure, retinal tear, dellen, macular hole, subretinal deposits, eye inflammation, eye irritation, eye pain, and maculopathy; 2 ocular SAEs had severe consequences.

Introduction

Disease Background

IRD consists of a broad group of genetic retinal disorders that are associated with progressive visual dysfunction. These conditions are caused by mutations in more than 260 different genes, including the *RPE65* gene.[1] These genetic retinal disorders used to be grouped together based on their clinical manifestations and findings (phenotype). However, they are clinically heterogeneous and vary widely in their pathogenesis, progression, and mutation inheritance.[10,16] For patients with IRD caused by biallelic mutations in the *RPE65* gene, common clinical diagnoses include LCA and RP.[2] As mutations in a gene are more directly linked to the underlying molecular pathogenesis, it is now considered more appropriate to categorize this group of disorders by the individual disease-causing gene (genotype).[3]

Regardless of the clinical diagnoses, IRD caused by biallelic *RPE65* mutations eventually leads to complete blindness.[4,17-20] Symptoms of LCA can become evident from as young as 2 to 3 months of age.[2,4] There is a progressive and profound reduction of VA, concentric reduction of VFs, night blindness (nyctalopia), and nystagmus.[2,4-6] Patients with a clinical diagnosis of RP tend to have a more variable onset and slower progression of disease than those with LCA. Symptoms of RP usually begin around 10 years of age and these patients tend to have better preservation of visual function compared with those with a clinical diagnosis of LCA.[7,8] Visual impairment in individuals with *RPE65*-mediated IRD can present at a range of ages, from infancy to adolescence, with the unavoidable progression to complete blindness occurring as early as the preschool years or as late as the third decade of life.[4,9] The first symptom of the disease is night blindness, resulting in difficulty seeing in dim light, such as at dusk or at night.[7,8] Individuals with normal night vision can almost fully adapt to dim light in 15 to 30 minutes, whereas patients with *RPE65*-mediated IRD take much longer or are unable to adapt at all.[8] The progressive nature of *RPE65*-mediated IRD is well documented, with profound deterioration over time in both VF and VA. There is no evidence of spontaneous sustained improvement in measures of visual function in any individual.[9] Retinal sensitivity also declines with age, with patients ultimately losing the ability to detect light of any intensity.[21]

A diagnosis of biallelic *RPE65*-mediated IRD is typically suspected in patients who present with signs and symptoms suggestive of certain early onset, rod-mediated IRD. However, due to the complexity of signs and symptoms, a diagnosis based solely on clinical exam and electroretinogram is insufficient. A comprehensive assessment of medical and ocular history, clinical eye exams, OCT imaging, assessment of VFs, and an analysis of family history with genetic testing are essential in accurately diagnosing and classifying IRD.[10,11] In the natural history study by Chung et al.,[9] the 70 patients with confirmed biallelic *RPE65*-mediated IRD had initially received 20 distinct clinical diagnoses. On average, each patient received 3 diagnoses over their care, and clinical diagnoses of both RP and LCA were made in 13% of the patients.[9] For this reason, clinical diagnoses without molecular confirmation are insufficient to establish definitive diagnoses of *RPE65*-mediated IRD.[11] Genetic testing in IRD plays an important role in improving the accuracy of diagnosis and prognosis and is critical for guiding treatment decisions and identifying appropriate patients for clinical trials.[10] Genetic testing is recommended for patients with presumed genetically caused IRD.[22,23] Such testing may involve single-gene analyses, panel-based tests that

include many IRD disease genes, or exome and genome sequencing of DNA derived from blood or saliva samples.[24] A summary of *RPE65* mutation testing is provided in Appendix 5.

IRDs typically manifest in children and young people, with more than half of patients having severe visual impairment before adulthood, which has a profound impact on the HRQoL of patients and their families as well as on the patients' ability to perform ADL.[9,25] Patients with visual impairment and blindness rate their HRQoL similar to patients who have AIDS, a stroke, or kidney transplant.[26] Individuals who are blind attach high value to vision loss and are willing to give up 60% of their lifetime to regain normal VA.[26]

The exact prevalence and incidence of *RPE65*-mediated IRD is uncertain. No Canadian-specific data were identified. Fighting Blindness Canada (FBC) has estimated that RP affects between 1 in 3,500 to 1 in 4,000 people,[7] based on US and European estimates. The mutations in the *RPE65* gene account for 0.81%[27] to 1.85%[17] of RP diagnoses. LCA is less common: FBC estimates a prevalence of from 1 in 33,330 to 1 in 50,000 people,[28] based on US estimates. Mutations in the *RPE65* gene account for 6.8%[29] to 15.5%[29] of LCA diagnoses. Using these prevalence rates for LCA and RP, and the percentage of patients with the *RPE65* gene, CADTH Common Drug Review reviewers estimated that the number of individuals with *RPE65*-mediated IRD in Canada ranges from 129 to 378, as noted in the CADTH Pharmacoeconomic Review Report. The sponsor estimated that the number of individuals with *RPE65*-mediated IRD in Canada ranges from 42 to 277.[12]

Standards of Therapy

In Canada, there are currently no pharmacological treatment options for patients with IRDs. The clinical experts consulted by CADTH for this review indicated that the standard of care for patients with IRD is supportive in nature and focuses on monitoring, psychological support, mobility training, and visual rehabilitation to maintain the patients' ability to perform ADL and improve HRQoL. Depending on the level of visual impairment, low-vision services such as the use of guide dogs, low-vision aids that maximize existing vision, assistive devices (e.g., canes), and technology (e.g., Braille embossers) may help individuals compensate for the progressive vision loss that is associated with this disease.[30] However, none of these options treat the underlying disease mechanism or alter the natural history of IRD.

Drug

The indication for voretigene neparvovec is for the treatment of adult and pediatric patients with vision loss due to IRD caused by confirmed biallelic *RPE65* mutations and who have sufficient viable retinal cells. Disease-causing biallelic *RPE65* mutations should be confirmed by an accredited laboratory using validated assay methods.[13]

The indication for voretigene neparvovec in the US is for the treatment of patients with confirmed biallelic *RPE65* mutation-associated retinal dystrophy. Patients must have viable retinal cells as determined by the treating physician(s).[31] In Europe, the indication is similar to the indication in Canada.[32] The sponsor's reimbursement request is similar to the Health Canada indication.

Voretigene neparvovec is a vg/mL concentrate for solution for subretinal injection. The dose of voretigene neparvovec for each eye is 1.5×10^{11} vg. Each dose is to be delivered into the subretinal space in a total volume of 0.3 mL. The individual administration procedure for

each eye is to be performed on separate days within a close interval, but no fewer than 6 days apart.[13]

Starting 3 days prior to the administration of voretigene neparvovec in the first eye, it is recommended that an immunomodulatory regimen be initiated at a dose equivalent to 1 mg/kg per day (maximum of 40 mg per day) of prednisone and continued at the same dose for 4 more days (including the day of administration). This is followed by a tapering dose during the next 10 days at a dose equivalent to 0.5 mg/kg per day (maximum of 20 mg per day) of prednisone for the first 5 days and 0.5 mg/kg of prednisone every other day (maximum of 20 mg per day) for the last 5 days. Initiation of the immunomodulatory regimen for the second eye should follow the same schedule and supersede completion of the immunomodulatory regimen of the first eye. Prior to initiation of the immunomodulatory regimen and prior to administration of voretigene neparvovec, the patient must be checked for symptoms of active infectious disease of any nature and, in the event of such infection, the start of treatment must be postponed until after the patient has recovered.[13]

Voretigene neparvovec is a vector-based gene therapy designed to deliver a normal copy of the gene encoding the *RPE65* gene to cells of the retina in persons with reduced or absent levels of biologically active RPE65. RPE65 is produced in the retinal pigment epithelial cells and converts all-*trans*-retinol to 11-*cis*-retinol, which subsequently forms the chromophore, 11-*cis*-retinal, during the visual (retinoid) cycle. The visual cycle is critical in phototransduction, which refers to the biological conversion of a photon of light into an electrical signal in the retina. Mutations in the *RPE65* gene lead to reduced or absent levels of RPE65 isomerohydrolase activity, blocking the visual cycle, resulting in impairment of vision and, ultimately, complete blindness. Voretigene neparvovec is a modified virus that contains a working copy of the *RPE65* gene. After injection, it delivers this gene into the cells of the retina, the layer at the back of the eye that detects light. This enables the retina to produce the proteins needed for vision. The virus that is used to deliver the gene does not cause disease in humans.[13]

Voretigene neparvovec is contraindicated in patients who are hypersensitive to voretigene neparvovec or to any ingredient in the formulation; it is also contraindicated in patients with ocular or periocular infection, and in patients with active intraocular inflammation.[13]

Voretigene neparvovec has the following warnings and precautions in the product monograph:[13]

- Endophthalmitis may occur following any intraocular surgical procedure or injection.
- Permanent decline in VA may occur following subretinal injection of voretigene neparvovec.
- Retinal abnormalities may occur during or following vitrectomy, including retinal tears or epiretinal membrane or retinal detachment.
- Increased intraocular pressure may occur after subretinal injection of voretigene neparvovec.
- Patients should avoid air travel or travel to high elevations until the air bubble formed following administration of voretigene neparvovec has completely dissipated from the eye. A rapid increase in altitude while the air bubble is still present can cause a rise in eye pressure and irreversible vision loss.
- Subretinal injection of voretigene neparvovec, especially vitrectomy surgery, is associated with an increased incidence of cataract development and/or progression.

Stakeholder Engagement

Patient Group Input

This section was prepared by CADTH staff based on the input provided by patient groups.

About the Patient Groups and Information Gathered

One patient group submission was received for this review, which was authored jointly by Canada's largest blindness organizations: Fighting Blindness Canada (FBC), the Canadian Council of the Blind (CCB), the CNIB Foundation, and Vision Loss Rehabilitation Canada (VLRC). All 4 organizations have close ties to the vision loss community and a vested interest in ensuring patient views are comprehensively integrated into the health technology assessment process. Founded in 1974, FBC is a private funder of ophthalmology research in Canada and is aimed at raising and directing funds to accelerate the development and availability of treatments and cures. The CCB was founded in 1944 by returning blind veterans and schools of the blind. It is a membership-based registered charity that is focused on improving quality of life for those with vision loss through awareness, peer mentoring, socializing, sports, advocacy, health promotion, and illness prevention. The CNIB Foundation is a non-profit organization that is dedicated to delivering programs and advocacy to empower people impacted by blindness. VLRC is a provider of rehabilitation therapy for people with vision loss, helping to maximize their remaining vision by promoting safety and enhancing essential skills for daily activities.

The submission is mainly oriented around a survey that was prepared in consultation with an independent researcher; the analysis was done by FBC and the submission was developed collaboratively by all of the submitting organizations. With the exception of VLRC, all of the organizations received funding of more than $50,000 from Novartis, though each specified that the funding was for educational programs and not directed specifically toward this patient input submission.

In addition to the survey, findings from 2 case studies were reported: 1 conducted in the UK and the Republic of Ireland that was aimed at understanding the socioeconomic implications of IRDs, and the other was a summarized interview with a Canadian parent whose child received Luxturna. Finally, a white paper containing over 300 survey responses from various stakeholders (patients, caregivers, researchers, clinicians, policy-makers, and more) was provided that captures some of the complexities of living with vision loss in Canada.

Disease Experience

Information in this section of the submission was collected through a 60-question online survey developed and hosted by the FBC. The online survey is part of a broader mixed-methods research project, consisting of survey data and qualitative interviews, aimed at understanding the physical, psychological, and practical challenges associated with IRDs. The survey collected responses from 537 Canadian patients (mean age 51 years; 54% female) living with a variety of IRDs. The majority of respondents were located in Ontario (48%), British Columbia (17%), and Alberta (14%). Many of them indicated living with 1 or more eye-related comorbidities, with cataracts being the most frequent (34%), followed by glaucoma (7%), macular degeneration (5%), and others. Approximately 71% of the survey group respondents reported having RP, 4% reported LCA, and the remaining reported a

broader spectrum of IRDs. Seven individuals had the *RPE65* mutation treated by voretigene neparvovec, which in each case was confirmed by genetic testing.

IRDs are a group of rare genetic conditions caused by any one of more than 250 possible genetic mutations. The diseases typically lead to degenerative vision loss (the progressive loss of sight over time) and are accompanied by a range of complex physical, psychological, and economic burdens. The most common IRD (reported in approximately 1 in 4,000 Canadians) is RP, a disorder characterized by damage to the photoreceptor cells of the retina that leads to a reduced ability to perceive light and a progressive loss of vision. RP is usually diagnosed during childhood or adolescence, although a small group of patients develop symptoms later in life. During the early stages of the disease, peripheral and night vision are affected first, followed by a progressive form of tunnel vision (narrowing of the VF); later stages are characterized by the loss of central vision and near or total blindness. A less common (affecting approximately 2 to 3 per 100,000 newborns) but more severe IRD is LCA, which is characterized by an earlier onset and more rapid progression of visual impairment. RP and LCA can be caused by mutations in more than 64 and 17 possible genes, respectively, with disease type and symptoms determined by the gene(s) affected.

Participants in the survey had varying degrees of vision loss: no night vision (41%), good central vision (34%), some useful central vision (32%), some useful peripheral vision (27%), some light perception (14%), good overall vision (10%), good peripheral vision (7%), no vision (5%), and shadows only (3%). Even though VA was maintained for some participants, 67% of the patients in the overall group were diagnosed as legally blind, over 70% had moderate to severe low vision, and 12% had near or total blindness. Notably, a higher percentage of participants with LCA reported a loss of visual perception, consistent with its early childhood onset and rapid visual deterioration.

The survey report showed the daily challenges and impacts on daily life associated with IRDs. Among the survey participants, 54% responded that IRD had negatively affected their employment or school status, with 22% rating the impact of their disease on their ability to perform job or school responsibilities as very severe. Patients reported struggling through school, requiring assistive technologies, or having to quit schooling as a result of the daily challenges being too difficult to overcome. Maintaining a stable career was flagged as being equally challenging. Patients faced discrimination from employers with respect to finding a job, keeping a job, having to settle for jobs below or not in accordance with their skill level, or being laid off without disability benefits or a pension. One patient summarized the disease experience as having an enormous and onerous impact:

> I struggled all through elementary school and junior high. It wasn't until grade 9 when I got my first CCTV and computer with a screen reader and magnifier that I finally could do my own homework without help. I got taken out of gym class instead of having adapted activities, and this affected by (sic) weight, which has become a lifelong struggle. At work, I have been discriminated against. I have failed interviews when I have used adaptive technology to complete them.

Patients reported challenges in almost all aspects of life. When asked how much their eyesight interferes with a range of activities, participants indicated some degree of interference with most daily activities, including "mobility and getting around," "hobbies/leisure," "socializing and interacting with others," "looking after your appearance," "reading a book or a newspaper," and "using a phone or iPad." Patients worry about their condition getting worse; struggle with challenges presented by daily activities, including

parenting; experience long wait times for appointments; feel anxiety and uncertainty about the future and the impact of their diseases on their families; and, in some cases, are impacted by a lack meaningful work, education, or a social life.

In addition to physical challenges, responses to the survey show that IRDs present a considerable psychological burden to patients that results, in many cases, from anxiety over worsening sight. This emotional and psychological strain is persistent and prolonged since the underlying condition is progressive in nature. The majority of the respondents (97%) think about their IRD and/or vision loss at least once a week and sometimes often. Patients expressed concerns with general safety both in and outside of home, progressive worsening of eyesight, and coping with everyday life. All of this contributed to a general sense of depression and a negative impact on mental health, with survey respondents reporting anxiety (71%), stress (73%), fear (64%), anger (62%), loss of confidence (58%), isolation (47%), employment barriers (43%), loneliness (41%), discrimination (37%), and lack of self-worth (34%). Discrimination and stigma were foregrounded as having a psychological impact as well, with over half of the patients reporting differential treatment by others in the workplace, educational institutions, or general public.

At the same time, the psychological and professional impacts of IRDs penetrate into broader social contexts since these are typically interconnected. Three-quarters of the survey participants reported slight or severe negative effects on their social life resulting from IRDs. Withdrawing from relationships and activities was shown to be common, particularly after dark and when transportation is required. Among the patients struggling to maintain social lives, the dangers of anxiety, isolation, loneliness, and depression are most acute. The negative impact of IRDs often extends to the patient's family, and 2-thirds of the surveyed patients reported negative effects on their families ranging from slightly negative to very negative. Concerns over genetic inheritance of the disease and a general apprehensiveness about the future were the most common among family members.

Experience With Treatment

No pharmacological or surgical treatments were described in the patient group submission. However, a wide variety of modifications or aids such as canes, magnifiers, and specialized laptops are available to patients for daily activities. The following modifications or aids were reportedly used most often: canes, magnifiers, specialized books with enlarged fonts or audiobooks, special applications for mobile phones, modified laptops, screen readers, ergonomic adaptations, Braille, and voice-run applications or machines (e.g., Google Home). While these devices and aids are supportive in nature, they are used to navigate some of the day-to-day challenges of work and school outlined previously.

A case report was provided in the submission, taking the form of an informal interview by FBC staff of a parent of a child treated with Luxturna. The child had early symptoms of visual impairment at 2 months of age, shown by nystagmus (rapid involuntary movements of the eye) and other signs, within a few months of being born. Based on a genetic test, a confirmed diagnosis of LCA with biallelic *RPE65* mutation was made at the age of 10 months. After missing the age cut-off for a clinical trial for Luxturna, the child was provided access to the treatment as a special case following a request made to the government of Quebec. Both testing and treatment procedures were reportedly streamlined, and associated costs were covered. The parent reported that Luxturna substantially benefited the child's vision. Prior to treatment, the child was light sensitive and had poor vision in dark or dim settings. Even during the day, their vision was far from normal, with gaps or blind spots reported that made it difficult to read, play, and identify objects or people. Many of

these complications were tied to an experience that could be described as “Swiss cheese vision,” according to the parent, where sight is obscured by a multitude of gaps or blind spots. The post-treatment results were almost immediately noticeable: the child was able to identify things and recognize people much better. This ultimately led to an overall and extensive improvement in the child’s confidence and self-reliance. The parent also felt relieved from some of the complexities of caregiving, spending a substantially lower amount of time assisting with the child’s daily activities, such as getting dressed and completing schoolwork. The child was reportedly much more self-sufficient, playing independently and being active, and relying on a magnifying lens for reading only on rare occasions. While it was acknowledged that the treatment did not lead to perfect vision and had no impact on VA, the improvements in low-light vision and to “Swiss cheese” gaps were substantial compared with pre-treatment life. The parent was cognizant of the uncertain longevity of the drug, but still considered the treatment to be life-changing due to the additional years of improved vision incurred by the treatment.

Improved Outcomes

Among the surveyed patients, most reported their emotional well-being would improve significantly with a new treatment, especially if the treatment were to recover overall sight, cure the condition entirely, improve night vision and mobility at night, or facilitate regular day-to-day activities such as social interactions, maintaining personal relationships, work, and study. Even if the treatments were to only enhance vision and mobility at night, most surveyed participants reported that this would improve their overall quality of life. Many respondents also indicated that a stabilization of vision from the new treatment would be valuable, something that would at least halt progression even if retinal damage could not be reversed.

Additional Details on Genetic Testing

Genetic testing is an important factor in the diagnosis of underlying gene mutations and, as a result, is an essential step in determining the appropriate treatment pathway for patients with IRDs. Among the survey respondents, over one-third did not receive a genetic test or meet with a genetic counsellor. Patients who underwent testing were more likely to be aware of their genetic mutation. Tests are also a helpful way to determine the family history of the genetic condition. A genetic history of the survey participants’ family members was either not done or was unknown in 60% of cases. Most cited difficulty in getting tested or receiving genetic counselling, as well as receiving inconclusive results, as the reasons for the low uptake of genetic testing. This underscores the need for a contemporary genetic test that is accurate and reliable, with minimum barriers to access. While genetic tests are covered by all public health plans across Canada, the well-known shortage of genetic counsellors and related infrastructure is a likely factor in the underutilization of genetic tests, illustrating an unmet need and area of improvement needed in ocular treatment.

Clinician Input

All CADTH review teams include at least 1 clinical specialist with expertise in the diagnosis and management of the condition for which the drug is indicated. Clinical experts are a critical part of the review team and are involved in all phases of the review process (e.g., providing guidance on the development of the review protocol, assisting in the critical appraisal of clinical evidence, interpreting the clinical relevance of the results, and providing guidance on the potential place in therapy). In addition, as part of the voretigene

neparvovec review, a panel of 6 clinical experts from across Canada was convened to characterize unmet therapeutic needs, assist in identifying and communicating situations where there are gaps in the evidence that could be addressed through the collection of additional data, promote the early identification of potential implementation challenges, gain further insight into the clinical management of patients living with a condition, and explore the potential place in therapy of the drug (e.g., potential reimbursement conditions). A summary of this panel discussion follows.

Diagnosis

Patients are diagnosed with IRDs clinically when they exhibit vision loss, loss of central or peripheral VFs, and/or nyctalopia or colour-vision defects in a clinical and hereditary manner consistent with inherited retinal disease. Usually, additional testing with visual electrophysiology is performed to confirm the phenotype from the parameters exhibited clinically. In the majority of patients with mutations in known genes, a diagnosis of inherited retinal disease is dependent on molecular genetic testing. This testing is usually guided by the findings of the clinical and electrophysiologic phenotyping and may consist of panel testing, where a predetermined set of genes is researched based on the clinical parameters, such as a macular dystrophy panel or cone-rod dystrophy panel. Alternatively, when the clinical signs are helpful, a single-gene analysis could be accessed faster. In Canada, molecular genetic testing is performed by specialists certified in medical genetics whose training does not encompass the ophthalmic examinations and investigations necessary for adequate phenotyping. Hence, the process of molecular genetic testing in adults with inherited retinal disease is separated structurally from the essential process of phenotyping required to direct and interpret genetic testing. This poses unique challenges to the implementation of care in adult patients. In Quebec, the current process to complete phenotyping with electrophysiology takes weeks, and patients with phenotype characterizations completed in 2017 and 2018 still await molecular genetic testing.

The involvement of biallelic *RPE65* mutations is confirmed through molecular genetic analysis that is performed in laboratories that are certified for the molecular diagnosis of genetic eye diseases. To report these mutations, these certified laboratories follow the American College of Medical Genetics guidelines, which are very reliable.[33]

Treatment Goals

In the absence of treatment for inherited retinal disease, the ultimate objective currently is to help each patient achieve their self-described life goals despite the limitation in their visual functioning. These goals could include learning Braille and adapting to a guide dog in order to maneuver in a university setting, or shifting from playing soccer to coaching as a milder form of a rod photoreceptor disease sets in.

The most important thing for patients is “not to go completely blind.” Most inherited retinal degeneration is relentless and progressive, and patients fear becoming totally blind with no light perception. The most important outcome for patients is to maintain or improve their mobility, independence, and ability to achieve their life goals. The ideal treatment would result in delaying or stopping disease progression and maintaining the vision the patients had before they initiated therapy. Hence, if the treatment slows down or stops the retinal degeneration and preserves vision (including day and night vision) with minimal adverse events, that would be an ideal treatment. Other treatment goals would be improving functional vision and visual function, maintaining or improving the patients’ ability to perform ADL, and maintaining or improving HRQoL. The improvement in rod function, cone

function, VF, and VA measurements must correlate with and result in enhancing the patients' ability to perform ADL.

Unmet Needs

Currently, there is no approved pharmacological drug for the treatment of patients with vision loss due to IRD caused by confirmed biallelic *RPE65* mutations. The current treatment is supportive care, where the treatment of patients who have progressive and inevitable vision loss revolves mainly around counselling the patient. If the patient is a child, the counselling would be delivered to the caregivers of the child and would include referring them to institutions that may be able to advise them on how to raise a child with low vision and whose vision is expected to deteriorate. Other supportive care for visually disabled children and young adults comprises mainly assistive devices, including magnifying devices for school or work, which may be provided through school boards or the CNIB Foundation. Patients are given mobility training and use sunglasses for photophobia and corrective glasses. However, these supportive-care measures have limited impact on maintaining the patients' ability to perform ADL or maintaining HRQoL or occupational functioning; they reduce neither the severity of the disease nor its progression.

The most vulnerable patients with *RPE65* mutations appear to be those aged 10 to 20 years, as patients in that age group lose considerable VF and start to lose their cone function and rod function. Therefore, it would be better to intervene before patients reach that age group, but age on its own should not be the only criteria for treating patients.

Place in Therapy

Given there is currently no pharmacologic treatment available, and the only management available for patients with vision loss due to IRD caused by confirmed biallelic *RPE65* mutations is supportive care, voretigene neparvovec would cause a shift in the current treatment paradigm.

Patient Population

Patients best suited for treatment with voretigene neparvovec are those who are at least 4 years of age with sufficient viable retinal cells. Patients who are younger than 4 years of age may benefit from treatment with voretigene neparvovec; however, the surgery and administration of treatment in that age group is difficult and the assessment of treatment response is influenced by the inability to measure and monitor visual function, as the outcome measurements are not designed for young children under 4 years of age.

The clinical experts indicated that measuring viable retinal cells is not a straightforward procedure. The panellists also indicated there is no benchmark or threshold of viable retinal cells that can be used to objectively define the sufficient viable retinal cells criterion. The clinical experts also indicated that a numerical cut-off could not be applied universally throughout all OCT technologies and generations because of the changes in OCT and differences between measurements on different OCT technologies. In clinical practice, the presence of sufficient viable retinal cells would be determined by the treating physician using OCT examinations and measuring the area of remaining viable photoreceptors, supplemented by tests of VA and visual function.

Biallelic *RPE65* mutations display clinical heterogeneity. This means that the exact same mutations in 1 individual may result in blindness and nystagmus at birth (which is labelled as LCA), whereas in another individual, the same mutations may display only VF restriction

and nyctalopia in adulthood with a much slower progression (historically labelled RP). The reason for this clinical heterogeneity is poorly understood. Hence, the need for RPE65 therapy should not be based on the age of onset of disease. The panellists also indicated that patients who might benefit from treatment should not be chosen based on whether they have LCA or RP, as patients diagnosed with either LCA or RP caused by *RPE65* mutations are both suited for treatment with voretigene neparvovec. While patients with LCA or RP differ in terms of their age of diagnosis, both have an unmet need and there is no difference in the types or severity of mutations between these 2 groups.

The clinical experts indicated that younger patients might respond to treatment better than adult patients, mainly because adult patients usually have more extensive structural damage; however, the panellists added that age on its own should not be a criterion for treating patients but rather the presence of sufficient viable retinal cells.

Given there is some risk of side effects associated with the administration of voretigene neparvovec, if a patient has better than 20/60 VA in both eyes, the panellists indicated that, while these patients might benefit from treatment with voretigene neparvovec, they would be more cautious and have a detailed discussion with the patient or their caregiver about the potential risks and benefits of the procedure. Similarly, if a patient has very poor vision and very poor remaining function, then physicians would have a detailed discussion with the patient or their caregiver about the potential risks and benefits of the procedure and whether the patient would benefit from the procedure. The panellists also indicated that disease progression also leads to structural retinal damage that can increase the risk of side effects from the vitrectomy procedure and the subretinal injection of the vector. Hence, patients with less damage (better vision) would be more suited from a procedure point of view. At the same time, procedure risks are greater when treating children and could have more serious consequences for vision health; hence, the risks versus benefits of the procedure should be carefully weighed.

Assessing Response to Treatment

The clinical experts indicated that, in practice, clinicians would use testing to assess VA and VF, plus or minus electrophysiology tests. The experts also indicated that FST testing is very informative for showing improvement in symptoms. However, the FST is not used in clinical treatment settings. The bilateral MLMT also has not been widely used in clinical practice and is currently not available in any Canadian centre. The patient's HRQoL and ability to perform ADL should also be assessed in clinical practice. The administration of these questionnaires and subsequent assessments by a multi-disciplinary team may be necessary to distinguish minor improvements in patient functioning, whether personal or occupational.

After receiving treatment, a patient would be seen for assessment on day 1 and again after 1 week, 1 month, and 3 months, and then annually thereafter. A clinically meaningful response to treatment would show an improvement of 3 lines in VA or an improvement in VF. In addition, improvements in the patient's ability to perform ADL under low-light conditions would be considered a clinically meaningful improvement.

The clinical experts indicated it is expected that treatment response will wane over time; however, it is uncertain how long the treatment effect will last and when the treatment effect will start waning. Some clinical experts indicated they might re-treat patients with voretigene neparvovec after treatment effect waned; however, the panellists also indicated that for the

time being, there is no evidence indicating how safe it is to repeat a vitrectomy, and there are many unknowns, risks, and uncertainties associated with re-treatment.

Prescribing and Administration

Most of the panellists indicated that voretigene neparvovec should be initiated and administered by a retinal surgeon experienced in performing sub-macular injection and managing its complications. The genetic diagnosis and the pre- and post-surgical evaluations (safety and efficacy) should be confirmed by a specialist in inherited retinal diseases. Patient selection must be done by an inherited retinal disease specialist with specific fellowship training in the area of inherited retinal diseases, genetics, and clinical end points. It is expected that there would be a very close collaboration between a retinal surgeon or a retinal surgeon with experience in children, and an ocular geneticist who is familiar with IRD. One panellist indicated that, in adult patients, candidacy determination and clinical monitoring for voretigene neparvovec therapy should be performed by an ocular geneticist skilled in retinal disease, and the surgery for implementation of the therapy should be performed by a vitreoretinal surgeon. In pediatric patients, candidacy determination and clinical monitoring for voretigene neparvovec therapy should be performed by a pediatric ophthalmologist skilled in pediatric ocular hereditary disease, and the surgery for implementation of the therapy should be performed by a vitreoretinal surgeon. Given the unique risks that the implementation of this therapy poses in this population, vitreoretinal surgeons should have formal fellowship training or considerable clinical experience. All panellists indicated that for adult and pediatric patients, a medical geneticist is required to complete the genetic testing.

The clinical experts indicated that some patients might experience adverse events such as weight gain, sleeplessness, anxiety, or acne due to the immunomodulatory regimen that patients must receive before and after the administration of voretigene neparvovec. They also indicated that while these adverse events are temporary, patients must be made aware of them.

Clinical Evidence

The clinical evidence included in the review of voretigene neparvovec is presented in 3 sections. The first section, the systematic review, includes pivotal studies provided in the sponsor's submission to CADTH and Health Canada, as well as those studies that were selected according to an a priori protocol. The second section is intended to include indirect evidence; however, no indirect evidence was submitted by the sponsor, nor was any indirect evidence that met the selection criteria specified in the review identified from the literature. The third section includes sponsor-submitted long-term extension studies and additional relevant studies that were considered to address important gaps in the evidence included in the systematic review.

Systematic Review (Pivotal and Protocol-Selected Studies)

Objectives

To perform a systematic review of the beneficial and harmful effects of voretigene neparvovec administered as a single dose of 1.5×10^{11} vg in each eye via subretinal injection for the treatment of adult and pediatric patients with vision loss due to IRD caused by confirmed biallelic *RPE65* mutations and who have sufficient viable retinal cells.

Methods

Studies selected for inclusion in the systematic review included pivotal studies provided in the sponsor's submission to CADTH and Health Canada, as well as those meeting the selection criteria presented in Table 3.

Of note, the systematic review protocol presented subsequently was established prior to the granting of a Notice of Compliance by Health Canada.

Table 3: Inclusion Criteria for the Systematic Review

Patient population	Adult and pediatric patients with vision loss due to inherited retinal dystrophy caused by confirmed biallelic *RPE65* mutations and who have sufficient viable retinal cells. Subgroups: • age • clinical phenotype • viable retinal cells.
Intervention	A single dose of 1.5×10^{11} vg of voretigene neparvovec in each eye administered via subretinal injection.
Comparators	Standard of care.
Outcomes	**Efficacy outcomes** • Functional vision (e.g., ability to perform activities of daily living, MLMT)[a] • Visual function (e.g., visual acuity, visual field, contrast sensitivity, and light sensitivity)[a] • HRQoL for patients • Caregiver burden (e.g., HRQoL)[a] • Disease progression[a]

	Harms outcomes • AEs, SAEs, WDAEs, mortality • Notable harms and harms of special interest related to treatment: cataract development or progression, endophthalmitis, increased intraocular pressure, retinal abnormalities, and ocular inflammation • Notable harms (procedure-related events): vision loss, retinal detachment, retinal tear, and macular changes
Study design	Published and unpublished phase III and IV RCTs.

AE = adverse event; HRQoL = health-related quality of life; MLMT = multi-luminance mobility test; RCT = randomized controlled trial; SAE = serious adverse event; vg = vector genome; WDAE = withdrawal due to adverse event.

[a] These outcomes were identified as being of particular importance to patients in the input received by CADTH from patient groups.

The literature search for clinical studies was performed by an information specialist using a peer-reviewed search strategy according to the *PRESS Peer Review of Electronic Search Strategies* checklist (www.cadth.ca/resources/finding-evidence/press).[34]

Published literature was identified by searching the following bibliographic databases: MEDLINE All (1946–) through Ovid, Embase (1974–) through Ovid, and PubMed. The search strategy comprised both controlled vocabulary, such as the National Library of Medicine's MeSH (Medical Subject Headings), and keywords. The main search concept was Luxturna (voretigene neparvovec). Clinical trial registry searched: The US National Institutes of Health's clinicaltrials.gov.

No filters were applied to limit the retrieval by study type. Retrieval was not limited by publication date or by language. Conference abstracts were excluded from the search results. See Appendix 1 for the detailed search strategies.

The initial search was completed on May 22, 2020. Regular alerts updated the search until the meeting of the CADTH Canadian Drug Expert Committee (CDEC) on September 16, 2020.

Grey literature (literature that is not commercially published) was identified by searching relevant websites from the following sections of the *Grey Matters: A Practical Tool For Searching Health-Related Grey Literature* checklist (www.cadth.ca/grey-matters):[35] Health Technology Assessment (HTA) Agencies, Health Economics, Clinical Practice Guidelines, Drug and Device Regulatory Approvals, Advisories and Warnings, Drug Class Reviews, Clinical Trials Registries, and Databases (Free). Google was used to search for additional internet-based materials. In addition, the sponsor of the drug was contacted for information regarding unpublished studies. See Appendix 1 for more information on the grey literature search strategy.

Two CADTH clinical reviewers independently selected studies for inclusion in the review based on titles and abstracts, according to the predetermined protocol. Full-text articles of all citations considered potentially relevant by at least 1 reviewer were acquired. Reviewers independently made the final selection of studies to be included in the review, and differences were resolved through discussion.

Findings From the Literature

One study was identified from the literature for inclusion in the systematic review (Figure 1). The included study is summarized in Table 4. A list of excluded studies is presented in Appendix 2.

Figure 1: Flow Diagram for Inclusion and Exclusion of Studies

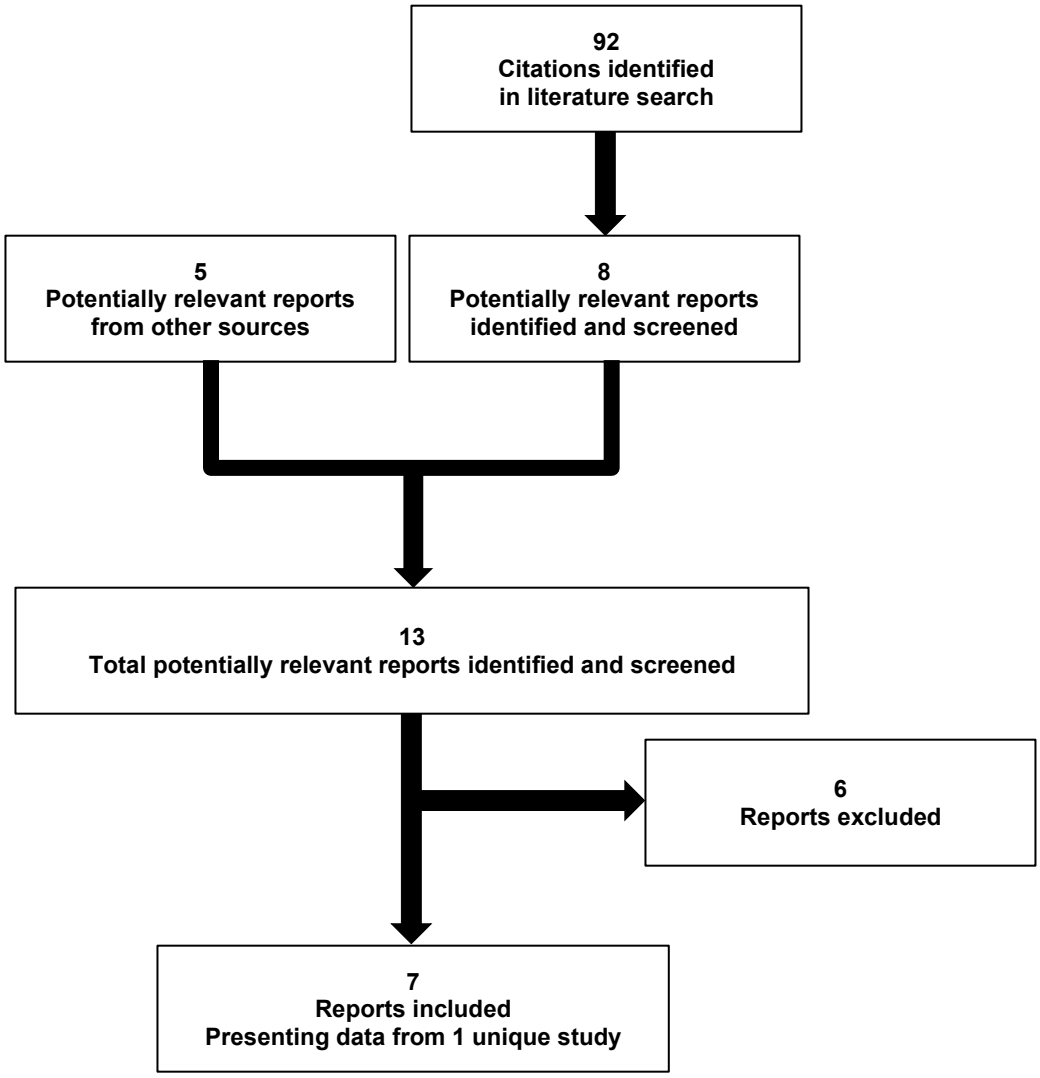


Table 4: Details of Included Studies

		Study 301
Designs & Populations	**Study design**	Phase III, open-label RCT.
	Location	US
	Randomized (N)	31

	Inclusion criteria	• Diagnosis of LCA due to *RPE65* mutations; molecular diagnosis was performed, or confirmed, by a CLIA-certified laboratory.
		• 3 years old or older.
		• Visual acuity worse than 20/60 (both eyes) and/or visual field less than 20° in any meridian as measured by III4e isopter or equivalent (both eyes).
		• Sufficient viable retinal cells as determined by non-invasive means, such as OCT and/or ophthalmoscopy, and by at least 1 of the following:
		  ○ an area of retina within the posterior pole of > 100 µm thickness shown on OCT
		  ○ ≥ 3 disc areas of retina without atrophy or pigmentary degeneration within the posterior pole
		  ○ a remaining visual field within 30° of fixation as measured by III4e isopter or equivalent.
		• Patients must be evaluable on mobility testing (the primary efficacy end point) to be eligible for the study. Evaluable was defined as:
		  ○ The ability to perform mobility testing within the luminance range evaluated in the study. Individuals must receive an accuracy score of ≤ 1 during screening mobility testing at 400 lux or less to be eligible; individuals with an accuracy score of > 1 on all screening mobility test runs at 400 lux, or those who refuse to perform mobility testing at screening, were excluded.
		  ○ The inability to pass mobility testing at 1 lux. Individuals must fail screening mobility testing at 1 lux to be eligible; individuals who pass 1 or more screening mobility test runs at 1 lux were excluded.

	Exclusion criteria	• Unable or unwilling to meet study requirements, including receiving bilateral subretinal vector administrations. • Use of retinoid compounds or precursors that could potentially interact with the biochemical activity of the RPE65 enzyme; individuals who discontinue use of these compounds for 18 months may become eligible. • Prior intraocular surgery within the past 6 months. • Known sensitivity to medications planned for use in the perioperative period. • Pre-existing eye conditions or complicating systemic diseases that would preclude the planned surgery or interfere with the interpretation of the study. Complicating systemic diseases would include those in which the disease itself, or the treatment for the disease, can alter ocular function. Examples are malignancies whose treatment could affect central nervous system function. Patients with diabetes or sickle cell disease would be excluded if they had any manifestation of advanced retinopathy. Also excluded were patients with immunodeficiency (acquired or congenital). • Individuals of childbearing potential who are pregnant or unwilling to use effective contraception for 4 months following vector administration. • Individuals incapable of performing mobility testing (the primary efficacy end point) for a reason other than poor vision, including physical or attentional limitations. • Any prior participation in a study in which a gene therapy vector was administered. • Participation in a clinical study with an investigational drug in the past 6 months. • Any other condition that would not allow the potential patient to complete follow-up examinations during the course of the study or, in the opinion of the investigator, makes the potential patient unsuitable for the study.
Drugs	**Intervention**	Patients received a dose of 1.5×10^{11} vg of voretigene neparvovec in a total subretinal volume of 300 µL in each eye. The individual administration procedures were to be performed on separate days no more than 18 days apart (12 days ± 6 days).

		Study 301
	Comparator(s)	No intervention. Patients in the control group became eligible to receive voretigene neparvovec at a dose of 1.5×10^{11} vg in a total subretinal volume of 300 µL in each eye (within 18 days), 1 year after their baseline evaluations, provided they still met all eligibility criteria.
DURATION	**Phase**	
	Open-label	1 year
	Follow-up	10 years
OUTCOMES	**Primary end point**	Change from baseline in bilateral MLMT performance at 1 year.
	Secondary and exploratory end points	Secondary: • FST testing: Average light sensitivity (averaged over both eyes) for white light at 1 year as compared with baseline light sensitivity testing • monocular mobility testing change score: Change from baseline to 1 year in the MLMT score for the first eye • visual acuity: Average change in visual acuity (averaged over both eyes) at 1 year compared with baseline. Exploratory: • visual field testing by Humphrey computerized testing for static fields • visual field testing by Goldmann perimetry for kinetic fields • contrast sensitivity • Visual Function Questionnaire.
NOTES	**Publications**	Russell et al.[4]

CLIA = Clinical Laboratory Improvement Amendments of 1988; FST = full-field sensitivity threshold; LCA = Leber congenital amaurosis; MLMT = multi-luminance mobility testing; OCT = optical coherence tomography; RCT = randomized controlled trial; RPE65 = retinal pigment epithelium 65 kDa protein.

Note: Five additional reports were included (sponsor's submission,[12] FDA clinical and statistical reviews,[36,37] and European Medicines Agency assessment report[38]).

Source: Russell et al.,[4] Maguire et al.,[39] Clinical Study Report for Study 301.[15]

Description of Studies

One trial (Study 301) met the inclusion criteria of the CADTH review protocol. Study 301 (N = 31) was a phase III, open-label, randomized controlled trial designed to evaluate the efficacy and safety of sequential subretinal injection of voretigene neparvovec into each eye in pediatric and adult patients diagnosed with LCA due to *RPE65* mutations. Randomization, which was to follow screening and confirmation of study eligibility, was to occur in a 2:1 ratio of intervention (voretigene neparvovec) to control and use a block design stratified by age (≥ 10 years versus < 10 years) and mobility testing passing level (pass at ≥ 125 lux versus < 125 lux) as determined at screening. A total of 31 patients in 2 study sites in the US were randomized to either the treatment group (n = 21) or the control group (n = 10). The study enrolled international patients, including 1 patient from Canada.

Patients randomized to the voretigene neparvovec group received a dose of 1.5×10^{11} vg of voretigene neparvovec in each eye; the non-simultaneous, subretinal injections were to occur within an 18-day period (12 days ± 6 days). Patients randomized to the control group did not receive voretigene neparvovec, sham injection, or corticosteroids for a period of at least 1 year from baseline evaluations. Following repeated retinal and visual function analysis, including mobility testing, at 1 month, 3 months, 6 months, and 1 year, patients in the control group were crossed over to receive non-simultaneous injections of 1.5×10^{11} vg of voretigene neparvovec in each eye (within 18 days) after 1 year of randomization, provided they still met all study eligibility criteria.

The most recent CSR addendum from the sponsor provided updated results of available efficacy and safety data through July 2, 2018. This CSR was finalized [redacted]. At the time of the data cut-off for this CSR, all [redacted] in the voretigene neparvovec treatment group had [redacted] after the second-eye injection study visit and [redacted] in the control/voretigene neparvovec treatment group had completed the [redacted] study visit. In addition, [redacted] in the voretigene neparvovec treatment group had [redacted] after the second-eye injection study visit and [redacted] in the control/voretigene neparvovec treatment group had completed the [redacted] after the second-eye injection study visit.

Populations

Inclusion and Exclusion Criteria

Patients enrolled in Study 301 had to be 3 years of age or older with best-corrected VA worse than 20/60 in both eyes (i.e., in each eye) and/or a VF of less than 20° in both eyes (i.e., in each eye). Molecular diagnosis or confirmation of diagnosis of *RPE65* mutations by a Clinical Laboratory Improvement Amendments of 1988 (CLIA)–certified laboratory was required; homozygotes or compound heterozygotes were both eligible. Rather than establishing an upper age limit, eligibility was based on whether potential participants retained sufficient viable retinal cells, defined as at least 1 of the following:

- an area of retina within the posterior pole of greater than 100 µm thickness as shown on OCT
- 3 or more disc areas of retina without atrophy or pigmentary degeneration within the posterior pole based on ophthalmoscopy
- a remaining VF within 30° of fixation.

Additionally, patients participating in the study had to be evaluable by the primary efficacy end point, mobility testing. Patients who were able to pass the mobility course at screening, in the time allotted at the lowest illumination to be evaluated (1 lux), were considered too close to normal function with respect to ability to navigate in dim light conditions; these patients were not eligible to enrol in the study. Patients who were unable to perform the mobility course at screening with an accuracy score of 1 or less at the highest illumination to be evaluated (400 lux) were considered to have extensive disease progression such that they would be less likely to achieve measurable, clinically meaningful benefit; these patients were not eligible to enrol in the study. (An accuracy penalty was assigned for every collision with obstacles, stepping off the course, or having to be redirected by the tester; the accuracy score was then obtained by dividing the number of accuracy penalties by the total number of obstacles. The patient could receive more points than there were obstacles, and a perfect score was 0.) Patients were also excluded from Study 301 if they had prior intraocular surgery within the past 6 months, or if they had any prior participation in a study in which a gene therapy vector was administered.

Baseline Characteristics

In Study 301, there was a larger percentage of females (n = 18; 58%) than males (n = 13; 42%) enrolled, and patients were primarily White (68%). At the time of randomization, the patient ages ranged from 4 to 44 years (mean = 15.1; SD = 10.9 years). Patients randomized to the voretigene neparvovec group were younger (median = 11; range = 4 to 44 years) than those randomized to the control group (median = 14; range = 4 to 31 years). A larger percentage of patients with baseline MLMT performance (< 125 lux) were enrolled

in the voretigene neparvovec group (57%) than in the control group (40%). On average, patients in the control group had better full-field light sensitivity than patients randomized to the voretigene neparvovec group. Patients randomized to the voretigene neparvovec group had, on average, better acuity (less VA loss), and better VF (as measured by Goldmann VF V4e, Humphrey VF [foveal sensitivity], and Humphrey VF [macula threshold]) than patients randomized to the control group. Study 301 enrolled 1 patient from Canada. A summary of baseline characteristics is presented in Table 5.

Table 5: Summary of Baseline Characteristics in Study 301 (ITT)

Baseline characteristics	Voretigene neparvovec n = 21	Control (n = 10)
Age (years)		
Mean (SD)	14.7 (11.8)	15.9 (9.5)
Median (range)	11 (4 to 44)	14 (4 to 31)
Age (at screening), n (%)		
Age < 10 years	9 (43)	4 (40)
Age ≥ 10 years	12 (57)	6 (60)
Sex, n (%)		
Male	9 (43)	4 (40)
Female	12 (57)	6 (60)
Race, n (%)		
White	14 (67)	7 (70)
Asian	3 (14)	2 (20)
American Indian or Alaska Native	2 (10)	1 (10)
Black or African American	2 (10)	0
Country of residence, n (%)		
US	17 (81)	6 (60)
Netherlands	1 (5)	2 (20)
Belgium	0	1 (10)
Canada	1 (5)	0
India	1 (5)	0
Italy	0	1 (10)
Mexico	1 (5)	0
Past ocular history, n (%)		
Patients reporting any abnormal ocular history	21 (100)	10 (100)
Cataracts	1 (5)	1 (10)
Lenticular opacities	1 (5)	2 (20)
Nystagmus	21 (100)	10 (100)
Retina abnormalities	21 (100)	10 (100)
Strabismus	8 (38)	5 (50)
Other	3 (14)	2 (20)
Baseline ophthalmic examination, n (%)[a]		
Anterior segment inflammation (grade < 1)	21 (100)	10 (100)
Cornea (grade < 1)	21 (100)	10 (100)
Intraocular pressure (grade < 1)	21 (100)	9 (100)
Optic nerve (grade < 1)	21 (100)	10 (100)
Posterior segment inflammation (grade < 1)	21 (100)	10 (100)
Retinal inflammation (grade < 1)	21 (100)	10 (100)
Nystagmus present	20 (95)	10 (100)

Baseline characteristics	Voretigene neparvovec n = 21	Control (n = 10)
Distribution of lowest light levels passed bilaterally at baseline lux level on the MLMT, n (%)		
1 lux	0	0
4 lux	4 (19)	1 (10)
10 lux	5 (24)	2 (20)
50 lux	7 (33)	5 (50)
125 lux	3 (14)	1 (10)
250 lux	0	0
400 lux	0	0
> 400 lux[b]	2 (10)	1 (10)
MLMT (at screening),[c] n (%)		
Pass at < 125 lux	12 (57)	4 (40)
Pass at ≥ 125 lux	9 (43)	6 (60)
Full-field light sensitivity testing: white light (log_{10}[cd.s/m^2]) at baseline		
Baseline, mean (SE)	−1.29 (0.09)	−1.65 (0.14)
Visual acuity (LogMAR) using Holladay scale for off-chart results at baseline		
Baseline, mean (SE)	1.18 (0.14)	1.29 (0.21)
Goldmann VF V4e (sum total degrees) at baseline		
Baseline, mean (SD)	888.7 (487.8)	788.2 (482.9)
Goldmann VF III4e (sum total degrees) at baseline		
Baseline, mean (SD)	332.9 (413.3)	427.1 (372.0)
Humphrey VF, foveal sensitivity (dB) at baseline		
Baseline, mean (SD)	22.4 (6.8)	17.6 (8.9)
Humphrey VF, macula threshold (dB) at baseline		
Baseline, mean (SD)	16.1 (5.5)	14.4 (8.0)

cd.s/m^2 = candela second per square metre; dB = decibel; ITT = intention to treat; LogMAR = logarithm of the minimum angle of resolution; MLMT = multi-luminance mobility testing; SD = standard deviation; SE = standard error; VF = visual field.

[a] Denominators are the number of patients assessed for each exam for the specified eye.

[b] Patients in the > 400 lux row did not pass at 400 lux, the highest light level tested in the study.

[c] As determined by the eye with the worst passing result.

Source: Clinical Study Report for Study 301.[15]

Interventions

A dose of 1.5 × 10^{11} vg voretigene neparvovec in a total volume of 300 µL was injected into the subretinal space of each eye. The non-simultaneous, subretinal injections occurred within an 18-day period (12 days ± 6 days).

Patients randomized to the control group did not receive voretigene neparvovec, sham injection, or corticosteroids for a period of at least 1 year from baseline evaluations. Following repeated retinal and visual function analysis, including mobility testing, at 1 month, 3 months, 6 months, and 1 year, patients in the control group were crossed over to receive non-simultaneous injections of 1.5 × 10^{11} vg of voretigene neparvovec into each eye (within 18 days) after 1 year of randomization, provided they still met all study eligibility criteria.

Voretigene neparvovec was administered using a commercially available cannula designed for subretinal injection. Surgery was performed under general anesthesia supplemented by retrobulbar anesthetic to minimize intra-operative eye movement and post-operative

discomfort. Voretigene neparvovec injections were administered in the ophthalmology surgical suite at the Children's Hospital of Philadelphia or the University of Iowa. The surgical procedure used to deliver voretigene neparvovec was optimized and standardized over the course of the phase I studies to minimize risks related to vector administration. Four surgeons, in addition to the surgeon who performed the injections for the phase I studies, were successfully trained and included in Study 301. A qualified pharmacist with specific training in this protocol was responsible for receipt of the gene therapy material from the sponsor and its storage and preparation on the day of surgery.

Patients received systemic corticosteroids beginning 3 days before the first administration of voretigene neparvovec. The initial dose was 1 mg/kg per day of prednisone for 7 days, with a maximum prescribed dose of 40 mg per day, regardless of the weight of the patient; this was followed by 0.5 mg/kg per day of prednisone for an additional 5 days, with a maximum prescribed dose of 20 mg per day, regardless of the weight of the patient. The prednisone was then tapered further to 0.5 mg/kg every other day (maximum 20 mg every other day, regardless of weight) until 3 days prior to the second administration of voretigene neparvovec. The prednisone regimen surrounding the second injection was the same as that for the first injection, namely 1 mg/kg per day of prednisone for 7 days (maximum 40 mg per day, regardless of weight) followed by 0.5 mg/kg per day of prednisone for an additional 5 days (maximum 20 mg per day, regardless of weight). This regimen began 3 days prior to the second injection and continued through 8 days after the second injection. Introduction of the prednisone regimen surrounding the second injection superseded the taper of the regimen surrounding the first injection. Patients were on systemic corticosteroids for a minimum of 18 days up to a maximum of 30 days, depending on the timing of the second injection. The corticosteroids were given to minimize inflammation associated with the surgical procedure and to reduce the potential for an immune response to the voretigene neparvovec capsid and transgene product.

Once they were enrolled in the study, patients were not permitted to take or use any of the following medications until 1 year after the second-eye injection:

- investigational drugs, other than voretigene neparvovec
- high dose of vitamin A (more than 7,500 retinol-equivalent units or > 3,300 IU per day)
- tretinoin-containing skin cream
- isotretinoin
- sildenafil or related compounds used to treat erectile dysfunction
- hydroxychloroquine, chloroquine, thioridazine, or any related retinotoxic compounds.

Of note, periocular injection of the various corticosteroids used in the phase I trial were discontinued in the phase III trial to decrease the incidence and severity of elevated intraocular pressure and cataract formation or progression.

Outcomes

A list of efficacy end points identified in the CADTH review protocol that were assessed in the clinical trials included in this review is provided in Table 6. These end points are further summarized subsequently. A detailed discussion and critical appraisal of the outcome measures is provided in Appendix 4.

Table 6: Summary of Outcomes of Interest Identified in the CADTH Review Protocol

Outcome measure	Study 301
Change in bilateral MLMT performance at year 1 relative to baseline	Primary
Change in white light FST averaged over both eyes at year 1 relative to baseline	Secondary[a]
Change in assigned first-eye MLMT performance at year 1 relative to baseline	Secondary[a]
Change in visual acuity averaged over both eyes at year 1 relative to baseline	Secondary[a]
Visual field testing by Humphrey computerized testing for static fields	Exploratory
Visual field testing by Goldmann perimetry for kinetic fields	Exploratory
Contrast sensitivity	Exploratory
Visual Function Questionnaire	Exploratory

FST = full-field sensitivity threshold; MLMT = multi-luminance mobility testing.

[a] The primary efficacy end point was considered statistically significant if the permutation test P value was less than 0.05. The secondary efficacy end points were formally tested only if the primary efficacy end point was statistically significant. To control the overall type I error rate, the 3 secondary end points were tested hierarchically in the following order: change in the FST test result, MLMT score change using the first-treated eye, and change in visual acuity. Each of the 3 end points was tested at a 2-sided type I error rate of 0.05.

Source: Clinical Study Report for Study 301.[15]

Multi-Luminance Mobility Testing

The primary efficacy end point, designed to measure the effect of the intervention on functional vision, was the change in bilateral MLMT performance (change in lux score for the lowest passing light level) at 1 year relative to baseline.

The MLMT quantifies a patient's ability to navigate an obstacle course under varying environmental illuminations, including very low light levels. A person with 20/20 vision (20/20 vision is a term used to express the clarity or sharpness of vision measured at a distance of 20 feet (6 m); if a person has 20/20 vision, that person can see clearly at 20 feet what should normally be seen at that distance; if a person has 20/100 vision, it means that person must be as close as 20 feet to see what a person with 20/20 vision can see at 100 feet [30 m][40]) would be able to complete the course at 1 lux with no or minimal errors, which corresponds to the level of light available on a moonless summer night, or from an indoor night light. At the other end of the scale would be a light level of 400 lux, which equates to the level of light available within a well-illuminated indoor setting. Passing the MLMT, at any light level, is defined as completing the course at the specified light level with fewer than 4 errors (corresponding to an accuracy score of less than 0.25, where lower scores mean better accuracy) and within 3 minutes. The test had 12 course configurations to reduce learning effect. After 40 minutes of dark adaptation, patients completed the course with 1 eye patched, then completed a new randomly assigned configuration with the other eye patched and, finally, completed another randomly assigned configuration using both eyes. This was repeated at various light levels progressing from lower to higher luminance. Audiovisual recordings of MLMT were independently graded by 2 masked, trained reviewers and an adjudicator, if needed, at a separate time and location.

Patients were evaluated for accuracy and speed on the MLMT at 7 standardized light levels (1 lux, 4 lux, 10 lux, 50 lux, 125 lux, 250 lux, and 400 lux). Each light level was assigned a discrete lux score from −1 to 6, with lower light levels corresponding to higher lux scores (i.e., a score of −1 was assigned for more than 400 lux and a score of 6 was assigned for 1 lux). Baseline testing was used to establish the lowest level of illumination at which each patient could pass the MLMT, with the change score defined as the difference in lux scores

relative to baseline. A positive change score indicates passing the MLMT at a lower light level.

The MLMT was developed and validated by a group of researchers in collaboration with the sponsor of voretigene neparvovec; it was first implemented and standardized in 2 phase I trials (studies 101 and 102) and subsequently used in the pivotal phase III trial (Study 301).[41] Chung et al.[41] assessed the validity and reliability of the MLMT in collaboration with the sponsor in a prospective, observational study in 54 participants with 20/20 vision and visually impaired individuals with IRDs (of which 20 were diagnosed with LCA and 4 with RP). In this study, construct validity was demonstrated, as was high inter-observer, test-retest, and intra-observer reproducibility; responsiveness was demonstrated with questionable methodology. Given that the MLMT does not assess the ability to see at light levels lower than 1 lux, the authors noted there is a potential ceiling effect.[41] The sponsor indicated that an average change of 1 light level in passing the MLMT was considered clinically significant, although no supporting evidence was presented.[12] However, the FDA indicated that an MLMT score change of 1 may represent a background fluctuation occurring in both the treatment and control groups, and considered a clinically meaningful MLMT score change to be 2 or greater.[37] Similarly, the EMA indicated that any clinically relevant change in MLMT with voretigene neparvovec would need to exceed 1 light level; however; the EMA did not indicate what change would be considered clinically relevant.[38]

FST Testing

The FST is a measure of light sensitivity of the entire VF and is aimed at detecting the lowest luminance of a flash detected by a patient. In this test, flashes of various luminance (range spanning approximately 80 dB) are presented to individuals who are then asked to press a response button if they are able to see the visual stimulus.[42] FST results are expressed in dB, which are then converted to $\log_{10}(cd.s/m^2)$ to accommodate different dB conversion rates. Smaller values of dB and $cd.s/m^2$ indicate better sensitivity, and negative $\log_{10}(cd.s/m^2)$ values indicate better sensitivity.[15] No evidence was found in the literature for validity and responsiveness of the FST in patients diagnosed with IRD or other conditions. Acceptable inter-visit variability was found among patients with inherited retinal degeneration and individuals with 20/20 vision. [43] Acceptable coefficients of repeatability were reported among individuals considered legally blind (RP, macular disease, optic nerve disease, diabetic retinopathy, and other retinal diseases) and individuals with 20/20 vision.[44] The CSR for Study 301, referencing a study by Bittner et al.,[14] suggested a clinical significance threshold of 10 dB or 1 log change for the FST test; however, the methodology used to estimate the MID is unclear.[14]

Visual Acuity

VA is a measure of the ability of the eye to distinguish shapes and the details of objects, also knowns as optotypes (individual letters of a standardized size and contrast presented to the test taker for VA assessment), from a set viewing distance.[15] Two commonly used tools for testing VA include the Snellen eye chart and the Early Treatment Diabetic Retinopathy Study (ETDRS) chart. In Study 301, the VA was measured using the ETDRS eye chart. ETDRS charts present a series of 5 letters of equal difficulty on each row, with standardized spacing between letters and rows, for a total of 14 lines (70 letters). ETDRS letter score can be calculated when 20 or more letters are read correctly at a distance of 4.0 m; the VA letter score is equal to the total number of letters read correctly at 4.0 m plus 30. If fewer than 20 letters are read correctly at 4.0 m, the VA letter score is equal to the total number of letters read correctly at 4.0 m (number recorded on line 1), plus the total number

of letters read correctly at 1.0 m in the first 6 lines. Therefore, the ETDRS letter score could result in a maximum score of 100.[45,46] The standard chart-testing distance is 4.0 m; however, shorter distances may be used when vision is severely impaired.[47,48] The scoring for EDTRS charts is designed to produce a LogMAR suitable for statistical analysis in which individual letters score 0.02 log units.[49] The CSR of Study 301 indicated that a 0.1 improvement in LogMAR score corresponded to a 5-letter improvement (or the equivalent of 1 line) on an ETDRS eye chart. Notably, a decrease in LogMAR score represented an improvement in VA.[15]

Among patients in Study 301 who were unsuccessful in correctly identifying the line of largest letters on the ETDRS chart, off-chart VA measurements were collected by counting fingers and evaluating hand-motion perception, light perception, and no light perception. For off-chart VA measurements, LogMAR values were assigned using the scale adapted from Holladay et al.[50] and Lange et al.[51] The CSR for Study 301 indicated that the Lange method was implemented in response to a concern by the EMA and the study's data and safety monitoring board that a difference of 1 log-unit step between counting fingers and hand-motion perception in the Holladay method could present a biased estimate of a treatment effect (improvement or reduction in LogMAR). The Lange scale was therefore recommended as a sensitivity analysis for VA assessment, in which the 1 log-unit step between counting fingers and hand motion is reduced to a 0.3 log-unit step.[15,51]

Studies investigating the validity and responsiveness of the ETDRS were not found in the literature. A number of studies assessing the reliability of the ETDRS reported it as mixed or uncertain, with test-retest reliability and coefficients of reproducibility reported using unclear methodology. Assessments were done in patients with various forms of vision loss (e.g., RP, macular disease, optic nerve disease, diabetic retinopathy, and other retinal diseases). An IRD-specific MID for ETDRS was not identified in the literature. A study based on patients with macular edema estimated that a meaningful change in ETDRS is typically defined as greater than 3 lines (15 letters, equivalent to 0.3 LogMAR).[52] An improvement of 0.3 LogMAR (i.e., ≤ LogMAR 0.3) was considered clinically meaningful by the FDA in its medical review of voretigene neparvovec.[37]

Visual Field

A VF test, also known as a perimetry test, is a measure of an individual's entire scope of vision, including central and peripheral vision. With this test, the VF of each eye is mapped individually and blind spots (scotomas), as well as more subtle areas of dim vision, can be detected. In Study 301, VF was assessed using both Goldmann perimetry for kinetic fields and Humphrey testing for static fields, allowing for evaluation of different regions of the retina. For Goldmann VF testing, the stimulus sizes vary, expressed using a Roman numeral (indicating the Goldmann size of the stimulus), an Arabic number, and a letter (indicating the attenuation of the light). In Study 301, both the size III4e and size V4e test stimuli were used; the III4e target being 1/16th smaller in total area and one-quarter of the diameter of the V4e target. The test stimulus at baseline was used to determine which test stimulus would be used for follow-up visits; if the size III4e test stimulus resulted in individuals being able to see and reliably perform the VF testing, this size was used for each subsequent visit and vice versa. In contrast, the Humphrey testing for static fields (macula sensitivity and foveal sensitivity thresholds) always used a size V isopter. Humphrey VF macula threshold testing targets both cone and rod photoreceptor cells, while Humphrey foveal sensitivity testing targets the most central, cone-enriched region of the macula. Results for the Goldmann perimetry were presented as the sum of total degrees; higher sum totals indicate a greater area of functional, light-sensitive retina corresponding

to a greater field of vision for the patient. Results for the Humphrey VF testing was reported in dB of attenuation, or dimming, ranging from 0 dB (the brightest, unattenuated stimulus) to 51 dB (the dimmest, maximally attenuated stimulus); higher numbers (dB) corresponded to higher retinal sensitivity.[15]

A literature search conducted by CADTH found limited evidence of the validity of the Goldmann perimetry and Humphrey testing methods. Overall, the reliability of both Goldmann and Humphrey VF perimetry can be considered mixed and inconclusive. Notably, the National Institute for Health and Care Excellence (NICE) considered measures of VF as unreliable because of inter-test variability.[53] CADTH's literature search found no studies reporting an MID for either Goldmann or Humphrey perimetry. The FDA reported a change of 7 dB for Humphrey macula threshold to be clinically meaningful, based on expert opinion.[37] The sponsor reported an improvement of approximately 4 dB to be clinically meaningful in VF kinetics; however, no supporting evidence was found. Furthermore, it is not clear if the aforementioned MID is applicable to both kinetic and static fields.

The 25-Item Visual Function Questionnaire

The 25-item Visual Function Questionnaire (VFQ-25) is a 25-item version of the original 51-item National Eye Institute VFQ that is available in both self-administered and interviewer-administered formats. This patient-reported survey measures the effect of visual disability and visual symptoms on generic health domains such as emotional well-being and social functioning, in addition to ADL. The VFQ-25 includes 25 vision-targeted questions that represent 11 vision-related constructs: near vision (3 items), distance vision (3 items), social functioning (2 items), vision-specific role difficulties (2 items), dependency (3 items), mental health symptoms (4 items), driving difficulties (3 items), limitations with peripheral (1 item) and colour vision (1 item), and ocular pain (2 items). In addition, the VFQ-25 includes a single item that assesses general health status. Patients (or their parent or guardian, where applicable) provide a response to each of the 25 questions using a Likert scale ranging from 0 (worst visual function) to 100 (best visual function). Subscale-specific items are averaged to create summary scores for each subscale, and subscale scores are averaged to estimate the overall composite score.[54] It should be noted that the VFQ used in Study 301 was modified by the investigative team to evaluate the ADL that are dependent on vision, or have a vision element, with items related to HRQoL removed. The adaptations were done to accommodate IRD-associated poor vision and to include a pediatric population. The scoring system was changed too, with the perceived difficulty of these activities rated on a numerical scale of 0 to 10 (0 being the most difficult), before the average of the responses is taken to determine the numeric score for each individual.[15] Both the 25- and 51-item versions of the VFQ have been extensively validated in various ophthalmic conditions and in a variety of age groups;[55-58] however, the adapted version of the VFQ used in Study 301 was not assessed psychometrically. A number of studies assessed a clinically meaningful change for the standard VFQ-25;[49,59] however, an MID was not assessed for the adapted version of the scale, nor was any evidence for an MID for that version found in the literature search. Given the modifications made to the original VFQ, the MIDs identified in the literature for that measure were not considered directly generalizable to the version used in Study 301.

Safety

TEAEs were defined for the intervention group as adverse events with onset on or after the first injection, or onset prior to the first injection but with worsened severity after the first injection. For the control group, similar conventions apply with respect to the first day of

baseline evaluations. Presentations of TEAEs included all adverse events through 30 days after a patient discontinued from the study.

Statistical Analysis

The sample size and power calculation in Study 301 were based on simulations using a Wilcoxon rank-sum test with an exact P value. It was estimated that, with a 2:1 randomization ratio, a sample size of 16 patients in the treatment group and 8 patients in the control group would provide nearly 100% simulated power to detect an MLMT score change of 1 or more for the treatment arm compared with the control arm (no change) at a 2-sided type I error rate of 0.05.

The primary efficacy end point was the patient's bilateral performance (no eye patching) on the MLMT, measured by a change score 1 year following vector administration, compared with the patient's baseline bilateral MLMT performance. A non-parametric permutation test based on a Wilcoxon rank-sum test as the observed test statistic and an exact method for the corresponding P value was used for the analysis of the primary efficacy end point. The Wilcoxon rank-sum test statistic used the average rank when observations had the same value (i.e., were tied). The primary efficacy outcome was to be tested at a 2-sided type I error rate of 0.05.

The analysis of the secondary end point, MLMT score change using the first-treated eye, was to use the same statistical approach described for the primary efficacy end point.

For FST, a patient's response is light sensitivity as measured in dB, which was converted programmatically to the $\log_{10}(cd.s/m^2)$ to accommodate different dB conversion rates. VA was converted to the LogMAR. For the analysis of FST and VA, a separate model was used to assess the magnitude of the difference in response by comparing results 1 year after injections with those at baseline. A linear contrast from a repeated measures general linear model assessing change in response was used to estimate the magnitude of these effects. The models used the following categorical covariates: time as a fixed effect (defined by analysis windows, i.e., no restriction imposed on the trajectory of the mean outcome over time), treatment group, and time by treatment interaction.

The model used an unstructured correlation structure to model within-patient correlations. The estimated mean change from baseline to 1 year and its 95% CI was calculated from the model.

All statistical tests used 2-sided significance criteria of alpha = 0.05.

To maintain strict control of type I error rate, the 3 secondary outcomes were to be tested hierarchically:

- If the primary outcome was statistically significant, FST was to be tested at a 2-sided type I error rate of 0.05.
- If FST was statistically significant, the monocular mobility change score was to be tested at a 2-sided type I error rate of 0.05.
- If the monocular test was statistically significant, VA was to be tested at a 2-sided type I error rate of 0.05.

Handling of Missing Data

Missing data for the primary efficacy end point were to be handled as follows:

- If patients were removed from the study on the day of randomization, these patients were to be assigned a score change of 0.
- If 1 of the 2 MLMT scores using both eyes was missing at baseline or at 1 year, the data for that individual eye for the same light level was used to impute the missing score.
- If all data were missing for the baseline assessments, the screening data were to be used. If all data were missing for the 1-year assessments, the day 180 data were to be used to impute the missing data.
- If the light levels tested at baseline produced only passing scores, the screening results were to be used to establish the necessary cut-off levels.

Only 2 patients in the intention-to-treat population had missing data that affected the primary outcome. The 2 patients had only baseline data because they were removed from the study on the day of randomization and prior to any intervention; these patients were assigned a change score of 0 at year 1 for both the bilateral and unilateral tests.

The models for FST and VA did not impute for missing data.

Subgroup Analysis

No subgroup analyses were planned for Study 301. Subgroup analyses were conducted for age (< 18 years at first injection versus ≥ 18 years at first injection), which were post hoc analyses; no rational was provided for the age cut-off used in this post hoc subgroup analysis.

Sensitivity Analysis

The intention-to-treat (ITT) population was used as the primary analysis population for all analyses of efficacy end points. The modified ITT (mITT) and per-protocol (PP) populations were used for sensitivity and supportive analyses of the primary and key secondary efficacy end points.

For VA, measurements were collected using the scale adapted from Holladay et al.[50] to assign LogMAR for off-chart vision measurements. Following the start of the study, both the EMA and the study's data and safety monitoring board expressed the opinion that this scale, with estimates suggesting a difference of 1 log-unit step between counting fingers and hand-motion perception, could present a biased estimate of any treatment effect (improvement or reduction in LogMAR) for patients with off-chart measurements at baseline; both groups therefore recommended sensitivity analyses for VA using the scale proposed by Lange et al.,[51] in which the 1 log-unit step between counting fingers and hand motion is reduced to a 0.3 log-unit step.

Analysis Populations

The ITT population included all randomized patients.

The mITT population included all randomized patients who did not withdraw, or were not withdrawn, prior to any of the following people knowing the treatment assignment: the patient, parent, principal investigator, or medical monitor.

The per-protocol population included all patients in the ITT population who met all inclusion and exclusion criteria and did not withdraw or were not withdrawn prior to the patient,

parent, principal investigator, or medical monitor knowing the treatment assignment. For the voretigene neparvovec group, the per-protocol population excluded patients who did not receive treatment in both eyes.

The safety population included all patients in the voretigene neparvovec group who received an injection in either eye and all patients in the control group who did not withdraw, or were not withdrawn, prior to any of the following people knowing the treatment assignment: the patient, parent, principal investigator, or medical monitor.

Results

Patient Disposition

Overall, 36 patients were screened; 5 patients did not pass screening and 31 patients were randomized.

Patients failed screening due to the following reasons: mobility test performance (2 patients, where 1 patient passed at 1 lux and another patient had an accuracy score of greater than 1 at 400 lux during screening), attentional limitations (1 patient), screening VA and VF testing (1 patient), and lack of voluntary consent (1 patient).

Of the 31 patients randomized to the voretigene neparvovec group, 20 (95%) received bilateral injections of voretigene neparvovec. One patient was randomized to the voretigene neparvovec group but was withdrawn by the investigator prior to receiving the study drug (based on baseline OCT findings [severe retinal atrophy or degeneration]); 1 patient was randomized to the control group but was discontinued early due to withdrawn consent. All other patients completed the study through year 1 (Table 7).

After 1 year, 9 patients from the control group received voretigene neparvovec.

At the time of the data cut-off for the CSR, which provided updated results of available efficacy and safety data through [redacted], patients who received treatment with voretigene neparvovec had [redacted] after the second-eye injection visit, and [redacted] patients in the control/voretigene neparvovec group had [redacted] after the second-eye injection visit. In addition, [redacted] patients who were randomized to the voretigene neparvovec group had [redacted] after the second-eye injection visit and [redacted] patients in the control/voretigene neparvovec group had [redacted] after the second-eye injection visit (Table 7).

Table 7: Patient Disposition in Study 301

	Voretigene neparvovec	Control
Screened, n	36	
Randomized, n	21	10
Injected in both eyes, n (%)	20 (95)	NA
Discontinued early, n (%)		
Prior to any intervention	1 (5)	1 (10)
Reason for discontinuing early, n (%)		
Physician decision	1 (5)	0
Withdrawal by patient	0	1 (10)
Completed 1-year assessment, n (%)	20 (95)	9 (90)
ITT population,[a] n (%)	21 (100)	10 (100)
mITT population,[b] n (%)	20 (95)	9 (90)
PP population,[c] n (%)	19 (90)	9 (90)
Safety population,[d] n (%)	20 (95)	9 (90)
One year after randomization	**Voretigene neparvovec (n = 20)**	**Control/voretigene neparvovec[e] (n = 9)**
Injected in both eyes, n (%)	20 (100)	9 (100)
Completed 1 year after the second-eye injection, n (%)	20 (100)	9 (100)
Completed 2 years after the second-eye injection, n (%)	20 (100)	9 (100)
Completed 3 years after the second-eye injection, n (%)	20 (100)	8 (89)[f]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

ITT = intention to treat; mITT = modified intention to treat; NA = not applicable; PP = per protocol; vg = vector genome.

[a] Includes all randomized patients.

[b] Includes all randomized patients who did not withdraw, or were not withdrawn, prior to any of the following people knowing the treatment assignment: the patient, parent, principal investigator, or medical monitor. Patients in the mITT population were categorized by their randomized treatment assignment.

[c] Includes all ITT population patients who met all inclusion and exclusion criteria and who did not withdraw, or were not withdrawn, prior to the patient, parent, principal investigator, or medical monitor knowing the treatment assignment. For the intervention group, the PP population excluded patients who did not receive both injections.

[d] Includes all randomized patients who received an injection in either eye for the voretigene neparvovec group and all control-group patients who did not withdraw, or were not withdrawn, prior to any of the following people knowing the treatment assignment: the patient, parent, principal investigator, or medical monitor.

[e] Patients in the control group could receive voretigene neparvovec at a dose of 1.5×10^{11} vg in a total subretinal volume of 300 µL in each eye (within 18 days) 1 year after their baseline evaluations, provided they still met all eligibility criteria.

[f] One patient in the control/voretigene neparvovec group chose not to return to the intervention site for the 3-year assessment visit. Phone contact was made with the patient, questionnaires were completed, and concomitant medications and adverse events were discussed.

Source: Clinical Study Report for Study 301.[15]

Exposure to Study Treatments

Most patients (90%) in the voretigene neparvovec group received the protocol-specified, estimated total subretinal volume of 300 µL, while 2 patients (10%) received a reduced estimated total subretinal volume of either 200 µL or 250 µL in the right (first-injected) eye, resulting in estimated total effective doses of 1.0 $x10^{11}$ vg or 1.25 $x10^{11}$ vg for these patients' eyes, respectively.

A total of 20 intervention patients received bilateral injections of voretigene neparvovec. The overall mean (SD) time from randomization to first injection was 34.3 (27.9) days. This

was on the low end of the allowable window of 90 days, thereby minimizing the difference between treatment groups with respect to the timing of the year 1 assessments (± 30 days). The overall time from the first injection to the second injection was 8.8 (2.6) days.

The overall mean number of injection-retinotomy sites was 1.3 (SD = 0.4), with a similar number of sites used in the first (mean = 1.4; SD = 0.5) and second (mean = 1.2; SD = 0.5) eyes. The overall mean volume of vector injected was 305.0 µL (SD = 22.4 µL), with a similar volume injected into the first (mean = 310.0 µL; SD = 44.7 µL) and second (mean = 300.0 µL; SD = 0.0 µL) eyes. The overall mean estimated total subretinal volume delivered was 296.3 µL (SD = 12.2 µL), with a similar estimated total subretinal volume for the first (mean = 292.5 µL; SD = 24.5 µL) and second (mean = 300.0 µL; SD = 0.0 µL) eyes. The overall mean estimated air-fluid exchange was 83.0% (SD = 4.0%), with a similar estimated air-fluid exchange for the first (mean = 82.8%; SD = 4.1%) and second (mean = 83.3%; SD = 5.9%) eyes.

In patients in the control group who received voretigene neparvovec 1 year after randomization, the overall mean time from first to second injection was 7.7 (SD = 1.4) days. Most patients (90%) received the protocol-specified, estimated total subretinal volume of 300 µL. One patient (10%) received a reduced estimated total subretinal volume of 250 µL in the left (first-injected) eye.

Efficacy

Only those efficacy outcomes and analyses of subgroups identified in the review protocol are reported subsequently. See Appendix 3 for detailed efficacy data.

Multi-Luminance Mobility Test

The mean bilateral MLMT change score after 1 year was 1.8 (SD = 1.1) for the voretigene neparvovec group and 0.2 (SD = 1.0) for the control group. The difference in change from baseline in bilateral MLMT between voretigene neparvovec and the control treatment groups at 1 year was 1.6 (95% CI, 0.72 to 2.41; P = 0.001) in favour of voretigene neparvovec (Table 8).

For the secondary end point of the monocular MLMT change score after 1 year for the first eye, the mean change from baseline was 1.9 (SD = 1.2) for the voretigene neparvovec group and 0.2 (SD = 0.6) for the control group. The difference in change from baseline in monocular MLMT between voretigene neparvovec and the control treatment groups at 1 year was 1.7 (95% CI, 0.89 to 2.52; P = 0.001) in favour of voretigene neparvovec (Table 8).

For the supportive efficacy end point of the monocular MLMT change score after 1 year for the second assigned eye, the mean change from baseline was 2.1 (SD = 1.2) for the voretigene neparvovec group and 0.1 (SD = 0.7) for the control group. The difference in change from baseline in monocular MLMT between voretigene neparvovec and the control treatment groups at 1 year was 2.0 (95% CI, 1.14 to 2.85; nominal P < 0.001) (Table 8).

Results from the sensitivity analysis using the mITT and per-protocol populations for the primary outcome (change in bilateral MLMT performance at year 1 relative to baseline), and the sensitivity analyses using mITT population for the secondary efficacy end point (monocular MLMT change score after 1 year relative to baseline for the first assigned eye) and the supportive efficacy end point (monocular MLMT change score after 1 year relative

to baseline for the second assigned eye) were in the same direction as the primary analyses.

Figure 2 and Figure 3 present different magnitudes of patients' MLMT score change using both eyes from baseline to 1 year. Eleven patients (52%) in the voretigene neparvovec group had an MLMT score change of 2 or higher, of which 6 patients had an MLMT score change of more than 2. In contrast, only 1 patient or 10% of the control group had a score change of 2, and none of the patients in the control group had a score change greater than 2. Similar results for the MLMT score change using individual eyes were reported, where 15 patients (71%) in the treatment group had a score change of 2 or more when using each individual eye, while no patients in the control group had a score change of 2 or more.

The observed changes in the mean bilateral MLMT change score were stable through 4 years after the second-eye injection for the voretigene neparvovec group and through 3 years after the second-eye injection for the control/voretigene neparvovec group. For all patients who received treatment with voretigene neparvovec (n = 29), the mean (SD) bilateral MLMT change score 1 year after the second-eye injection was 1.9 (1.2), and these [redacted]. While only [redacted] from the voretigene neparvovec group had data available [redacted] after the second-eye injection; the mean bilateral MLMT change score appeared to [redacted] (Table 21).

Table 8: MLMT at Year 1 in Study 301

	Voretigene neparvovec n = 21	Control (n = 10)
Bilateral MLMT performance at 1 year compared with baseline (ITT)[a]		
Number of patients contributing to the analysis	21	10
Change from baseline, mean (SD)	1.8 (1.1)	0.2 (1.0)
Treatment group difference versus control (95% CI)	1.6 (0.72 to 2.41)	Reference
P value	0.001	
Monocular MLMT performance at 1 year compared with baseline for the first assigned eye (ITT)[a]		
Number of patients contributing to the analysis	21	10
Change from baseline, mean (SD)	1.9 (1.2)	0.2 (0.6)
Treatment group difference versus control (95% CI)	1.7 (0.89 to 2.52)	Reference
P value	0.001	
Monocular MLMT performance at 1 year compared with baseline for the second assigned eye (ITT)[a]		
Number of patients contributing to the analysis	21	10
Change from baseline, mean (SD)	2.1 (1.2)	0.1 (0.7)
Treatment group difference versus control (95% CI)	2.0 (1.14 to 2.85)	Reference
P value	< 0.001[b]	

CI = confidence interval; ITT = intention to treat; MLMT = multi-luminance mobility testing; SD = standard deviation.

[a] The observed 2-sided P value is from a Wilcoxon rank-sum test using an exact method. The permutation test P value was computed from all possible permutations.

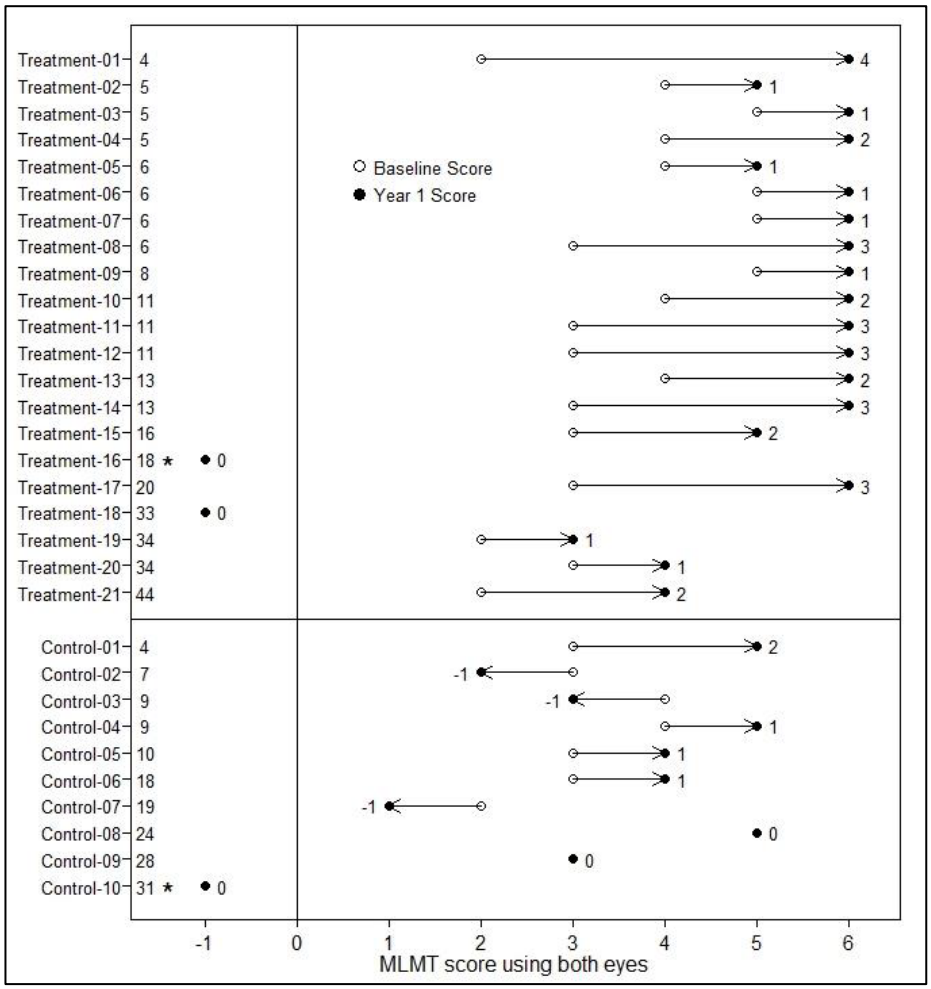
[b] P value was not adjusted for multiple testing.

Source: Clinical Study Report for Study 301.[15]

Figure 2: Bilateral MLMT Change Scores at Year 1 From Baseline, by Individual (ITT)

Figure 2 contained confidential information and was removed at the request of the sponsor.

Figure 3: Bilateral MLMT Scores at Baseline and Year 1, by Individual (ITT)



ITT = intention to treat; MLMT = multi-luminance mobility testing.

Source: FDA statistical review.[36]

The Full-Field Light Sensitivity Threshold

The FST results averaged over both eyes showed a mean change from baseline to year 1 of −2.08 (SE = 0.29) $\log_{10}$(cd.s/m^2) for the voretigene neparvovec group and 0.04 (SE = 0.44) $\log_{10}$(cd.s/m^2) for the control group, for a mean between-group treatment difference of −2.11 (95% CI, −3.19 to −1.04; P < 0.001) $\log_{10}$(cd.s/m^2) in favour of voretigene neparvovec (Table 9).

For the FST for the first assigned eye at 1 year, the difference between the voretigene neparvovec group and the control group was −2.33 (95% CI, −3.44 to −1.22; nominal P < 0.001) $\log_{10}$(cd.s/m^2) (Table 9).

For the FST for the second assigned eye at 1 year, the difference between the voretigene neparvovec group and the control group was −1.89 (95% CI, −3.03 to −0.75; nominal P = 0.002) $\log_{10}$(cd.s/m^2) (Table 9).

For the FST for blue light over both eyes, the difference between the voretigene neparvovec group and the control group was −2.10 (95% CI, −3.32 to −0.88; nominal P = 0.01) $\log_{10}$(cd.s/m^2) (Table 9).

For the FST for red light over both eyes, the difference between the voretigene neparvovec group and the control group was −1.46 (95% CI, −2.06 to −0.87; nominal P < 0.001) $\log_{10}$(cd.s/m^2) (Table 9).

Results from the sensitivity analysis using the mITT population for the secondary efficacy end point (change in white light FST averaged over both eyes at year 1 relative to baseline), and the exploratory end points (change in white light FST for the first assigned eye and change in white light FST for the second assigned eye at year 1 relative to baseline) were in the same direction as the primary analyses.

There were [redacted] after the second-eye injection for the voretigene neparvovec group and [redacted] after the second-eye injection for the control/voretigene neparvovec group. For both treatment groups, the [redacted] in FST performance was [redacted] following voretigene neparvovec administration, reflecting more than a [redacted] in light sensitivity. For [redacted] who received treatment with voretigene neparvovec [redacted], the mean change from injection baseline at [redacted] after the second-eye injection was [redacted] and these [redacted]. While the [redacted] voretigene neparvovec group had a change from baseline [redacted] after the second-eye injection, the FST [redacted] (Table 22).

Table 9: The Full-Field Light Sensitivity Threshold at Year 1 in Study 301

	Voretigene neparvovec n = 21	Control (n = 10)
Full-field light sensitivity testing: white light ($\log_{10}$[cd.s/m^2]) at 1 year (ITT)[a,b]		
Number of patients contributing to the analysis	20	9
Baseline, mean (SE)	−1.29 (0.09)	−1.65 (0.14)
At 1 year, mean (SE)	−3.36 (0.28)	−1.61 (0.42)
Change from baseline, mean (SE)	−2.08 (0.29)	0.04 (0.44)
Treatment group difference versus control (95% CI)	−2.11 (−3.19 to −1.04)	Reference
P value	< 0.001	
Full-field light sensitivity testing: white light ($\log_{10}$[cd.s/m^2]) for first assigned eye at 1 year (ITT)[b,c]		
Number of patients contributing to the analysis	20	9
Baseline, mean (SE)	−1.23 (0.10)	−1.65 (0.14)
At 1 year, mean (SE)	−3.44 (0.30)	−1.54 (0.44)
Change from baseline, mean (SE)	−2.21 (0.30)	0.12 (0.45)
Treatment group difference versus control (95% CI)	−2.33 (−3.44 to −1.22)	Reference
P value	< 0.001[d]	

	Voretigene neparvovec n = 21	Control (n = 10)
Full-field light sensitivity testing: white light (log_{10}[cd.s/m^2]) for second assigned eye at 1 year (ITT)[b,e]		
Number of patients contributing to the analysis	20	9
Baseline, mean (SE)	−1.35 (0.09)	−1.64 (0.14)
At 1 year, mean (SE)	−3.28 (0.29)	−1.69 (0.44)
Change from baseline, mean (SE)	−1.93 (0.31)	−0.04 (0.46)
Treatment group difference versus control (95% CI)	−1.89 (−3.03 to −0.75)	Reference
P value	0.002[d]	
Full-field light sensitivity testing: blue light (log_{10}[cd.s/m^2]) at 1 year (ITT)[b]		
Number of patients contributing to the analysis	20	9
Baseline, mean (SE)	−1.64 (0.11)	−1.99 (0.17)
At 1 year, mean (SE)	−3.61 (0.30)	−1.87 (0.44)
Change from baseline, mean (SE)	−1.97 (0.34)	0.13 (0.49)
Treatment group difference versus control (95% CI)	−2.10 (−3.32 to −0.88)	Reference
P value	0.001[d]	
Full-field light sensitivity testing: red light (log_{10}[cd.s/m^2]) at 1 year (ITT)[b]		
Number of patients contributing to the analysis	20	9
Baseline, mean (SE)	−1.21 (0.11)	−1.69 (0.16)
At 1 year, mean (SE)	−2.51 (0.18)	−1.53 (0.26)
Change from baseline, mean (SE)	−1.30 (0.17)	0.16 (0.24)
Treatment group difference versus control (95% CI)	−1.46 (−2.06 to −0.87)	Reference
P value	< 0.001[d]	

cd.s/m^2 = candela second per square metre; CI = confidence interval; ITT = intention to treat; SE = standard error.

[a] All measures were averaged over both eyes and then analyzed.

[b] Changes, 95% CIs, and P values were estimated using a repeated measures model with time, treatment, and time by treatment interaction.

[c] All measures were averaged across the first eye and then analyzed.

[d] P value was not adjusted for multiple testing.

[e] All measures were averaged across the second eye and then analyzed.

Source: Clinical Study Report for Study 301.[15]

Visual Acuity

For VA, when using the Holladay scale for off-chart results, the mean change across both eyes from baseline to year 1 was −0.16 (SE = 0.07) LogMAR for the voretigene neparvovec group and 0.01 (SE = 0.10) LogMAR for the control group, for a mean between-group treatment difference of −0.16 LogMAR (95% CI, −0.41 to 0.08; P = 0.17), an 8-letter difference (Table 10).

Using the Lange scale for off-chart VA results, the mean change across both eyes from baseline was −0.18 (SE = 0.04) LogMAR (a 9-letter improvement) for the voretigene neparvovec group and −0.03 (SE = 0.06) LogMAR (a 1.5-letter improvement) for the control group, for a mean between-group treatment difference of −0.16 LogMAR (95% CI, −0.31 to −0.01; nominal P = 0.035), a difference of 7.5 letters (Table 10).

Results from the sensitivity analysis using the mITT population for the secondary efficacy end point (change in VA averaged over both eyes at year 1 relative to baseline), and the exploratory end points (change in VA for the first assigned eye and change in VA for the

second assigned eye at year 1 relative to baseline) were in the same direction as the primary analyses.

The treatment effect of voretigene neparvovec on VA appeared to be maintained through 3 years after the second-eye injection, based on results in all patients who received treatment with voretigene neparvovec using the Holladay scale and Lange scale for off-chart results, where the mean change from injection baseline was −0.14 (SD = 0.30) LogMAR (a 7-letter improvement) at 1 year after the second-eye injection and −0.13 (SD = 0.32) LogMAR (a 7-letter improvement) at 3 years after the second-eye injection (Table 23). At [redacted] after the second-eye injection, the treatment effect of voretigene neparvovec on VA appeared to be [redacted], where the mean (SD) change from injection baseline was [redacted], that was mainly because [redacted] in the voretigene neparvovec group [redacted] after the second-eye injection visit (Table 23).

Table 10: Visual Acuity at Year 1 in Study 301

	Voretigene neparvovec n = 21	Control (n = 10)
Visual acuity (LogMAR) using Holladay scale for off-chart results at 1 year (ITT)[a]		
Number of patients contributing to the analysis	21	10
Baseline, mean (SE)	1.18 (0.14)	1.29 (0.21)
At 1 year, mean (SE)	1.03 (0.17)	1.30 (0.25)
Change from baseline, mean (SE)	−0.16 (0.07)	0.01 (0.10)
Treatment group difference versus control (95% CI)	−0.16 (−0.41 to 0.08)	Reference
P value	0.17	
Visual acuity (LogMAR) using Holladay scale for off-chart results for first assigned eye at 1 year (ITT)[b]		
Number of patients contributing to the analysis	21	10
Baseline, mean (SE)	1.31 (0.15)	1.37 (0.22)
At 1 year, mean (SE)	1.14 (0.19)	1.34 (0.28)
Change from baseline, mean (SE)	−0.17 (0.11)	−0.03 (0.16)
Treatment group difference versus control (95% CI)	−0.14 (−0.53 to 0.25)	Reference
P value	0.46	
Visual acuity (LogMAR) using Holladay scale for off-chart results for second assigned eye at 1 year (ITT)[c]		
Number of patients contributing to the analysis	21	10
Baseline, mean (SE)	1.06 (0.14)	1.21 (0.20)
At 1 year, mean (SE)	0.91 (0.15)	1.19 (0.21)
Change from baseline, mean (SE)	−0.15 (0.04)	−0.02 (0.06)
Treatment group difference versus control (95% CI)	−0.13 (−0.28 to 0.01)	Reference
P value	0.072	
Visual acuity (LogMAR) using Lange scale for off-chart results at 1 year (ITT)[a]		
Number of patients contributing to the analysis	21	10
Baseline, mean (SE)	1.16 (0.10)	1.15 (0.14)
At 1 year, mean (SE)	0.98 (0.11)	1.13 (0.17)
Change from baseline, mean (SE)	−0.18 (0.04)	−0.02 (0.06)
Treatment group difference versus control (95% CI)	−0.16 (−0.31 to −0.01)	Reference
P value	0.035[d]	

	Voretigene neparvovec n = 21	Control (n = 10)
Visual acuity (LogMAR) using Lange scale for off-chart results for first assigned eye at 1 year (ITT)[b]		
Number of patients contributing to the analysis	21	10
Baseline, mean (SE)	1.26 (0.10)	1.23 (0.15)
At 1 year, mean (SE)	1.06 (0.12)	1.21 (0.18)
Change from baseline, mean (SE)	−0.20 (0.05)	−0.02 (0.07)
Treatment group difference versus control (95% CI)	−0.18 (−0.36 to −0.01)	Reference
P value	0.044[c]	
Visual acuity (LogMAR) using Lange scale for off-chart results for second assigned eye at 1 year (ITT)[c]		
Number of patients contributing to the analysis	21	10
Baseline, mean (SE)	1.06 (0.10)	1.07 (0.14)
At 1 year, mean (SE)	0.91 (0.11)	1.05 (0.16)
Change from baseline, mean (SE)	−0.15 (0.04)	−0.02 (0.06)
Treatment group difference versus control (95% CI)	−0.13 (−0.27 to 0.01)	Reference
P value	0.076	

CI = confidence interval; ITT = intention to treat; LogMAR = logarithm of the minimum angle of resolution; SE = standard error.

[a] All measures were averaged over both eyes and then analyzed. Changes, 95% CIs, and P values were estimated using a repeated measures model with time, treatment, and time by treatment interaction.

[b] All measures were averaged across the first eye and then analyzed.

[c] All measures were averaged across the second eye and then analyzed.

[d] P value was not adjusted for multiple testing.

Source: Clinical Study Report for Study 301.[15]

Visual Field Testing

For the Goldmann VF V4e analysis, the mean change from baseline to year 1 was 78.8 (SD = 156.9) sum total degrees for the voretigene neparvovec group and −7.2 (SD = 341.4) sum total degrees for the control group, resulting in a mean treatment effect difference of 86.0 (95% CI, −186.1 to 358.1) sum total degrees (P = 0.67). For the Goldmann VF III4e analysis, the mean change from baseline to year 1 was 302.1 (SD = 289.6) sum total degrees for the voretigene neparvovec group and −76.7 (SD = 258.7) sum total degrees for the control group, resulting in a mean treatment effect difference of 378.7 (95% CI, 145.5 to 612.0) sum total degrees (nominal P = 0.006) (Table 11).

For the Humphrey VF foveal sensitivity analysis, the mean change from baseline to year 1 was 2.4 dB (SD = 9.7 dB) for the voretigene neparvovec group and 2.3 dB (SD = 5.3 dB) for the control group, resulting in a mean treatment effect difference of 0.04 dB (95% CI, −7.1 dB to 7.2 dB; P = 0.18). For the Humphrey VF macula threshold analysis, the mean change from baseline to year 1 was 7.7 dB (SD = 6.2 dB) for the voretigene neparvovec group and −0.2 dB (SD = 1.7 dB) for the control group, resulting in a mean treatment effect difference of 7.9 dB (95% CI, 3.5 dB to 12.2 dB; nominal P < 0.001) (Table 11).

For long-term treatment effect, for all patients who received treatment with voretigene neparvovec, the Goldmann VF III4e analysis showed a mean change from injection baseline of 267.4 (SD = 276.2) sum total degrees at 1 year after the second-eye injection of voretigene neparvovec, 268.6 (SD = 300.7) sum total degrees at 2 years after the second-eye injection, and 265.7 (SD = 248.8) sum total degrees at 3 years after the second-eye injection. For all patients who received treatment with voretigene neparvovec, the mean change from injection baseline for Humphrey VF foveal sensitivity was 2.6 dB (SD = 10.1 dB) at 1 year after the second-eye injection, 3.7 dB (SD = 8.3 dB) at 2 years after the

second-eye injection, and 3.3 dB (SD = 8.4 dB) at 3 years after the second-eye injection. For all patients who received treatment with voretigene neparvovec, the mean change from injection baseline for Humphrey VF macula threshold was 6.9 dB (SD = 7.5 dB) at 1 year after the second-eye injection, 6.7 dB (SD = 7.2 dB) at 2 years after the second-eye injection, and 6.6 dB (SD = 5.8 dB) at 3 years after the second-eye injection (Table 24).

Table 11: Goldmann Perimetry and Humphrey Computerized Testing at Year 1 in Study 301

	Voretigene neparvovec n = 21	Control (n = 10)
Goldmann VF V4e (sum total degrees) (ITT)[a]		
Number of patients contributing to the analysis	20	10
Baseline, mean (SD)	888.7 (487.8)	788.2 (482.9)
At 1 year, mean (SD)	1,032.8 (592.2)	778.8 (301.6)
Change from baseline, mean (SD)	78.8 (156.9)	−7.2 (341.4)
Treatment group difference versus control (95% CI)	86.0 (−186.1 to 358.1)	Reference
P value	0.67	
Goldmann VF III4e (sum total degrees) (ITT)[a]		
Number of patients contributing to the analysis	20	10
Baseline, mean (SD)	332.9 (413.3)	427.1 (372.0)
At 1 year, mean (SD)	673.9 (423.7)	397.8 (367.3)
Change from baseline, mean (SD)	302.1 (289.6)	−76.7 (258.7)
Treatment group difference versus control (95% CI)	378.7 (145.5 to 612.0)	Reference
P value	0.006[b]	
Humphrey VF, foveal sensitivity (dB) (ITT)[a]		
Number of patients contributing to the analysis	20	10
Baseline, mean (SD)	22.4 (6.8)	17.6 (8.9)
At 1 year, mean (SD)	25.8 (9.1)	21.5 (8.9)
Change from baseline, mean (SD)	2.4 (9.7)	2.3 (5.3)
Treatment group difference versus control (95% CI)	0.04 (−7.1 to 7.2)	Reference
P value	0.18	
Humphrey VF, macula threshold (dB) (ITT)[a]		
Number of patients contributing to the analysis	20	10
Baseline, mean (SD)	16.1 (5.5)	14.4 (8.0)
At 1 year, mean (SD)	24.0 (8.0)	15.8 (7.4)
Change from baseline, mean (SD)	7.7 (6.2)	−0.2 (1.7)
Treatment group difference versus control (95% CI)	7.9 (3.5 to 12.2)	Reference
P value	< 0.001[b]	

CI = confidence interval; dB = decibel; ITT = intention to treat; SD = standard deviation; VF = visual field.

[a] All measures were averaged over both eyes and then analyzed. The observed 2-sided P value is from a Wilcoxon rank-sum test.

[b] P value was not adjusted for multiple testing.

Source: Clinical Study Report for Study 301.[15]

Visual Function Questionnaire

For the patient-completed surveys, the mean change from baseline to year 1 was 2.6 (SD = 1.8) for the voretigene neparvovec group and 0.1 (SD = 1.4) for the control group, for a mean between-group treatment difference of 2.4 (95% CI, 1.0 to 3.8; nominal P = 0.001) (Table 12).

For the parent-completed surveys, the mean change from baseline to year 1 was 3.9 (SD = 1.9) for the voretigene neparvovec group and −0.2 (SD = 1.3) for the control group, for a mean between-group treatment difference of 4.0 (95% CI, 2.1 to 6.0; nominal P = 0.002) (Table 12).

For long-term treatment effect for the patient-completed surveys, for all patients who received treatment with voretigene neparvovec, the mean VFQ change from injection baseline was 2.2 (SD = 1.8) at 1 year after the second-eye injection, 1.8 (SD = 1.6) at 2 years after the second-eye injection, and 2.0 (SD = 1.7) at 3 years after the second-eye injection. For the parent-completed surveys, the mean VFQ change from injection baseline was 3.6 (SD = 1.8) at 1 year after the second-eye injection, 3.5 (SD = 1.7) at 2 years after the second-eye injection, and 3.4 (SD = 1.8) at 3 years after the second-eye injection (Table 25).

Table 12: Visual Function Questionnaire Average Scores (ITT) in Study 301

	Voretigene neparvovec n = 21	Control (n = 10)
Average score (patient)		
Number of patients contributing to the analysis	21	9
Baseline, mean (SD)	4.4 (1.4)	4.9 (1.5)
At 1 year, mean (SD)	7.0 (1.9)	5.1 (1.8)
Change from baseline, mean (SD)	2.6 (1.8)	0.1 (1.4)
Treatment group difference versus control (95% CI)	2.4 (1.0 to 3.8)	Reference
P value	0.001[a,b]	
Average score (parent)		
Number of patients contributing to the analysis	15	5
Baseline, mean (SD)	3.6 (1.3)	3.3 (1.7)
At 1 year, mean (SD)	7.5 (1.5)	3.1 (1.8)
Change from baseline, mean (SD)	3.9 (1.9)	−0.2 (1.3)
Treatment group difference versus control (95% CI)	4.0 (2.1 to 6.0)	Reference
P value	0.002[a,b]	

CI = confidence interval; ITT = intention to treat; SD = standard deviation.

[a] The observed 2-sided P value is from a Wilcoxon rank-sum test.

[b] P value was not adjusted for multiple testing.

Source: Clinical Study Report for Study 301.[15]

Subgroup Results by Age

A post hoc subgroup analysis was conducted for age (< 18 years at first injection versus ≥ 18 years at first injection).

The mean (SD) bilateral MLMT scores at injection baseline were [redacted] for the older (≥ 18 years at first injection) and younger patients (< 18 years at first injection), respectively. The mean (SD) MLMT sum change score 1 year after first injection was [redacted] for patients who were 18 years or older at first injection and [redacted] for patients who were younger than 18 years of age at first injection (Table 26).

For FST, patients who were 18 years or older at first injection had a mean (SD) for FST white light of [redacted] $\log_{10}(cd.s/m^2)$ at injection baseline, whereas the mean (SD) for FST white light for patients who were younger than 18 years at first injection was [redacted] $\log_{10}(cd.s/m^2)$. The mean (SD) FST white-light results for both eyes 1 year after

first injection was [redacted] log_{10}(cd.s/m^2) for patients who were 18 years or older at first injection and [redacted] log_{10}(cd.s/m^2) for patients who were younger than 18 years of age at first injection (Table 26).

For VA, patients who were 18 years or older at first injection had a mean (SD) of [redacted] LogMAR at injection baseline, whereas patients who were younger than 18 years at first injection had a mean (SD) of [redacted] LogMAR. The mean (SD) change from injection baseline to 1 year after first injection was [redacted] LogMAR [redacted] for patients who were 18 years or older at first injection and, for patients who were younger than 18 years of age at first injection, it [redacted] LogMAR, corresponding to an [redacted] on a standard eye chart (Table 26).

For VF, patients who were 18 years or older at first injection had mean (SD) results for Goldmann III4e of [redacted] sum total degrees and a Humphrey macula threshold of [redacted] dB analyses at injection baseline, whereas patients who were younger than 18 years at first injection had mean (SD) results for Goldmann III4e of [redacted] sum total degrees and a Humphrey macula threshold of [redacted] dB analyses at injection baseline. Patients who were 18 years or older at first injection had a mean (SD) change from injection baseline to 1 year after first injection of [redacted] sum total degrees for Goldmann III4e and [redacted] dB for Humphrey macula threshold, whereas patients who were younger than 18 years at first injection had mean (SD) changes of [redacted] sum total degrees for Goldmann III4e and [redacted] dB for Humphrey macula threshold (Table 26).

Harms

Only those harms identified in the review protocol are reported. See Table 13 for detailed harms data.

Adverse Events

All patients in Study 301 experienced at least 1 TEAE. During the first year after the randomization period, the most frequently reported TEAEs in the voretigene neparvovec group were leukocytosis (in 45% of patients); vomiting (in 40% of patients); nasopharyngitis, headache, and pyrexia (in 35% of patients for each); oropharyngeal pain, cough, and nausea (in 30% of patients for each); increased intraocular pressure (in 20% of patients); and cataract and hematuria (in 15% of patients for each). The most frequently reported TEAEs in the control group were oropharyngeal pain (in 44% of patients); upper respiratory tract infection (in 33% of patients); and vomiting, nasopharyngitis, and headache (in 22% of patients each) (Table 13). Overall, 13 patients (65%) in the voretigene neparvovec group had at least 1 TEAE considered to be related to the study drug administration procedure. In the voretigene neparvovec group, the TEAEs most often considered to be probably related to the administration procedure were cataract and increased intraocular pressure (n = 3 [15%] patients for each).

At the time of the data cut-off for the CSR, which provided updated safety data results through July 2, 2018, including follow-up for up to 5 years after the second injection for some patients, all 29 patients who received treatment with voretigene neparvovec experienced at least 1 TEAE. The most frequently reported TEAEs were headache (13 patients; 45%), leukocytosis (11 patients; 38%), and nausea and vomiting (10 patients; 34% for each) (Table 27). The most frequently reported ocular TEAEs were cataract and increased intraocular pressure (5 patients; 17% for each), and retinal deposits and retinal

tear (3 patients; 10% for each). Nineteen (66%) patients, including 13 patients (65%) in the voretigene neparvovec group and 6 patients (67%) in the control/voretigene neparvovec group, experienced TEAEs between the first injection and the data cut-off that were considered to be related to the study drug administration procedure. The most commonly occurring TEAEs considered related to the study drug administration procedure were cataract (5 patients; 17%), increased intraocular pressure (4 patients; 14%), nausea and retinal tear (3 patients; 10% for each).

Serious Adverse Events

During the first year after the randomization period, 2 patients (10%) in the voretigene neparvovec group experienced 3 SAEs at time points distant from vector administration. These SAEs were an adverse drug reaction and convulsion associated with a pre-existing complex seizure disorder in 1 patient and complications from oral surgery in another patient. These SAEs were considered unlikely to be related to the study drug or study drug administration procedure.

At the time of the data cut-off for the CSR, which provided updated results of safety data through July 2, 2018, including follow-up for up to 5 years after the second injection for some patients, 7 SAEs occurred in 5 patients in Study 301, including convulsion (1 event), adverse drug reactions (2 events) and retinal disorder (1 event of foveal thinning and loss of vision), retinal detachment (1 event), pneumonia (1 event), and menorrhagia (1 event) (Table 27). The retinal disorder and retinal detachment SAEs were considered related to the administration procedure.

Withdrawals Due to Adverse Events

No discontinuations from the study due to adverse events were reported in Study 301.

Mortality

No deaths were reported in Study 301.

Notable Harms

During the first year after the randomization period, 3 patients in the voretigene neparvovec group experienced a cataract, which were all mild in severity. Two patients in the voretigene neparvovec group experienced a retinal tear; 1 of the events was mild in severity and the other was moderate in severity; both events resolved with no sequelae. One patient developed an asymptomatic full-thickness macular hole that was noted 10 days after the first injection; this event improved to mild macular degeneration (thinning) that was noted 37 days after the first injection and recovered or resolved with no sequelae 90 days after the second injection.

During the follow-up, 1 patient who was originally in the control group and crossed over to voretigene neparvovec experienced a retinal disorder, which was foveal thinning and a loss of central vision. This retinal disorder was related to the subretinal injection in this patient, who had pre-existing atrophy of the retina. Another patient experienced 1 SAE of retinal detachment in the first-injected eye beginning 1,498 days after the voretigene neparvovec injection. This event was moderate in severity and considered related to the administration procedure.

Table 13: Summary of Harms at 1 Year in Study 301 (Safety)

	Voretigene neparvovec N = 20	Control (N = 9)
Patients with ≥ 1 AE		
n (%)	20 (100)	9 (100)
Most common events,[a] n (%)		
Leukocytosis	9 (45)	0
Vomiting	8 (40)	2 (22)
Nasopharyngitis	7 (35)	2 (22)
Headache	7 (35)	2 (22)
Pyrexia	7 (35)	1 (11)
Oropharyngeal pain	6 (30)	4 (44)
Cough	6 (30)	1 (11)
Nausea	6 (30)	1 (11)
Intraocular pressure increased	4 (20)	0
Cataract	3 (15)	0
Hematuria	3 (15)	1 (11)
Upper respiratory tract infection	2 (10)	3 (33)
Epistaxis	2 (10)	0
Nasal congestion	2 (10)	0
Diarrhea	2 (10)	1 (11)
Abdominal pain upper	2 (10)	0
Eye inflammation	2 (10)	0
Retinal tear	2 (10)	0
Adverse drug reaction	2 (10)	0
Animal bite	2 (10)	0
Ear infection	1 (5)	1 (11)
Hypertension	1 (5)	1 (11)
Patients with ≥ 1 SAEs		
n (%)	2 (10)	0
Most common events, n (%)		
Adverse drug reaction	2 (10)	0
Convulsion	1 (5)	0
Patients who stopped treatment due to AEs		
n (%)	0	0
Deaths		
n (%)	0	0
Notable harms, n (%)		
Treatment related notable harms		
Cataract development or progression	0	0
Endophthalmitis	0	0
Increased intraocular pressure	0	0
Retinal abnormalities	0	0

	Voretigene neparvovec N = 20	Control (N = 9)
Ocular inflammation	0	0
Procedure-related notable harms		
Vision loss	0	0
Retinal detachment	0	0
Retinal tear	2 (10)	0
Macular changes	1 (5)	0
Cataract	3 (15)	0

AE = adverse event; SAE = serious adverse event.

[a] AEs reported in more than 1 patient.

Source: Clinical Study Report for Study 301.[15]

Critical Appraisal

Internal Validity

Study 301 used accepted methods to conceal allocation where a randomization list was generated by an independent biostatistician. However, because it was an open-label trial, patients were aware of the treatment allocation following randomization. Therefore, the evaluation of patient-reported outcomes (such as those assessed by VFQ) and adverse events may be biased by treatment knowledge. The treatment effect on these subjective outcomes can potentially be overestimated as a consequence of the patient's expectation of the efficacy of a new drug. Certain steps have been taken to ensure appropriate blinding of the assessment of the primary outcome measure (MLMT) and it seems there was potential for an unbiased outcome assessment using the MLMT, despite the open-label design. On the other hand, while the trial states that orientation and mobility assessors were masked, there is insufficient detail provided in the CSR to judge if adequate blinding of outcome assessment was performed for all secondary outcome measures (e.g., VF and VA). However, it is considered unethical to perform sham subretinal surgery, and the procedure itself is not without risk.

Study 301 included 31 patients, which provided nearly 100% simulated power to detect an MLMT score change of 1 or more; however, a score change of 1 point might not be clinically relevant, as discussed subsequently, and it does not seem that the study was powered to detect an MLMT score change of 2 or more for the between-group difference. In addition, when the sample size is small, differences in 1 or 2 patients can have a substantial impact on results.

It is likely that patients in the voretigene neparvovec arm of Study 301 received best supportive care as background care; therefore, the treatment evaluated in this trial may be considered to be voretigene neparvovec plus best supportive care. There is limited detail about the nature of the best supportive care received by patients in the control arm of Study 301 and in the background of the treatment arm; the comparator is therefore unclear. It is also unclear whether the level of supportive care was similar and balanced between treatment groups. As some best supportive care interventions, such as the use of visual aids, may be associated with improvements in visual function, the presence or absence of such interventions in the control arm of Study 301 would aid interpretation of the reported effect size.

Imbalances in the baseline patient characteristics between the voretigene neparvovec and control groups included age and visual performance. It is unclear whether differences in baseline age between the voretigene neparvovec and control groups might introduce a risk of bias; on the one hand, the clinical experts indicated that age on its own should not be a criterion for treating patients but rather the presence of sufficient viable retinal cells. The imbalance at baseline in MLMT performance following assignment to the voretigene neparvovec or control group may bias the observed treatment effect estimate; however, the direction of the bias is unclear. As differences in MLMT, VA, and VF were noted at baseline in Study 301, and as these outcomes represent the key clinical data, there is uncertainty associated with the effect estimate for these outcomes, given that a larger percentage of patients with a baseline MLMT performance of less than 125 lux were enrolled in the voretigene neparvovec group (57%) than in the control group (40%). Also, at baseline, patients randomized to the voretigene neparvovec group had, on average, better acuity (less VA loss), and better VF than patients randomized to the control group. However, there is no evidence to indicate how visual performance at baseline could affect the treatment effect and, if bias exists, it is not possible to judge the direction of the bias.

While there were unexpected dropouts, these were accounted for appropriately, as the patients who discontinued were removed from the study on the day of randomization and these patients were assigned a change score of 0 at year 1 for both bilateral and unilateral tests. There was a 10% dropout rate in the control group compared with a 5% dropout rate in the voretigene neparvovec group. Explanations for discontinuation were unrelated to treatment with voretigene neparvovec and included personal reasons and severe retinal atrophy that precluded participation.

Scores on the MLMT may underestimate the treatment effect of voretigene neparvovec due to the potential ceiling effect, where patients who passed the test at the second-lowest light level at baseline were able to achieve only a maximum increase of 1 unit. Although 62% of patients in the voretigene neparvovec group achieved the maximum possible increase following administration compared with none in the control group, the observed mean increase of 1.8 in MLMT score may be an underestimate of the treatment effect because of this ceiling effect. The MLMT also potentially has a floor effect, where 1 patient who was in the voretigene neparvovec group did not pass the MLMT at the highest light level at baseline and also failed at year 1. It is worth noting that out of the 5 patients who did not pass screening, 2 patients (40%) were not eligible based on mobility test performance, with 1 patient excluded due to ceiling effect and another patient excluded due to floor effect. In addition, it is not clear whether the 12 unique MLMT course configurations were of equivalent difficulty.

The MLMT was developed by the sponsor and there is some uncertainty over the accepted definition of a clinically relevant improvement. The sponsor noted that an MLMT change score from baseline of one or more lux levels should be considered clinically meaningful. However, the FDA indicated that an MLMT score change of 1 level may represent a background fluctuation occurring in both the treatment and control groups, and that an MLMT score change considered clinically meaningful would be 2 levels or greater.[37] Similarly, the EMA indicated that any clinically relevant change in MLMT with voretigene neparvovec would need to exceed 1 light level, given that patients in the control group had a change in score of 1 ± 1 over 1 year; however, the EMA did not indicate what difference would be considered clinically meaningful.[38]

The clinical experts indicated that the outcome measures of VA and VF used in the trial are highly specialized, requiring administration by special centres, and are not generally used in clinical practice. Also, it is worth noting that the NICE guidance document for voretigene neparvovec indicated that VA and VF are often considered unreliable because of inter-test variability.[53]

No subgroup analyses were planned for Study 301. A post hoc subgroup analysis by age (< 18 years at first injection versus ≥ 18 years at first injection) was conducted, but the statistical methodology for this analysis was not provided. In addition, these analyses were post hoc and thus should be considered hypothesis-generating. Also, it is not clear if this subgroup analysis was conducted at the request of a regulator, and no rationale was provided for the age cut-off.

The controlled comparison is up to 1 year. No comparison between treatment groups was conducted beyond 1 year.

For Goldmann VF V4e, 1 year and 2 years after the second-eye injection, only 11 patients (55%) were included in the voretigene neparvovec group and 1 patient (11.1%) was included in the control/voretigene neparvovec group. Three years after the second-eye injection and four years after the second-eye injection, no patients were included in the control/voretigene neparvovec group. Why there were that many missing data was not reported and there was no accounting for missing data. This calls the study finding of this outcome into question.

Contrast sensitivity, an outcome identified as important by the clinical experts, was an exploratory outcome in Study 301; however, the CSR provided by the sponsor did not report detailed results for this outcome, and the clinical reviewer was not able to interpret the results from this outcome.

The patient-reported outcome measure, the VFQ, which consisted of 25 questions pertaining to ADL that are dependent upon vision or have a vision component, was completed by patients or their parent or guardian, where applicable. Also, for the parent-completed questionnaire, the need for “proxy reporting” by parents or caregivers adds uncertainty as to the validity of the information reported. Finally, the VFQ used was modified by the investigative team from the VFQ-25. While the VFQ-25 is validated and has an estimated MID, the VFQ used in Study 301 was not assessed psychometrically and the MIDs identified in the literature for the VFQ-25 were not considered directly generalizable to the version used in Study 301. All of these limitations call the VFQ study findings into question.

External Validity

While the patients in Study 301 were required to have a diagnosis of LCA due to *RPE65* mutations to be enrolled in the pivotal trial of voretigene neparvovec, the clinical experts indicated that the results reported for these patients should be generalizable to patients with RP as long as they have confirmed biallelic *RPE65* mutations. Voretigene neparvovec is designed to deliver a normal copy of the *RPE65* gene to cells of the retina in individuals with reduced or absent levels of biologically active RPE65[13] and is, therefore, intended to treat the underlying mechanism of the disease, which is the same in patients with confirmed *RPE65* mutations, regardless of clinical phenotype.[13]

The clinical experts indicated that measuring viable retinal cells is not a straightforward procedure, and that the methods used in the pivotal trial to determine whether a patient has

sufficient viable retinal cells do not give a complete picture of the health of the retina or the number of viable photoreceptors. They also indicated that the technology to assess the structure and function of the retina has evolved since the trial. The clinical experts indicated that OCT can measure the thickness of the retina; however, it may not inform the treating physician whether there are viable retinal cells. The clinical experts also indicated that a numerical cut-off could not be applied universally throughout all OCT technologies and generations because of the OCT changes and the differences between measurements on different OCT technologies. In clinical practice, the presence of sufficient viable retinal cells would be determined by the treating physician using OCT examinations measuring the area of remaining viable photoreceptors, which would be supplemented by tests for the function of the retina, such as VA and visual function.

Study 301 aims to follow patients for up to 15 years after treatment. The current duration of follow-up is limited to 5 years and longer-term efficacy and safety data for voretigene neparvovec are awaited. While, as per the latest cut-off date of the CSR, the effect of voretigene neparvovec appears to be maintained for at least up to 5 years, the duration of treatment effect beyond 5 years is unclear. In addition, the clinical experts consulted by CADTH anticipated a waning in the effect over time. The clinical experts also indicated there is no information on whether patients who lose treatment effect would benefit from re-treatment.

Results of Study 301 may not be generalizable to patients who do not share the characteristics of the trial population, such as patients who are younger than 4 years of age, and patients who have better or worse baseline functional vision and visual function than those enrolled in the trial. The clinical experts consulted by CADTH did not anticipate that voretigene neparvovec would be used in patients who are younger than 4 years of age, as it is difficult to perform the procedure in these patients. Study 301 excluded patients who had participated in a study in which a gene therapy vector was administered, The clinical experts indicated that, currently, there is limited knowledge regarding whether patients who have received other gene therapies should receive voretigene neparvovec, or if patients can receive more than 1 gene therapy.

Study 301 included only 31 patients; however, this needs to be considered in view of the rarity of the condition being treated. On the other hand, differences in 1 or 2 patients can have a substantial impact on results and the ability to capture rare adverse events.

Indirect Evidence

No indirect evidence was submitted by the sponsor. An independent literature search was conducted by CADTH, but no indirect evidence was identified that met the inclusion criteria of the CADTH review protocol.

Other Relevant Evidence

This section includes submitted long-term extension studies and additional relevant studies included in the sponsor's submission to CADTH that were considered to address important gaps in the evidence included in the systematic review.

Single-Arm Studies: Study 101 and Study 102

The sponsor submitted 2 supportive trials, Study 101 and Study 102. Study 101 was a dose-escalation study that assessed the safety and tolerability of 3 different doses of

voretigene neparvovec administered via subretinal administration to 1 eye (first-treated eye) of patients with a molecular diagnosis of LCA due to *RPE65* mutations. Study 101 also evaluated the clinical efficacy. Study 102 was a follow-on study of Study 101 in which patients received voretigene neparvovec treatment in the previously uninjected eye (second-treated eye). Details of the trials' characteristics are provided in Table 14.

Table 14: Details of Studies 101 and 102

		Study 101	Study 102
Designs & Populations	**Study design**	Phase I, open-label, single-arm, dose-escalation, safety study	Phase I, open-label, single-arm, safety study
	Locations	US	US
	Enrolled (N)	12	11

	Inclusion criteria	• 8 years of age or older at time of administration
		• Diagnosis of LCA
		• Molecular diagnosis confirmed due to *RPE65* mutations (homozygotes or compound heterozygotes) by a CLIA-certified laboratory
		• Visual acuity ≤ 20/160 or visual field less than 20° in the eye to be injected
			• Prior participation in initial phase I Study 101 with unilateral, subretinal administration of voretigene neparvovec
			• Visual acuity equal to or greater than light perception
			• Sufficient viable retinal cells in the contralateral, previously uninjected eye, as determined by OCT and/or ophthalmoscopy. Must have one of the following:
			  • an area of retina within the posterior pole of > 100 µm as shown on OCT
			  • ≥ 3 disc areas of retina without atrophy or pigmentary degeneration within the posterior pole
			  • a remaining visual field within 50° of fixation

	Exclusion criteria	• Insufficient viable retinal cells, as determined by OCT and/or ophthalmoscopy (e.g., areas of retina with thickness measurements less than 100 µm or absence of neural retina) • Neutralizing antibodies to AAV2 > 1:1,000 • Pre-existing eye conditions that would preclude the planned surgery or interfere with the interpretation of study end points (e.g., glaucoma, corneal or lenticular opacities) • Ocular surgery within the previous 6 months	• Pre-existing eye condition, such as glaucoma, or complicating systemic diseases • Prior intraocular surgery within the past 6 months
Drugs	**Intervention**	Patients received a single unilateral dose of voretigene neparvovec by subretinal administration. 3 dose cohorts were studied: • Dose cohort 1: 1.5×10^{10} vg of voretigene neparvovec in 150 µL administered subretinally	• 1.5×10^{11} vg of voretigene neparvovec in 300 µL administered subretinally into the contralateral eye (second-treated eyes)

		Study 101	Study 102
		• Dose cohort 2: 4.8 × 10^{10} vg of voretigene neparvovec in 150 µL administered subretinally • Dose cohort 3: 1.5 × 10^{11} vg of voretigene neparvovec in 300 µL administered subretinally	
	Comparator(s)	None	None
Duration	**Phase**		
	Open-label	1 year	1 year
	Follow-up	15 years after first injection	15 years after first injection
Outcomes	**Primary end point**	• Safety	• Safety
	Secondary and exploratory end points	• Visual acuity • Visual field testing (Goldmann perimetry) • Contrast sensitivity • Mobility testing • Quality of life assessments	• Mobility testing • FST testing • Visual acuity • Visual field testing (Goldmann perimetry) • Contrast sensitivity
Notes	**Publications**	Maguire et al.[60]	Bennett et al.[61]

AAV2 = adeno-associated virus type 2; CLIA = Clinical Laboratory Improvement Amendments of 1988; FST = full-field light sensitivity threshold; LCA = Leber congenital amaurosis; OCT = optical coherence tomography; vg = vector genome.

Source: Maguire et al.,[60] Bennett et al.,[61] Maguire et al.,[62] Simonelli et al.,[63] Testa et al.,[64] Bennett et al.,[65] and the Clinical Study Report for Study 101[66] and Study 102.[67]

Methods

Study 101 (N = 12) was a phase I, open-label, single-arm, dose-escalation study that assessed the safety and tolerability of 3 different doses of voretigene neparvovec administered via subretinal injection in 1 eye (first-treated eye) of patients with LCA due to *RPE65* mutations. Twelve patients, who were 8 years of age or older at the time of administration, were to receive a unilateral subretinal injection in the eye with the worse function (first-treated eye). Three doses of voretigene neparvovec were tested sequentially: 1.5 × 10^{10} vg of voretigene neparvovec, 4.8 × 10^{10} vg of voretigene neparvovec, and 1.5 × 10^{11} vg of voretigene neparvovec. The dose that was submitted for approval to Health Canada was 1.5 × 10^{11} vg of voretigene neparvovec.

Study 102 was a follow-on study to Study 101. Eleven of the 12 treated patients in Study 101 received a subretinal injection in the contralateral eye (second-treated eye) of 1 dose at 1.5 × 10^{11} vg of voretigene neparvovec in a total volume of 300 µL. One of the 12 patients in Study 101 was not treated in the second eye because they had glaucomatous changes in the eye to be treated, which was an exclusion criterion. The interval between the first- and second-eye injections ranged from 1.7 to 4.6 years.

The study duration of both Study 101 and Study 102 was 1 year, with an extended long-term follow-up planned for a total of up to 15 years.

The sponsor indicated that study participants were assessed for up to 7.5 years following administration of voretigene neparvovec, and observation is ongoing. However, the most recent CSR addendum submitted by the sponsor provided updated results of available efficacy and safety data through July 2, 2018. This CSR was finalized on May 19, 2019. At

the time of the data cut-off for this CSR, [redacted]

Populations

Inclusion and Exclusion Criteria

Patients enrolled in Study 101 had to be at least 8 years of age at the time of administration, with a molecular diagnosis of LCA due to *RPE65* mutations confirmed by a CLIA-approved laboratory. Also, patients had to have VA no better than 20/160 or a VF of less than 20° in the eye to be injected. Patients were excluded from Study 101 if they lacked sufficient viable retinal cells as determined by non-invasive means, such as OCT or ophthalmoscopy. Specifically, if indirect ophthalmoscopy reveals less than 1 disc area of retina that is not involved in complete retinal degeneration (indicated by geographic atrophy, thinning with a tapetal sheen, or confluent intraretinal pigment migration), those eyes were excluded. In addition, in eyes that OCT scans of sufficient quality could be obtained, if there were areas of the retina with thickness measurements of less than 100 µm or if a neural retina was absent, the eyes were not targeted for delivery of voretigene neparvovec. Patients were excluded if immunological studies showed the presence of neutralizing antibodies to adeno-associated virus type 2 above 1:1,000. In addition, patients were excluded if they had a pre-existing eye condition that would preclude the planned surgery or interfere with the interpretation of study end points (e.g., glaucoma, corneal or lenticular opacities), or if they had ocular surgery within the previous 6 months.

Patients enrolled in Study 102 had to be participants of Study 101, with VA no worse than light perception, and have sufficient viable retinal cells in the contralateral, previously uninjected eye, as determined by OCT or ophthalmoscopy. Patients had to have either an area of retina within the posterior pole of greater than 100 µm (as shown on OCT), 3 or more disc areas of retina with no atrophy or pigmentary degeneration within the posterior pole, or a remaining VF within 50° of fixation. Patients were excluded if they had a pre-existing eye condition, such as glaucoma or complicating systemic diseases, or if they had prior intraocular surgery within 6 months.

Baseline Characteristics

Twelve patients were enrolled in Study 101; of those, 7 patients (58%) were male. The mean age at the time of the study drug administration was 20.8 (SD = 11.2) years with a range of 8 to 44 years. All patients enrolled in the study were non-Hispanic; the majority of the patients were White (92%).

Study 102 enrolled 11 patients; of those, 6 patients (55%) were male. The mean age at the time of the study drug administration was 22.8 (SD = 10.28) years with a range of 11 to 46 years. All patients enrolled in the study were non-Hispanic; the majority of the patients were White (91%).

Interventions

The dose of voretigene neparvovec was defined based on both the vg and subretinal injection volume (µL). Three dose levels were sequentially tested for the first-eye injection in Study 101. In total, 3 patients received the low dose (1.5×10^{10} vg of voretigene neparvovec in 150 µL), 6 patients received the middle dose (4.8×10^{10} vg of voretigene

neparvovec in 150 µL), and 3 patients received the high dose (1.5×10^{11} vg of voretigene neparvovec in 300 µL).

For the second-eye injection in Study 102, the highest dose (1.5×10^{11} vg of voretigene neparvovec in 300 µL) was used, which was concluded to be safe in Study 101.

Patients in both studies were given systemic corticosteroids beginning 3 days before the vector injection. The initial dose was 1 mg/kg per day of prednisone for 10 days, with a maximum prescribed dose of 40 mg per day regardless of the weight of the patient; this was followed by 0.5 mg/kg per day for an additional 7 days with a maximum prescribed dose of 20 mg per day, regardless of the weight of the patient. The steroids were given to reduce the potential for an immune response to the voretigene neparvovec capsid and transgene product.

The following ocular corticosteroid and prophylactic antibiotics were allowed to be used during the studies:

- sub-tenon/retrobulbar infusion of 1 mL triamcinolone (40 mg/mL)
- subconjunctival injection of 0.5 mL of 4 mg/mL dexamethasone solution
- topical ocular dressing with 2.5 cm of prednisolone acetate 0.6%, gentamicin sulfate 0.3% or tobramycin 0.3%, or dexamethasone 0.1% ointment
- 0.5 mL of 50 mg/mL vancomycin or cefazolin.

Patients were not permitted to take the following medications once they were enrolled in the study:

- investigational drugs other than voretigene neparvovec
- high dose of vitamin A (> 7,500 retinol equivalent units or > 3,300 IU per day)
- tretinoin-containing skin cream
- isotretinoin
- sildenafil or related compounds used to treat erectile dysfunction
- hydroxychloroquine, chloroquine, thioridazine, or any related retinotoxic compounds.

Outcomes

In Study 101 and Study 102, the primary and secondary efficacy end points were not pre-specified. The following efficacy outcomes were assessed in Study 101 and Study 102:

- Mobility testing, which assessed the patient's ability to navigate an obstacle course:
  - In Study 101, mobility testing, under defined lighting conditions, evaluated the patient's ability to navigate through a course requiring the patient to follow a path defined by large black arrows on tiles on the floor and to avoid obstacles placed in and around the path. For this test, patient performance was timed and video-recorded to provide documentation. The mobility testing protocol was refined over the course of the study; this affected both the interpretation of the results and the number of patients considered evaluable.
    - Mobility testing was initially carried out only under office lighting conditions (200 lux to 250 lux). Additionally, for the patients followed initially at the former referral/follow-up site in Italy, testing was erroneously carried out with both eyes together (i.e., neither eye patched); thus, it was not possible to determine which eye, if either, was functioning better. Further, the lighting conditions and the degree of difficulty of the courses used at that site were not consistent between visits.

- For patients beginning in the middle-dose cohort, the mobility testing protocol was expanded to include testing under scotopic (low light) conditions, the conditions that elicit responses from rod photoreceptors. The test design was also refined during the course of the study such that courses were standardized to contain a specified number of turns and numbers of specific types of obstacles. The video-recording protocol was also refined, thereby enabling more accurate scoring of the tests. Prior to standardization, the "pass" versus "fail" designation was determined based on the percentage of obstacles avoided, where avoiding more than 75% of obstacles was typically designated as a pass. Also considered was whether the patient was able to identify 3 specific landmarks in the mobility course; if the patient was not able to identify 2 of the 3 landmarks, the result was designated as a fail, regardless of the percentage of obstacles avoided. When mobility testing was initiated, the investigators had thought the ability to identify landmarks might be informative; however, greater experience with the test showed this information could be captured by analyzing collisions with the various obstacles on the course. To optimize mobility testing procedures, the following standardization was implemented:

- Testing was carried out at baseline to determine the estimated lower light level (for each eye) at which a patient could carry out mobility testing.
- Light levels were standardized to a predetermined series of specified light levels.
- The course was standardized so that both intra-patient and inter-patient test results could be compared more accurately.
- Twelve different test configurations with equal difficulty in terms of distance, number of turns, and number of obstacles were developed, and these were selected randomly for each test in order to minimize any learning effect.
- Procedures for documenting the test performance (with a video recording) were optimized.
- Detailed instructions on how to carry out baseline and post-injection mobility testing, and how to score the tests, were assembled into a standard operating procedure.
- Independent masked readers scored all video recordings.
  - The scoring paradigm used, namely, weighting the accuracy and speed for each test according to a defined algorithm that places greater emphasis on accuracy than speed, includes both accuracy and time penalties but does not rely directly on landmark identification. To reduce redundancies in the test, the identification of landmarks was omitted, since identification of these landmarks was essentially already required for the timely and accurate completion of the course.
- In Study 102, the testing procedure was the same as described in Study 101. There were 2 types of reporting for mobility testing, namely, the non-standardized form used for the initial visits of the first 3 patients enrolled in Study 102 and the standardized form used for the remaining 8 patients.
- In refining the outcome measures for Study 301, the number of standardized light levels was decreased from 9 to 7 in order to show a greater distinction in patient performance between the testing levels; therefore, for data analyses and figure presentation, the light levels were consolidated as follows: 100 lux and 150 lux (Study 301 = 125 lux) and 200 lux and 250 lux (Study 301 = 250 lux). By using this conservative approach and combining the light levels, the maximum change score decreased from 8 to 6 units and some patients' change scores could be decreased by as much as 2 units.
- FST testing, a measure of retinal light sensitivity, with lower (more negative) numbers reflecting increased white-light sensitivity.

- VA measures were used to document any change in the ability to discern shapes or letters of different sizes.
- VF (Goldmann perimetry) parameters were designed to evaluate alterations in the function of different regions of the retina.

Statistical Analysis

The sample size for Study 101 was chosen based on the principle that a minimum number of patients should be exposed to potential risks in a phase I study in order to obtain valid information to proceed with further investigation. Overall, it was considered that the administration of voretigene neparvovec in up to 9 patients 8 year of age and older might allow for an estimation of the likely best dose and provide evidence of product safety. After preliminary evaluations of the first 6 patients injected with voretigene neparvovec, a total of 12 patients were included in the safety and efficacy population: 3 at the low dose of 1.5×10^{10} vg, 3 at the middle dose of 4.8×10^{10} vg, 3 additional patients who were subsequently added to the middle-dose cohort, and an additional 3 patients who received the high dose of 1.5×10^{11} vg.

No formal hypothesis testing was conducted. Descriptive statistics (mean, standard deviation, median, minimum, and maximum values) were presented for each of the evaluable parameters for change from baseline as well as at each time point. Missing values were treated as missing without any imputation.

Analysis Populations

The analyzed efficacy and safety populations included all patients who received the study drug.

Patient Disposition

Table 15 summarizes the patient disposition of Study 101 and Study 102. Twelve patients enrolled in Study 101; there were no discontinuations. One patient treated with the high dose in Study 101 was not enrolled in Study 102 due to glaucoma changes identified through examination (pre-existing).

Table 15: Patient Disposition for Studies 101 and 102

	Low-dose voretigene neparvovec (1.5×10^{10} vg)	Middle-dose voretigene neparvovec (4.8×10^{10} vg)	High-dose voretigene neparvovec (1.5×10^{11} vg)	Total
Study 101				
Enrolled, N	[redacted]	[redacted]	[redacted]	[redacted]
Treated, N	[redacted]	[redacted]	[redacted]	[redacted]
Discontinued, n	[redacted]	[redacted]	[redacted]	[redacted]
Study 102				
Enrolled, N	[redacted]	[redacted]	[redacted]	[redacted]
Treated, N	[redacted]	[redacted]	[redacted]	[redacted]
Discontinued, n	[redacted]	[redacted]	[redacted]	[redacted]

vg = vector genome.

Source: Clinical Study Report for Study 101[66] and Study 102.[67]

Exposure to Study Treatments

In Study 101, all patients in the low-dose group received the protocol-specified dose of the study drug (1.5×10^{10} vg). In the middle-dose group, 1 patient received 2 injection attempts due to foveal dehiscence with the first injection attempt. Due to this dehiscence and the loss of the study drug to the intravitreal space, the patient received an estimated subretinal volume of 100 µL (estimated dose 3.2×10^{10} vg) as opposed to the protocol-specified volume of 150 µL. In the high-dose group, 1 patient received a total injected volume of 350 µL for an estimated subretinal volume of 300 µL; thus, all patients in this dose group received the protocol-specified subretinal dose of the study drug (1.5×10^{11} vg).

In Study 102, all patients received the protocol-specified dose of the study drug. The total dose administered to all patients was 1.5×10^{11} vg and the total volume injected was 300 µL. There were no injection-related or surgical (non-injection) complications reported for any patient at the time of vector administration. All injections were successfully administered with 1 injection retinotomy site.

Efficacy

All data are descriptive, and no inferences can be made because there were no hypotheses being tested.

Mobility Testing

The mobility testing results reported subsequently should be interpreted with caution, as these results were obtained using a very primitive form of the MLMT.

In Study 101, 4 patients were considered non-evaluable given the inconsistent use of patching, as well as the variability of lighting conditions and test-course difficulty. Following vector administration, follow-up mobility testing using the injected eye indicated that 4 of the 8 evaluable patients were able to complete the mobility test at a light level that was at least 1 level darker than baseline.

In Study 102, monocular mobility testing was assessed for change from baseline for the eye injected in Study 102. Of the 11 patients who received voretigene neparvovec, 8 patients were considered evaluable for mobility testing. Three patients were not included for presentations of this parameter, as they were not evaluable on mobility testing. One year after their eyes were injected in Study 102, and using the eyes injected in Study 102, all 8 evaluable patients completed the mobility test at a light level that was at least 1 level darker than baseline (Table 16). Five of the 8 (63%) evaluable patients received the maximum attainable score. At the year 2 visit, similar results were recorded for 7 of the 8 evaluable patients, as 1 patient completed the mobility test at a light level that was 1 level darker than the level passed at year 1. Consistent with earlier findings, mobility testing results continued to show a relatively stable change from baseline in functional vision at lower light levels in all 8 evaluable patients through year 4, with 6 patients completing the mobility test at a light level that was 2 levels darker than baseline. At the time of database lock for the CSR provided, [redacted]. Of the evaluable patients with new data during the time period of this CSR report, most patients continued to show a relatively stable change from baseline in functional vision at lower light levels, where [redacted] evaluable patients completed the mobility test at a light level that was [redacted]

[redacted] patients completed the mobility test at a light level that was [redacted] than baseline, and [redacted] the mobility test at the [redacted] (Table 16).

Table 16: Summary of Change in Mobility Testing Scores From Baseline (Evaluable Patients) for Study 102

Patient	Number of years after eye injection						
	1 year	2 years	3 years	4 years	5 years	6 years	7 years
Patient 1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patient 2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patient 3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patient 4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patient 5	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patient 6	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patient 7	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patient 8	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

NA = not available.

Values represent the monocular mobility test change from baseline for the eye injected in Study 102.

[a] Patient failed the mobility test at the highest tested light level (400 lux).

Source: Clinical Study Report for Study 102.[67]

FST Testing

For the eye injected in Study 101, 4 of 7 (57%) evaluable patients at year 1 showed a decrease of 10 dB or more in FST compared with baseline. The change from baseline in all 7 evaluable patients ranged from −8.1 dB to −33.7 dB. At year 2, 3 of 7 (43%) evaluable patients showed a decrease of 10 dB or more in FST, with a change from baseline ranging from −5.7 dB to −29.1 dB in all evaluable patients, apart from 1 patient.

At the time of the database lock on May 19, 2019, for the Study 102 CSR, the available FST data indicated that light sensitivity in the Study 102 injected eye increased after injection for 8 of 11 patients; this increase was greater than the 10 dB cut-off considered clinically important in 7 of 11 patients, and was slightly below this cut-off for 1 patient. Light sensitivity remained essentially stable in the other 3 patients.

In Study 102 at the year 4 visit, the mean of the FST score of the injected eye was −13.09 (SE = 4.17 dB), which was a decrease when compared with the FST score of the injected eye at baseline (mean = 1.40; SE = 2.79 dB), indicating increased light sensitivity (i.e., responded at a lower light intensity) after injection, which was sustained over the follow-up period (Table 17).

Table 17: Full-Field Light Sensitivity Threshold Testing Results for Studies 101 and 102

	Eye injected in Study 101	Eye injected in Study 102
Baseline, mean (SE)	−6.71 (3.79)	1.40 (2.79)
Year 1, mean (SE)	−10.70 (3.50)	−16.63 (4.82)
Year 2, mean (SE)	−7.67 (4.36)	−14.54 (4.51)
Year 3, mean (SE)	−9.99 (3.13)	−15.75 (4.36)
Year 4, mean (SE)	−7.94 (3.50)	−13.09 (4.17)

SE = standard error.

Source: Clinical Study Report for Study 101[66] and Study 102.[67]

Visual Acuity

In Study 101 at year 1, for the eye injected in Study 101, 7 of 12 (58%) evaluated patients had a change in LogMAR score of 0.3 or more in comparison with baseline (corresponding to an improvement of at least 3 lines or 15 letters on the eye chart). At year 2, 4 of 11 (36%) evaluated patients had a change in LogMAR score of 0.3 or more in comparison with baseline. At year 3, 5 of 9 (56%) evaluated patients had a change in LogMAR score of 0.3 or more in comparison with baseline.

In Study 101, for the eye injected in Study 101, when using the Holladay scale for off-chart results for VA, the mean at baseline was 1.34 (SE = 0.17), and at year 4 it was 1.57 (SE = 0.21). When using the Lange scale for off-chart results for VA, the mean at baseline was 1.31 (SE = 0.15), and at year 4 it was 1.51 (SE = 0.17) (Table 17).

In Study 102, for the LogMAR scores based on the Holladay scale for off-chart results, 1 (9%) of the 11 patients assessed had a change in LogMAR score of 0.3 or more 1 year after receiving treatment in comparison with baseline.

In Study 102, for the eye injected in Study 102, when using the Holladay scale for off-chart results for VA, the mean at baseline was 1.21 (SE = 0.21), and at year 4 it was 1.42 (SE = 0.27). When using the Lange scale for off-chart results for VA, the mean at baseline was 1.14 (SE = 0.16), and at year 4 it was 1.29 (SE = 0.20). (Table 17).

Table 18: Visual Acuity Results for Studies 101 and 102

	Eye injected in Study 101 mean (SE)	Eye injected in Study 102 mean (SE)
Using scales adapted from Holladay 2004[50] for off-chart LogMAR assignments		
Baseline	1.34 (0.17)	1.21 (0.21)
Year 1	1.33 (0.17)	1.21 (0.18)
Year 2	1.44 (0.15)	1.22 (0.18)
Year 3	1.43 (0.16)	1.25 (0.18)
Year 4	1.57 (0.21)	1.42 (0.27)
Using scales adapted from Lange 2009[51] for off-chart LogMAR assignments		
Baseline	1.31 (0.15)	1.14 (0.16)
Year 1	1.30 (0.15)	1.18 (0.16)
Year 2	1.41 (0.13)	1.19 (0.16)
Year 3	1.40 (0.15)	1.22 (0.16)
Year 4	1.51 (0.17)	1.29 (0.20)

LogMAR = logarithm of the minimum angle of resolution; SE = standard error.

Source: Clinical Study Report for Study 101[66] and Study 102.[67]

Harms

Study 101

Adverse event results reported in Table 19 for Study 101 include at least 1 year of post-injection follow-up.

All patients experienced at least 1 TEAE. The most frequently reported TEAEs were conjunctival hyperemia (67%), pyrexia (58%), leukocytosis (50%), and abdominal discomfort and headache (42% each). Persistence or recurrence of conjunctival hyperemia was seen in 8 (67%) patients and classified as a TEAE. These events included findings of

surface irritation (eye), suture reaction, suture irritation, and/or suture allergy and were, in some cases, attributed to the use and persistence of slow-absorbing suture material at the incision site. Symptoms of foreign-body sensation were managed with topical steroids and antibiotic drops, standard post-operative care for subretinal surgery; in 1 patient, the discomfort was addressed by removal of the suture and topical treatment with antibiotic ointment (erythromycin).

Overall, 10 (83%) patients had at least 1 TEAE considered to be related to the study drug administration procedure; all of these TEAEs were considered mild in intensity. The TEAE most often considered to be probably related to the administration procedure was conjunctival hyperemia (67%). Additional adverse events considered related to the administration procedure were cataract, eye disorder, macular hole retinal tear, and endotracheal intubation complication, each of which was reported in 1 patient (8%).

One SAE was reported in Study 101; the SAE was considered unlikely to be related to the study drug or study drug administration procedure.

There were no TEAEs leading to study withdrawal and no deaths reported Study 101.

One patient developed an asymptomatic macular hole approximately 2 weeks after surgery. The event was mild in severity and considered probably related to the administration procedure. One patient experienced a cataract 1,442 days after surgery. This adverse event was considered probably related to the administration procedure. One patient experienced a peripheral retinal tear at the time of surgical visualization of the retina and prior to vector administration; this adverse event was considered probably related to the administration procedure.

Study 102

The adverse event results reported in Table 19 for Study 102 include at least 2 years of post-injection follow-up.

All patients experienced at least 1 TEAE. The most frequently reported TEAEs were pyrexia, influenza, increased blood creatinine, headache, hematuria, and proteinuria (36% for each) followed by cataract, dellen, abdominal discomfort, nausea, vomiting, and oropharyngeal pain (27% for each).

Overall, 7 (64%) patients had at least 1 TEAE considered to be related to the administration procedure; all of these TEAEs were considered mild or moderate in severity. The TEAEs most often considered to be probably related to the administration procedure were dellen (27%), cataract, and increased intraocular pressure (2% for each). Additional adverse events considered related to the administration procedure were eye inflammation, eye irritation, eye pain, maculopathy, and headache, which were each reported in 1 patient (9%).

One SAE was reported in this study, which was elevated intraocular pressure of grade 4 in the right eye resulting in hospitalization. The SAE was deemed related to the use of a depo-steroid injection for a known rare complication of vitrectomy (endophthalmitis). The event was poorly controlled with intraocular pressure–lowering drugs and eventually required filtration surgery, which restored intraocular pressure to normal. During the period of poorly controlled elevated intraocular pressure, the patient suffered optic nerve damage (optic atrophy) that did not reverse.

No deaths or discontinuations from the study due to adverse events were reported.

No inflammation or retinal tears were reported during or immediately after the surgery in this study.

Table 19: Summary of Harms for Studies 101 and 102

	Study 101 (N = 12)	Study 102 (N = 11)
Patients with ≥ 1 AE		
n (%)	12 (100)	11 (100)
Most common events[a]		
Influenza	4 (33)	4 (36)
Nasopharyngitis	4 (33)	2 (18)
Ear infection	3 (25)	1 (9.1)
Tracheitis	2 (17)	NR
Conjunctival hyperemia	8 (67)	NR
Pyrexia	7 (58)	4 (36)
Leukocytosis	6 (50)	1 (9.1)
Contusion	3 (25)	2 (18)
Corneal abrasion	2 (17)	NR
Fall	2 (17)	NR
Hypoglycemia	4 (33)	2 (18)
Hyperglycemia	3 (25)	2 (18)
Abdominal discomfort	5 (42)	3 (27)
Headache	5 (42)	4 (36)
Blood creatinine increased	4 (33)	4 (36)
Hematuria	4 (33)	4 (36)
Proteinuria	2 (17)	4 (36)
Cough	4 (33)	2 (18)
Acne	2 (17)	NR
Neck pain	2 (17)	NR
Abdominal pain	NR	2 (18)
Diarrhea	NR	2 (18)
Gastroesophageal reflux disease	NR	2 (18)
Nausea	1 (8)	3 (27)
Vomiting	NR	3 (27)
Cataract	1 (8)	3 (27)
Dellen	NR	3 (27)
Intraocular pressure increased	1 (8)	2 (18)
Excoriation	NR	2 (18)
Oropharyngeal pain	NR	3 (27)
Seasonal allergy	NR	2 (18)
Patients with ≥ 1 SAE		
n (%)	1 (8)	1 (9)
Most common events		
Anal fistula	1 (8)	0
Intraocular pressure	0	1 (9)

	Study 101 (N = 12)	Study 102 (N = 11)
Patients who stopped treatment due to AEs		
n (%)	0	0
Deaths		
n (%)	0	0

AE = adverse event; NR = not reported; SAE = serious adverse event.

[a] AEs reported in more than 1 patient.

Source: Clinical Study Reports for Study 101[66] and Study 102.[67]

Critical Appraisal

Internal Validity

Studies 101 and 102 are phase I, single-arm, open-label, non-randomized studies with a small sample size and are not considered to provide high-quality evidence to support the efficacy benefit of voretigene neparvovec.

No conclusions can be drawn regarding the clinical efficacy of voretigene neparvovec. Study 101 was a dose-escalating study and neither Study 101 nor Study 102 were designed or powered to assess the clinical efficacy of voretigene neparvovec. No formal hypothesis testing was conducted, and the within-group changes were also not designed to be inferential. Only descriptive statistics were presented.

Study 101 and Study 102 had small sample sizes, which would lead to large standard error and high variability. In addition, it was not reported how many patients were included in the calculation of the mean values.

During the course of Study 101, the mobility test was further refined and standardized, which affected both the number of patients considered evaluable and the interpretation of the results.

External Validity

The dosage of voretigene neparvovec used in 9 of the 12 patients enrolled in Study 101 was not the dosage approved by Health Canada; hence, the generalizability of the study results to the Canadian patient population is unclear, especially since the majority of patients in Study 101 received a dose that is lower than the dosage approved by Health Canada and, therefore, the treatment effect might have been underestimated.

Study 101 included 12 patients; differences in 1 or 2 patients can have a substantial impact on results and the ability to capture rare adverse events.

Compared with Study 101, Study 301 included a broader population, including younger patients (4 years of age and older in Study 301 versus 8 years of age and older in Study 101), and patients with less advanced disease (VA no better than 20/60 in Study 301 versus VA no better than 20/160 in Study 101). The more stringent criteria introduced in Study 102 for determining the number of viable retinal cells (≥ 3 disc areas of retina without atrophy or pigmentary degeneration within the posterior pole) was also used in Study 301, whereas in Study 101, patients were eligible if they had 1 or more disc areas of retina that were not involved in complete retinal degeneration. Finally, periocular injection of the various corticosteroids used in Study 101 and Study 102 was discontinued in Study 301 to decrease the incidence and severity of elevated intraocular pressure and cataract formation and/or progression.

Discussion

Summary of Available Evidence

One phase III, open-label, randomized controlled trial designed to evaluate the efficacy and safety of sequential subretinal injection of voretigene neparvovec into each eye met the inclusion criteria for the review. Study 301 enrolled a total of 31 patients diagnosed with LCA due to *RPE65* mutations in 2 study sites in the US (the study enrolled international patients, including 1 patient from Canada). Patients were randomized in a 2:1 ratio to either the voretigene neparvovec group (n = 21) or the control group (n = 10). Patients randomized to the voretigene neparvovec group received a dose of 1.5×10^{11} vg of voretigene neparvovec in each eye; these non-simultaneous, subretinal injections were to occur within an 18-day period (12 days ± 6 days). Patients randomized to the control group did not receive voretigene neparvovec, sham injection, or corticosteroids for a period of at least 1 year from baseline evaluations. Patients in the control group were crossed over to receive non-simultaneous injections of 1.5×10^{11} vg of voretigene neparvovec in each eye (within 18 days) after 1 year of randomization, provided they still met all study eligibility criteria. The primary end point was change in bilateral MLMT performance at year 1 relative to baseline. Secondary end points were change in white light FST averaged over both eyes at year 1 relative to baseline, change in assigned first eye MLMT performance at year 1 relative to baseline, and change in VA averaged over both eyes at year 1 relative to baseline. At the time of the data cut-off for the CSR provided by the sponsor, all 20 patients (100%) in the voretigene neparvovec treatment group had completed the 4 years after the second-eye injection study visit and 8 patients (89%) in the control/voretigene neparvovec treatment group had completed the 3 years after the second-eye injection study visit. In addition, 4 patients (20%) in the voretigene neparvovec treatment group had completed the 5 years after the second-eye injection study visit.

Study 101 and Study 102, which were phase I studies, were reviewed and critically appraised, and results are reported in the Other Relevant Evidence section of this report. Study 101 was a dose-escalation study that assessed the safety and tolerability of 3 different doses of voretigene neparvovec administered via subretinal administration in 1 eye (first-treated eye) of patients with LCA due to *RPE65* mutations. Study 101 also evaluated the clinical efficacy, but no formal hypothesis testing was conducted. Study 102 was a follow-on study of Study 101 in which patients received voretigene neparvovec treatment in the previously uninjected eye (second-treated eye).

Interpretation of Results

Efficacy

One meaningful aspect of health addressed by the MLMT is “functional vision” or how the individual uses vision to navigate around obstacles and from place to place. In patients with *RPE65* mutation–associated retinal dystrophy, nyctalopia is a hallmark of the disease, meaning that vision, particularly in dim light, is profoundly impaired. This limits or prevents the ability to perform multiple activities that are part of normal life, particularly those that take place in low-illuminance environments. The patient group input received for this submission indicated that this impairment is meaningful to the patients. In Study 301, treatment with voretigene neparvovec resulted in a statistically significant improvement in navigational ability in low-to-moderate light conditions (measured by MLMT score) 1 year

after non-simultaneous treatment of both eyes, where 11 patients (52%) in the voretigene neparvovec group had an MLMT score change of 2 or more (the difference considered meaningful by the FDA and EMA). In contrast, only 1 patient (10%) in the control group had a score change of 2, and none of the patients in the control group had a score change greater than 2. The difference in change from baseline in bilateral MLMT between the voretigene neparvovec and control treatment groups at 1 year was 1.6 (95% CI, 0.72 to 2.41; P = 0.001), which was statistically significant in favour of voretigene neparvovec; however, the difference between the treatment groups did not exceed the 2 points proposed by the regulatory agencies in the US. However, MLMT scores may underestimate the treatment effect of voretigene neparvovec due to the potential ceiling effect, where patients who passed the test at the second-lowest light level at baseline were only able to achieve a maximum 1 unit increase. Although 62% of patients in the voretigene neparvovec group achieved a score of 6 on the MLMT (the maximum possible score) following administration of voretigene neparvovec, compared with no patients in the control group achieving a score of 6 on the MLMT, the observed mean increase of 1.8 in MLMT score observed in the voretigene neparvovec group could be an underestimate of the within-groups magnitude of the change because of this ceiling effect. Improvements in the MLMT score observed at 1 year seemed to be maintained until the 4-year follow-up.

Based on the patient group input received for this submission, patients consider improved visual function to be an important outcome of treatment. The visual function outcomes assessed in Study 301 were FST, VA, and VF.

The FST results were transformed from dB, the original relative unit of measurement, to $\log_{10}(\text{cd.s/m}^2)$, in which cd.s/m^2 is the absolute unit of measurement; this conversion was necessary to compare testing performed with different dB conversion rates. For this analysis, smaller dB and smaller cd.s/m^2 values both indicate better sensitivity. $\text{Log}_{10}(\text{cd.s/m}^2)$ values indicate better sensitivity the more negative they are. The FST was a secondary outcome in Study 301 and it was adjusted for multiple testing. Patients treated with voretigene neparvovec experienced a mean improvement in FST of greater than 2 log units after 1 year, whereas mean FST did not change in the control arm; this was a statistically significant difference in favour of voretigene neparvovec, with a mean between-group treatment difference of −2.11 (95% CI, −3.19 to −1.04; P < 0.001) $\log_{10}(\text{cd.s/m}^2)$. While this between-group difference exceeded the estimated defined threshold of 10 dB or 1 log unit for a clinically meaningful difference, the methodology used to estimate this MID was unclear. Long-term data suggest the improvements were sustained for 4 years after the second-eye injection. The clinical experts explained that the changes seen would be substantial in terms of improving visual function.

VA results are presented in LogMAR units, where smaller values indicate better acuity (less VA loss). The off-chart VA measurements used adaptations of previously reported scales for assigning LogMAR (Holladay et al.[50] and Lange et al.[51]). For the VA analyses, a 0.1 improvement in LogMAR corresponded to a 5-letter improvement (or the equivalent of 1 line) on a standard eye chart. A longitudinal repeated measures analysis of VA using the Holladay scale averaged over both eyes from baseline to 1 year resulted in a between-groups difference that was neither statistically significant nor clinically meaningful. Using the Lange scale for off-chart VA results, the mean between-group treatment difference was −0.16 LogMAR (95% CI, −0.31 to −0.01; nominal P = 0.035) corresponding to a 7.5-letter improvement on the eye chart for people who had received voretigene neparvovec. While the difference between treatment groups using the Lange scale for off-chart VA seems to favour voretigene neparvovec, it is worth noting that this difference did not exceed the

estimated MID of 0.30 LogMAR or more. In addition, this analysis was not adjusted for multiple comparisons; thus, the level of significance is inflated and this should be considered when interpreting the results. By year 3, little further change was seen in VA for either arm after treatment. The clinical experts explained that even a small change in VA would be important for patients. The clinical experts also noted that, even if there were no improvement, preventing vision deterioration would be important for the patient's quality of life.

VF was tested in Study 301 as an exploratory end point. VF measures the peripheral retinal function while the eye is focused on a central point. Goldmann perimetry (assessing the full extent of the VF for each eye; frequently used in low-vision patients and those with nystagmus), as well as Humphrey computerized testing (evaluating the sensitivity of specific points in the central retina [macula and fovea]), were tested. Compared with baseline, VF patients who received voretigene neparvovec had higher Goldmann VF III4e and Humphrey VF macula scores at 1 year than patients in the control group. The between-group difference reported for the Humphrey VF macula threshold analysis exceeded the difference of 7 dB that is considered clinically meaningful by the FDA. Given that VF was assessed as an exploratory end point and was not adjusted for multiple statistical testing, no statistical interpretations should be made.

The clinical experts indicated that the VA and VF outcome measures used in the trial are highly specialized, requiring special centres to administer and generally not used in clinical practice. Also, it is worth noting that the NICE guidance for voretigene neparvovec indicated that VA and VF are often considered unreliable because of inter-test variability.[53]

It was also clear from the patient group input received for this submission that patients consider improved quality of life to be an important outcome of treatment. Although the VFQ indicated improvements in patients' ability to perform ADL, it did not contain any items to specifically assess HRQoL for patients. While both the 25- and 51-item versions of the VFQ have been extensively validated in various ophthalmic conditions and in a variety of age groups,[55-58] the adapted version of the VFQ used in Study 301 was not assessed psychometrically. Also, given the modifications made to the original VFQ, the MIDs identified in the literature for that measure were not considered directly generalizable to the version used in Study 301. In addition, the VFQ was susceptible to bias and intrinsic subjectivity due to the open-label design of the study. Hence, the potential effect of voretigene neparvovec on patients' HRQoL remains unknown. The lack of patient-reported outcomes was identified as a key limitation by the NICE guidance for voretigene neparvovec.[53]

Subgroup analyses of interest to this review were based on age, clinical phenotype, and number of viable retinal cells. No preplanned subgroup analyses were conducted in Study 301. The sponsor conducted a post hoc subgroup analysis by age (< 18 years at first injection versus ≥ 18 years at first injection); however, this analysis should be considered hypothesis-generating. Also, it is not clear if this subgroup analysis was conducted at the request of a regulator, and no rationale was provided for the age cut-off. The clinical experts consulted by CADTH indicated that greater clinical benefit from treatment is expected earlier in the condition when there are more viable retinal cells for the gene replacement therapy to restore. While patient age can be used as a proxy to estimate how advanced the condition is, due to the heterogeneous nature of *RPE65* mutation–associated retinal dystrophy and the differences in the age of onset and disease progression, retinal cell viability should be assessed on an individual patient basis, regardless of the patient's

age (as long as the patient is at least 4 years old), to determine if the specific individual is likely to respond to voretigene neparvovec.

For Study 301, long-term data for patients in the voretigene neparvovec group suggest durable improvements in visual performance across multiple end points for at least 4 years following voretigene neparvovec administration. Similarly, after crossing over to voretigene neparvovec, patients in the control/voretigene neparvovec group exhibited durability in visual performance improvements comparable with those observed in the voretigene neparvovec group, effects that were maintained for at least 3 years following bilateral voretigene neparvovec administration. Study 301 aims to follow patients for up to 15 years after treatment. The current duration of follow-up is limited to a maximum of 5 years, and longer-term efficacy and safety data for voretigene neparvovec are awaited. The duration of treatment effect is unclear and there is no information on whether patients who lose treatment effect would benefit from re-treatment. Also, given the small sample size (30 patients), there is uncertainty around the generalizability of the results observed for the long-term treatment effect. The clinical experts indicated that voretigene neparvovec would likely provide long-term benefits, although this was associated with substantial uncertainty.

Individuals with pathogenic *RPE65* mutations have significant retinal degeneration leading to worse functional vision over time.[68] Currently, it is unknown whether voretigene neparvovec has the potential to reduce or eliminate retinal degeneration. One challenge to assessing ongoing degeneration is the pace of deterioration in visual outcomes, where it could take up to a decade to observe worsening in visual outcome measures in this population.[69] Also, deterioration in vision tends to occur with age, and thus natural deterioration should not be confused with a waning of treatment effect.

Study 101 and Study 102 suggest there is sustained improvement in vision for up to 4 years; however, there is considerable uncertainty associated with the results of these studies. In addition to their evidence being of low quality, very few patients in Study 101 received 1.5×10^{11} vg of voretigene neparvovec, the dose under review (most patients received smaller doses). Also, in comparison with Study 101, Study 301 included a broader population, including younger patients (4 years of age and older in Study 301 versus 8 years of age and older in Study 101), and included patients with less advanced disease (VA no better than 20/60 in Study 301 versus VA no better than 20/160 in Study 101). Also, more stringent criteria were introduced for determining the number of viable retinal cells in Study 301. In Study 101, patients were eligible if they had 1 or more disc areas of retina that were not involved in complete retinal degeneration but, in Study 301, eligible patients were required to have 3 or more disc areas of retina without atrophy or pigmentary degeneration.

Harms

All patients in Study 301, Study 101, and Study 102 experienced at least 1 TEAE. Most adverse events were mild in severity and no patient had adverse events that led to study discontinuation or death. The most common adverse events were ocular events related to the subretinal injection of voretigene neparvovec and the concomitant use of systemic corticosteroids. These adverse events included conjunctival hyperemia, increased intraocular pressure, cataract, retinal abnormalities (retinal tear, macular hole, macular pucker, foveal thinning, retinal bleeding, foveal dehiscence), endophthalmitis, and loss of vision. Most of these events were temporary and responded to medical management. There were ongoing adverse events, including maculopathy, cataracts, and increased intraocular pressure. The most common ocular adverse reactions (incidence ≥ 5%) related

to the administration procedure were conjunctival hyperemia, cataract, increased intraocular pressure, retinal tear, dellen (areas of thinning of the corneal stroma), macular hole, subretinal deposits, eye inflammation, eye irritation, eye pain, and maculopathy (wrinkling on the surface of the macula). Two serious ocular adverse events with severe consequences were associated with the surgical procedure. One of these events occurred in the phase I trial (Study 101); this SAE was persistent elevated intraocular pressure which was attributed to the local administration of steroids used to treat inflammation. The patient suffered optic nerve damage (optic atrophy) that did not reverse. Periocular injection of the various corticosteroids used in the phase I trial were discontinued in Study 301 to decrease the incidence and severity of elevated intraocular pressure and cataract formation and/or progression. The other ocular SAE occurred in Study 301, where a patient who received voretigene neparvovec experienced a retinal disorder which was foveal thinning and loss of central vision that was related to the subretinal injection in this patient with pre-existing atrophy of the retina.

As described, the administration procedure for voretigene neparvovec may be associated with potential safety concerns. The product monograph indicates that voretigene neparvovec should be administered in the surgical suite under controlled aseptic conditions. According to the product monograph, voretigene neparvovec should be initiated and administered by a retinal surgeon experienced in performing subretinal surgery, but the clinical experts consulted for this review indicated that surgeons should also have experience with sub-macular injection and management of its complications.

Conclusions

Currently, there is no pharmacologic treatment available for the treatment of patients with vision loss due to IRD caused by confirmed biallelic *RPE65* mutations; the current standard of care is supportive in nature and focuses on monitoring, psychological support, mobility training, and visual rehabilitation to maintain the patients' ability to perform ADL and improve HRQoL. Based on the results of 1 phase III study (Study 301), voretigene neparvovec, compared with the control group, demonstrated a statistically significant improvement in functional vision under dim light conditions, as measured by the MLMT one year post treatment. Voretigene neparvovec also resulted in a statistically significant improvement in FST one year post treatment. No improvement was observed in VA and the effect of voretigene neparvovec on HRQoL is unknown. The improvements observed with voretigene neparvovec after one year appear to be maintained for up to 4 years; however, further data are required and there is uncertainty about the duration of treatment effect. Harms, although present, were related to the surgical aspects of administration; the most common ocular adverse events related to the administration procedure were conjunctival hyperemia, cataract, increased intraocular pressure, retinal tear, dellen, macular hole, subretinal deposits, eye inflammation, eye irritation, eye pain, and maculopathy; 2 ocular SAEs had severe consequences.

Appendix 1: Literature Search Strategy

Clinical Literature Search

OVERVIEW	
Interface:	Ovid
Databases:	MEDLINE All (1946 to present) Embase (1974 to present) **Note:** Subject headings have been customized for each database. Duplicates between databases were removed in Ovid.
Date of Search:	May 22, 2020
Alerts:	Bi-weekly search updates until project completion
Study Types:	No search filters were applied
Limits:	No date or language limits were used Conference abstracts: excluded

SYNTAX GUIDE	
/	At the end of a phrase, searches the phrase as a subject heading
*	Before a word, indicates that the marked subject heading is a primary topic; or, after a word, a truncation symbol (wildcard) to retrieve plurals or varying endings
MeSH	Medical Subject Heading
exp	Explode a subject heading
.ti	Title
.ab	Abstract
.dq	Candidate term word (Embase)
.ot	Original title
adj#	Requires terms to be adjacent to each other within # number of words (in any order)
.hw	Heading word; usually includes subject headings and controlled vocabulary
.kf	Author keyword heading word (MEDLINE)
.kw	Author keyword (Embase)
.pt	Publication type
.mp	Mapped term
.rn	Registry number
.yr	Publication year
medall	Ovid database code: MEDLINE All, 1946 to present, updated daily
oemezd	Ovid database code; Embase, 1974 to present, updated daily

MULTI-DATABASE STRATEGY	
Line #	**Search Strategy**
1	(voretigene* or luxturna* or aav2 hrpe65v2 or aav2hrpe65v2 or ltw 888 or ltw888 or spk rpe65 or VN-rzyl or 2SPI046IKD).ti,ab,kf,ot,hw,rn,nm.
2	1 use medall
3	*voretigene neparvovec/
4	(voretigene* or luxturna* or aav2 hrpe65v2 or aav2hrpe65v2 or ltw 888 or ltw888 or spk rpe65 or VN-rzyl).ti,ab,kw,dq.
5	3 or 4
6	5 use oemezd
7	6 not (conference review or conference abstract).pt.
8	2 or 7
9	remove duplicates from 8

CLINICAL TRIAL REGISTRIES	
ClinicalTrials.gov	Produced by the US National Library of Medicine. Targeted search used to capture registered clinical trials. Search updated prior to the completion of stakeholder feedback period. Search terms: (voretigene OR luxturna OR aav2 hrpe65v2 OR ltw 888 OR rpe65)

OTHER DATABASES	
PubMed	Searched to capture records not found in MEDLINE. Same MeSH, keywords, limits, and study types used as per MEDLINE search, with appropriate syntax used.

Grey Literature

Search dates:	May 12 to 19, 2020
Keywords:	voretigene or luxturna or rpe65
Limits:	None
Updated:	Search updated prior to the completion of stakeholder feedback period

Relevant websites from the following sections of the CADTH grey literature checklist *Grey Matters: A Practical Tool For Searching Health-Related Grey Literature* (www.cadth.ca/grey-matters) were searched:

- Health Technology Assessment Agencies
- Health Economics
- Clinical Practice Guidelines
- Drug and Device Regulatory Approvals
- Advisories and Warnings
- Drug Class Reviews
- Clinical Trial Registries
- Databases (Free)
- Internet Search

Appendix 2: Excluded Studies

Table 20: Excluded Studies

Reference	Reason for Exclusion
Bennett J, Ashtari M, Wellman J, et al. AAV2 gene therapy readministration in three adults with congenital blindness. Sci Transl Med. 2012;4(120):120ra115. Maguire AM, Simonelli F, Pierce EA, et al. Safety and efficacy of gene transfer for Leber's congenital amaurosis. N Engl J Med. 2008;358(21):2240-2248. Simonelli F, Maguire AM, Testa F, et al. Gene therapy for Leber's congenital amaurosis is safe and effective through 1.5 years after vector administration. Molecular Therapy: the Journal of the American Society of Gene Therapy. 2010;18(3):643-650. Bennett J, Wellman J, Marshall KA, et al. Safety and durability of effect of contralateral-eye administration of AAV2 gene therapy in patients with childhood-onset blindness caused by RPE65 mutations: a follow-on phase 1 trial. Lancet. 2016;388(10045):661-672. Maguire AM, High KA, Auricchio A, et al. Age-dependent effects of RPE65 gene therapy for Leber's congenital amaurosis: a phase 1 dose-escalation trial. Lancet. 2009;374(9701):1597-1605. Testa F, Maguire AM, Rossi S, et al. Three-year follow-up after unilateral subretinal delivery of adeno-associated virus in patients with Leber congenital amaurosis type 2. Ophthalmology. 2013;120(6):1283-1291.	Study design (phase I randomized controlled trial)

Appendix 3: Detailed Outcome Data

Table 21: Bilateral Multi-Luminance Mobility Test Change Score Post Injections in Study 301 (mITT Population)

	Voretigene neparvovec n = 20	Control/voretigene neparvovec n = 9	Total N = 29
Change score 1 year after the second-eye injection			
n	20	9	29
Mean (SD) compared with injection baseline	1.9 (1.0)	2.1 (1.6)	1.9 (1.2)
Change score 2 years after the second-eye injection			
n	20	9	29
Mean (SD) compared with injection baseline	1.9 (1.1)	2.1 (1.6)	1.9 (1.3)
Change score 3 years after the second-eye injection			
n	20	8	28
Mean (SD) compared with injection baseline	1.8 (1.0)	2.4 (1.5)	2.0 (1.1)
Change score 4 years after the second-eye injection			
n	[redacted]	[redacted]	[redacted]
Mean (SD) compared with injection baseline	[redacted]	[redacted]	[redacted]
Change score 5 years after the second-eye injection			
n	[redacted]	[redacted]	[redacted]
Mean (SD) compared with injection baseline	[redacted]	[redacted]	[redacted]

mITT = modified intention to treat; SD = standard deviation.

Source: Clinical Study Report for Study 301.[15]

Table 22: The Full-Field Light Sensitivity Threshold Post Injections in Study 301

	Full-field light sensitivity threshold: white light both eyes (log_{10}[cd.s/m^2]), observed estimates (mITT)[a]		
	Voretigene neparvovec n = 20	Control/voretigene neparvovec[b] n = 9	Total N = 29
Injection baseline			
n	19	9	28
Mean (SD)	−1.32 (0.44)	−1.61 (0.45)	−1.41 (0.46)
1 year after the second-eye injection			
n	20	9	29
Mean (SD)	−3.37 (1.48)	−4.47 (1.45)	−3.71 (1.54)
Change from baseline, mean (SD)	−2.10 (1.58)	−2.86 (1.49)	−2.34 (1.57)
2 years after the second-eye injection			
n	20	9	29
Mean (SD)	−3.51 (1.60)	−4.31 (1.31)	−3.76 (1.54)
Change from baseline, mean (SD)	−2.27 (1.65)	−2.69 (1.41)	−2.40 (1.57)
3 years after the second-eye injection			
n	20	8	28
Mean (SD)	−3.30 (1.33)	−4.50 (1.20)	−3.65 (1.38)

	Full-field light sensitivity threshold: white light both eyes (log_{10}[cd.s/m^2]), observed estimates (mITT)[a]		
	Voretigene neparvovec n = 20	Control/voretigene neparvovec[b] n = 9	Total N = 29
Change from baseline, mean (SD)	−2.04 (1.43)	−2.91 (1.05)	−2.30 (1.37)
4 years after the second-eye injection			
n	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]
5 years after the second-eye injection			
n	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]

cd.s/m^2 = candela second per square metre; mITT = modified intention to treat; SD = standard deviation; vg = vector genome.

[a] All measures were averaged over both eyes and then analyzed.

[b] Patients in the control group became eligible to receive voretigene neparvovec at a dose of 1.5 × 10^{11} vg in a total subretinal volume of 300 µL in each eye (within 18 days) 1 year after their baseline evaluations, provided they still met all eligibility criteria.

Source: Clinical Study Report for Study 301.[15]

Table 23: Visual Acuity Post Injections in Study 301

	Voretigene neparvovec n = 20	Control/voretigene neparvovec[a] n = 9	Total N = 29
Visual acuity (LogMAR) using Holladay off-chart, both eyes, observed estimates (mITT)[b]			
Injection baseline			
n	20	9	29
Mean (SD)	1.14 (0.37)	0.95 (0.33)	1.08 (0.36)
1 year after the second-eye injection			
n	20	9	29
Mean (SD)	0.97 (0.54)	0.87 (0.26)	0.94 (0.47)
Change from baseline, mean (SD)	−0.16 (0.34)	−0.09 (0.22)	−0.14 (0.30)
2 years after the second-eye injection			
n	20	9	29
Mean (SD)	0.98 (0.55)	0.89 (0.27)	0.95 (0.47)
Change from baseline, mean (SD)	−0.16 (0.36)	−0.06 (0.23)	−0.13 (0.32)
3 years after the second-eye injection			
n	20	8	28
Mean (SD)	0.98 (0.59)	0.85 (0.25)	0.94 (0.52)
Change from baseline, mean (SD)	−0.16 (0.35)	−0.06 (0.24)	−0.13 (0.32)
4 years after the second-eye injection			
n	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]
5 years after the second-eye injection			

	Voretigene neparvovec n = 20	Control/voretigene neparvovec[a] n = 9	Total N = 29
n	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]
Visual acuity (LogMAR) using Lange off-chart, both eyes, observed estimates (mITT)[a]			
Injection baseline			
n	20	9	29
Mean (SD)	1.12 (0.33)	0.95 (0.33)	1.07 (0.33)
1 year after the second-eye injection			
n	20	9	29
Mean (SD)	0.94 (0.42)	0.87 (0.26)	0.92 (0.37)
Change from baseline, mean (SD)	−0.18 (0.20)	−0.09 (0.22)	−0.15 (0.21)
2 years after the second-eye injection			
n	20	9	29
Mean (SD)	0.94 (0.42)	0.89 (0.27)	0.93 (0.38)
Change from baseline, mean (SD)	−0.17 (0.23)	−0.06 (0.23)	−0.14 (0.23)
3 years after the second-eye injection			
n	20	8	28
Mean (SD)	0.93 (0.44)	0.85 (0.25)	0.90 (0.39)
Change from baseline, mean (SD)	−0.19 (0.23)	−0.06 (0.24)	−0.16 (0.24)
4 years after the second-eye injection			
n	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]
5 years after the second-eye injection			
n	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]

LogMAR = logarithm of the minimum angle of resolution; mITT = modified intention to treat; SD = standard deviation; vg = vector genome.

[a] Patients in the control group became eligible to receive voretigene neparvovec at a dose of 1.5×10^{11} vg in a total subretinal volume of 300 µL in each eye (within 18 days) 1 year after their baseline evaluations, provided they still met all eligibility criteria.

[b] All measures were averaged over both eyes and then analyzed.

Source: Clinical Study Report for Study 301.[15]

Table 24: Goldmann Perimetry and Humphrey Computerized Testing Post Injections in Study 301

	Voretigene neparvovec n = 20	Control/voretigene neparvovec[a] n = 9	Total N = 29
Goldmann visual field V4e (sum total degrees) (mITT)[b]			
Injection baseline			
n	19	5	24
Mean (SD)	921.7 (477.6)	778.8 (301.6)	892.0 (444.8)
1 year after the second-eye injection			
n	11	1	12
Mean (SD)	1,032.8 (592.2)	442.0 (NA)	983.6 (589.8)
Change from baseline, mean (SD)	78.8 (156.9)	−37.0 (NA)	68.3 (152.9)
2 years after the second-eye injection			
n	11	1	12
Mean (SD)	1,006.8 (598.5)	458.5 (NA)	961.1 (592.2)
Change from baseline, mean (SD)	63.8 (153.8)	−20.5 (NA)	56.1 (148.1)
3 years after the second-eye injection			
n	11	0	11
Mean (SD)	884.8 (611.2)	NA	884.8 (611.2)
Change from baseline, mean (SD)	28.7 (130.0)	NA	28.7 (130.0)
4 years after the second-eye injection			
n			
Mean (SD)			
Change from baseline, mean (SD)			
Goldmann visual field III4e (sum total degrees) (mITT)[b]			
Injection baseline			
n	19	9	28
Mean (SD)	350.4 (416.9)	397.8 (367.3)	365.6 (395.4)
1 year after the second-eye injection			
n	20	9	28
Mean (SD)	673.9 (423.7)	592.1 (296.6)	648.5 (385.2)
Change from baseline, mean (SD)	302.1 (289.6)	194.3 (244.7)	267.4 (276.2)
2 years after the second-eye injection			
n	19	9	28
Mean (SD)	697.5 (389.7)	580.4 (276.7)	659.9 (356.4)
Change from baseline, mean (SD)	311.6 (295.3)	182.6 (309.9)	268.6 (300.7)
3 years after the second-eye injection			
n	19	2	21
Mean (SD)	625.9 (413.3)	687.8 (222.4)	631.8 (395.7)

	Voretigene neparvovec n = 20	Control/voretigene neparvovec[a] n = 9	Total N = 29
Change from baseline, mean (SD)	282.2 (256.5)	117.3 (91.6)	265.7 (248.8)
4 years after the second-eye injection			
n	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]
Humphrey visual field foveal sensitivity (dB) (mITT)[b]			
Injection baseline			
n	19	9	28
Mean (SD)	23.34 (5.49)	21.50 (8.91)	22.75 (6.66)
1 year after the second-eye injection			
n	20	9	29
Mean (SD)	25.83 (9.07)	24.72 (9.41)	25.48 (9.02)
Change from baseline, mean (SD)	2.37 (9.68)	3.22 (11.49)	2.64 (10.09)
2 years after the second-eye injection			
n	18	9	27
Mean (SD)	25.94 (8.52)	26.50 (4.83)	26.13 (7.40)
Change from baseline, mean (SD)	3.08 (8.48)	5.00 (8.30)	3.72 (8.31)
3 years after the second-eye injection			
n	20	2	22
Mean (SD)	26.60 (8.08)	25.25 (7.42)	26.48 (7.86)
Change from baseline, mean (SD)	3.03 (8.71)	6.00 (3.54)	3.31 (8.35)
4 years after the second-eye injection			
n	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]
Humphrey visual field macula threshold (dB) (mITT)[b]			
Injection baseline			
n	19	9	28
Mean (SD)	16.55 (5.31)	15.81 (7.37)	16.31 (5.92)
1 year after the second-eye injection			
n	20	9	29
Mean (SD)	23.98 (7.97)	21.04 (11.93)	23.07 (9.25)
Change from baseline, mean (SD)	7.66 (6.23)	5.23 (9.92)	6.88 (7.51)
2 years after the second-eye injection			
n	19	9	28
Mean (SD)	22.55 (8.66)	22.87 (9.36)	22.65 (8.72)
Change from baseline, mean (SD)	6.45 (7.35)	7.06 (7.23)	6.65 (7.17)

	Voretigene neparvovec n = 20	Control/voretigene neparvovec[a] n = 9	Total N = 29
3 years after the second-eye injection			
n	20	2	22
Mean (SD)	22.89 (6.90)	16.53 (18.21)	22.31 (7.89)
Change from baseline, mean (SD)	6.50 (5.77)	7.55 (8.13)	6.60 (5.78)
4 years after the second-eye injection			
n			
Mean (SD)			
Change from baseline, mean (SD)			

dB = decibels; mITT = modified intention to treat; NA = not applicable; SD = standard deviation; vg = vector genome.

[a] Patients in the control group became eligible to receive voretigene neparvovec at a dose of 1.5×10^{11} vg in a total subretinal volume of 300 µL in each eye (within 18 days) 1 year after their baseline evaluations, provided they still met all eligibility criteria.

[b] All measures were averaged over both eyes and then analyzed.

Source: Clinical Study Report for Study 301.[15]

Table 25: Visual Function Questionnaire Average Scores Post Injections in Study 301

	Voretigene neparvovec n = 20	Control/voretigene neparvovec[a] n = 9	Total N = 29
Average score (patient) (mITT)			
Injection baseline			
n	20	9	29
Mean (SD)	4.5 (1.4)	5.0 (1.7)	4.6 (1.5)
1 year after the second-eye injection			
n	20	9	29
Mean (SD)	7.0 (1.9)	6.5 (2.0)	6.9 (1.9)
Change from baseline, mean (SD)	2.6 (1.8)	1.5 (1.5)	2.2 (1.8)
2 years after the second-eye injection			
n	20	9	29
Mean (SD)	6.6 (1.9)	5.9 (1.8)	6.4 (1.9)
Change from baseline, mean (SD)	2.2 (1.5)	1.0 (1.5)	1.8 (1.6)
3 years after the second-eye injection			
n	20	2	22
Mean (SD)	6.4 (1.9)	5.6 (0.3)	6.3 (1.8)
Change from baseline, mean (SD)	1.9 (1.8)	2.7 (0.4)	2.0 (1.7)
4 years after the second-eye injection			
n			
Mean (SD)			
Change from baseline, mean (SD)			
Average score (parent) (mITT)			
Injection baseline			
n	15	5	20

	Voretigene neparvovec n = 20	Control/voretigene neparvovec[a] n = 9	Total N = 29
Mean (SD)	3.6 (1.3)	3.1 (1.8)	3.5 (1.4)
1 year after the second-eye injection			
n	15	5	20
Mean (SD)	7.5 (1.5)	6.0 (1.2)	7.1 (1.6)
Change from baseline, mean (SD)	3.9 (1.9)	2.9 (1.0)	3.6 (1.8)
2 years after the second-eye injection			
n	14	5	19
Mean (SD)	7.4 (1.4)	5.9 (1.3)	7.0 (1.5)
Change from baseline, mean (SD)	3.7 (1.9)	2.8 (0.8)	3.5 (1.7)
3 years after the second-eye injection			
n	14	1	15
Mean (SD)	7.1 (1.6)	4.9 (NA)	6.9 (1.6)
Change from baseline, mean (SD)	3.4 (1.9)	3.5 (NA)	3.4 (1.8)
4 years after the second-eye injection			
n	[redacted]	[redacted]	[redacted]
Mean (SD)	[redacted]	[redacted]	[redacted]
Change from baseline, mean (SD)	[redacted]	[redacted]	[redacted]

mITT = modified intention to treat; SD = standard deviation; vector genome.

[a] Patients in the control group became eligible to receive voretigene neparvovec at a dose of 1.5×10^{11} vg in a total subretinal volume of 300 µL in each eye (within 18 days) 1 year after their baseline evaluations, provided they still met all eligibility criteria.

Source: Clinical Study Report for Study 301.[15]

Table 26: Subgroup Efficacy Results by Age for Change From Injection Baseline to 1 Year After the Second-Eye Injection for All Patients Receiving Voretigene Neparvovec in Study 301

	Age at first injection	
	< 18 years n = 20	≥ 18 years n = 9
Mobility testing sum change score[a]		
Baseline, mean (SD)	[redacted]	[redacted]
Mean (SD) 1 year after the second-eye injection	[redacted]	[redacted]
Full-field light sensitivity threshold testing change (log_{10}[cd.s/m^2])[b]		
Baseline, mean (SD)	[redacted]	[redacted]
Mean (SD) 1 year after the second-eye injection	[redacted]	[redacted]
Goldmann visual field III4e change (sum total degrees)		
Baseline, mean (SD)	[redacted]	[redacted]
Mean (SD) 1 year after the second-eye injection	[redacted]	[redacted]
Humphrey visual field macula threshold change (mean dB)		
Baseline, mean (SD)	[redacted]	[redacted]
Mean (SD) 1 year after the second-eye injection	[redacted]	[redacted]

	Age at first injection	
	< 18 years n = 20	≥ 18 years n = 9
Visual acuity change (LogMAR)[c]		
Baseline, mean (SD)	[redacted]	[redacted]
Mean (SD) 1 year after the second-eye injection	[redacted]	[redacted]

cd.s/m^2 = candela second per square metre; dB = decibel; LogMAR = logarithm of the minimum angle of resolution; SD = standard deviation.

[a] Mobility testing results presented for the bilateral testing condition.

[b] Full-field light sensitivity threshold testing results are presented for the average of both eyes.

[c] Visual acuity results are presented for the average of both eyes using the Lange scale for off-chart results.

Source: Sponsor's submission.[12]

Table 27: Summary of Harms From First Injection Up to 5 Years After the Second-Eye Injection in Study 301

	Voretigene neparvovec n = 20	Control/voretigene neparvovec[a] n = 9	Total N = 29
Patients with ≥ 1 adverse event			
n (%)	20 (100)	9 (100)	29 (100)
Most common events,[b] n (%)			
Leukocytosis	9 (45)	2 (22)	11 (38)
Vomiting	8 (40)	2 (22)	10 (34)
Pyrexia	7 (35)	2 (22)	9 (31)
Nasopharyngitis	7 (35)	1 (11)	8 (28)
Headache	7 (35)	6 (67)	13 (45)
Nausea	6 (30)	4 (44)	10 (34)
Cough	6 (30)	2 (22)	8 (28)
Oropharyngeal pain	6 (30)	1 (11)	7 (24)
Cataract	4 (20)	1 (11)	5 (17)
Intraocular pressure increased	4 (20)	1 (11)	5 (17)
Hematuria	3 (15)	0	3 (10)
Eye inflammation	2 (10)	0	2 (7)
Retinal tear	2 (10)	1 (11)	3 (10)
Abdominal pain upper	2 (10)	1 (11)	3 (10)
Diarrhea	2 (10)	0	2 (7)
Adverse drug reaction	2 (10)	0	2 (7)
Upper respiratory tract infection	2 (10)	0	2 (7)
Animal bite	2 (10)	0	2 (7)
Epistaxis	2 (10)	0	2 (7)
Nasal congestion	2 (10)	2 (22)	4 (14)
Eye pain	1 (5)	1 (11)	2 (7)
Eye pruritus	1 (5)	1 (11)	2 (7)
Macular hole	1 (5)	1 (11)	2 (7)
Musculoskeletal pain	1 (5)	1 (11)	2 (7)

	Voretigene neparvovec n = 20	Control/voretigene neparvovec[a] n = 9	Total N = 29
Dizziness	1 (5)	1 (11)	2 (7)
Rash	1 (5)	1 (11)	2 (7)
Retinal deposits	0	3 (33)	3 (10)
Anxiety	0	2 (22)	2 (7)
Patients with ≥ 1 SAE			
n (%)	4 (20)	1 (11)	5 (17)
Most common event, n (%)			
Retinal detachment	1 (5)	0	1 (3)
Retinal disorder	0	1 (11)	1 (3)
Adverse drug reaction	2 (10)	0	2 (7)
Pneumonia	1 (5)	0	1 (3)
Convulsion	1 (5)	0	1 (3)
Menorrhagia[c]	1 (8)	0	1 (6)
Patients who stopped treatment due to adverse events			
n (%)	0	0	0
Deaths			
n (%)	0	0	0

SAE = serious adverse event; vg = vector genome.

[a] Patients in the control group became eligible to receive voretigene neparvovec at a dose of 1.5×10^{11} vg in a total subretinal volume of 300 µL in each eye (within 18 days) 1 year after their baseline evaluations, provided they still met all eligibility criteria.

[b] Adverse events were reported in more than 1 patient.

[c] Denominator includes only female patients.

Source: Clinical Study Report for Study 301.[15]

Appendix 4: Description and Appraisal of Outcome Measures

Aim

To describe the following outcome measures and review their measurement properties (validity, reliability, responsiveness to change, and MID):

- MLMT
- FST testing
- VA using Holladay and Lange scales for off-chart
- VF using Goldmann test for kinetic fields and Humphrey test for static fields
- adapted VFQ

Findings

Evidence from validation studies is summarized for all instruments according to the following measurement properties and depending on information availability: comprehensiveness, feasibility, validity (i.e., content, construct [convergent, discriminant], and criterion [concurrent, predictive] validity), reliability (internal consistency; i.e., inter-item correlations), reproducibility (i.e., test-retest [inter- and intra-rater] reliability); responsiveness, floor and ceiling effects, and scaling assumptions.

Interpretation of the reliability and validity metrics were based on the following criteria:

- Inter- and intra-rater reliability or agreement (kappa statistics or interclass coefficient): less than 0 to 0.2 = poor, 0.21 to 0.4 = fair, 0.41 to 0.6 = moderate, 0.61 to 0.8 = substantial, 0.81 to 1.00 = almost perfect agreement.[70]
- Internal consistency (Cronbach alpha) and test-retest reliability (≥ 0.7 is considered acceptable).[71]
- Validity, such as between-scale comparison (correlation coefficient, r ≤ 0.3 = weak, 0.3 to ≤ 0.5 = moderate, > 0.5 = strong).[72]

Table 28: Summary of Outcome Measures and Their Measurement Properties

Outcome measure	Type	Conclusions about measurement properties	MID
MLMT	Tool for mobility and functional vision assessment; measures accuracy and time needed to navigate a special MLMT course in various standardized lighting conditions	**Validity:** Construct validity demonstrated **Reliability:** High inter-observer, test-retest, and intra-observer reproducibility demonstrated **Responsiveness:** Demonstrated with questionable methodology Assessment done in individuals with 20/20 vision[a] and IRDs; developed and implemented by the study sponsor	In patients with IRD, the FDA suggests an MID of ≥ 2 light (lux) levels[37]
FST	Measure of light sensitivity; quantifies the dimmest flash of light detected	**Validity:** No evidence found	The CSR for Study 301, referencing a study by Bittner et al.[14] suggested

Outcome measure	Type	Conclusions about measurement properties	MID
		Reliability: Acceptable inter-visit variability and coefficients of repeatability found **Responsiveness:** Not assessed Assessments done in individuals with 20/20 vision and patients with IRDs, RP, macular disease, optic nerve disease, diabetic retinopathy, and other retinal diseases	a clinical significance threshold of 10 dB or 1 log change for the FST test, although methodology unclear
VA using ETDRS, adapted using Holladay and Lange scales for off-chart	Measure of central visual sharpness ETDRS charts present a series of 5 letters of equal difficulty on each row, with standardized spacing between letters and rows for a total of 14 lines (70 letters) Holladay and Lange adaptations done for off-chart VA	**Validity:** No evidence found **Reliability:** Results were mixed or uncertain; test-retest and coefficients of reproducibility reported using unclear methodology **Responsiveness:** No evidence found Assessments done in patients with various forms of vision loss (e.g., RP, macular disease, optic nerve disease, diabetic retinopathy, and other retinal diseases)	3 lines, 0.3 LogMAR, or 15 letters, although estimates are based on patient's macular edema[52]
VF using Goldmann testing for kinetic fields and Humphrey testing for static fields	Measure of entire scope of vision, including central and peripheral vision	Limited and weak evidence of validity and reliability found, with uncertain interpretation; responsiveness not assessed Assessment done in patients with IRDs as well as healthy individuals	FDA suggested an MID of 7 dB for Humphrey macula threshold, based on expert opinion[37]
Adapted VFQ-25	Original version: 25-item condition-specific HRQoL measure; measures visual disability and visual symptoms on generic health for 11 vision-related constructs, in addition to a single-item general health component Adapted version (used in Study 301) removed items related to HRQoL and replaced them with items measuring vision-dependent activities of daily living	No evidence of validity, reliability, or responsiveness found for the adapted version	Not found for the adapted version

CSR = Clinical Study Report; dB = decibel; ETDRS = Early Treatment Diabetic Retinopathy Study; FST = full-field sensitivity threshold; HRQoL = health-related quality of life; IRD = inherited retinal dystrophy; LogMAR = logarithm of the minimum angle of resolution; MID = minimal important difference; MLMT = multi-luminance mobility testing; RP = retinitis pigmentosa; VA = visual acuity; VF = visual field; VFQ-25 = 25-item Visual Function Questionnaire.

[a] The term 20/20 vision is used to express the clarity or sharpness of vision measured at a distance of 20 feet. If a person had 20/20 vision, that person can see clearly at 20 feet what should normally be seen at that distance; if a person has 20/100 vision, it means that person must be as close as 20 feet to see what a person with 20/20 vision can see at 100 feet.[40]

Multi-Luminance Mobility Testing

The MLMT is a mobility test designed to track functional vision changes over time. Functional vision is the ability to perform visually dependent ADL independently. It involves vision assessment at various distances and in a variety of settings and can also include orientation and mobility tasks.[41] The following is a description of how MLMT is conducted as well as the assignment and interpretation of scoring.

The MLMT was developed and validated by a group of researchers in collaboration with the sponsor of voretigene neparvovec. The MLMT was first implemented and standardized in 2 phase I trials (studies 101 and 102) and subsequently used in the pivotal phase III trial (Study 301). In this test, the objective for patients is to navigate a marked path (MLMT course, approximately 24 feet [7 m]) while avoiding obstacles in different lighting conditions. A total of 12 unique MLMT course configurations were developed, with course configuration randomly changed before each testing attempt to minimize memorization and learning effect. The path to follow in the MLMT course was indicated with standardized black arrows on a white background; each course layout featured the same number of arrows, turns, and obstacles.

The test was conducted in 7 standardized lighting conditions (at decreasing levels of luminance) to identify the lowest light level at which the patient could successfully navigate the course (termed the estimated lower light–sensitivity cut-off) as well as the illuminance level 1 level lower than the estimated lower light–sensitivity cut-off, at which the patient is unable to complete the course (termed sub-sensitivity cut-off). In other words, if a patient navigated the course accurately at a given light level, the next testing occurred at this same light level and an illuminance level below the specified light scale. If the patient appeared to have difficulty navigating accurately at a given light level, the light level was maintained or increased at the subsequent follow-up visit. Conversely, if a patient navigated the course accurately at a given light level but failed at the same light level during the subsequent test, the patient continued to be tested at higher light levels in order to identify a passing level (termed the supra-sensitivity cut-off). The light levels used in the trials ranged from 1 lux to 400 lux, corresponding to common real-world luminance levels, such as a moonless summer night or indoor nightlight to office environment or food court, respectively (Table 29).

Prior to actual testing, patients underwent a practice run of the course, with both eyes unpatched, then with alternate eyes patched. The practice course layout was different from the 12 configurations used for the actual testing runs. Up to 4 light levels were tested, including the estimated lower light–sensitivity cut-off as well as 1 or 2 light levels below and above the lowest light level for each eye. Patients were then dark-adapted (a procedure in which light is eliminated or reduced from a patient's view for a set period of time) for 40 minutes while the course was configured for testing. Following dark adaptation, testing proceeded from lower to higher light levels, first at the sub-sensitivity cut-off level, followed by testing at the estimated lower light–sensitivity cut-off and, finally, at the supra-sensitivity cut-off level. If the baseline estimated light-sensitivity cut-off differed significantly from 1 eye to the other, additional tests could be performed. The tester coordinated the whole process, providing cues to the patient to start and stop the test and notifying patients if they went off the course, risked injury or collision, or required redirection. No other verbal or physical cues were provided. The testing was done with each eye assessed individually (single eye patched in a randomized manner) and with both eyes unpatched. Following the completion of each test run, a new course configuration was introduced for the subsequent run. For

each testing visit, a patient was allowed to navigate the course a maximum of 12 times, with 12 different configurations. For each eye-patching condition, a passing level and a failing level were determined in the following way.[41]

The tests were recorded (using high-definition cameras capable of capturing clear images at low illuminance) and graded by independent assessors who were trained on MLMT scoring but were blinded to the study protocol, objectives, patient information, and results of other visual function tests. Each test was graded based on a predetermined algorithm that assesses both speed (time to complete the course) and accuracy (avoidance of obstacles). The following parameters were assessed for speed and accuracy: the number of collisions, the time to navigate the course, the number of times the test-taker went off course, the number of arrows that were bypassed, and the number of times the test-taker had to be redirected onto the course. A penalty was assigned for every collision with obstacles (regardless of form of contact) and other errors, accounting for errors in both speed and accuracy. The accuracy score was then obtained by dividing the number of accuracy penalties by the total number of obstacles (15). A perfect score was 0; a minimum accuracy score of 0.25 or less was required to pass, which translates to a maximum of 3 errors allowed for a pass. Time score is simply the time taken (in seconds) to complete the course, plus time penalties. Time penalties were assigned as follows: 15 seconds for each instance of being off course, 15 seconds for each tile bypassed, and 30 seconds for each redirect. A time score of less than 180 seconds was required to pass. Test-takers had to pass both the accuracy and time components of the test to pass each course run.[41]

In addition to the MLMT scoring algorithm described previously, an MLMT change-score algorithm was developed to quantify patient performance over time, in which each light level evaluated during the study was assigned a discrete lux score from −1 to 6, with lower light levels corresponding to higher lux scores (Table 29).

The MLMT change score was then determined by taking the difference in score code from the baseline lower light–sensitivity cut-off (i.e., the lowest light level at which a patient passed at baseline) to the lower light–sensitivity cut-off at the follow-up visit.[41]

Table 29: Lux Level and Corresponding MLMT Score Code and Real-World Environment

Illuminance (lux)	Score code	Corresponding environment
1	6	Moonless summer night, or indoor nightlight
4	5	Cloudless summer night with half moon, or outdoor parking lot at night
10	4	60 minutes after sunset in a city setting, or a bus stop at night
50	3	Outdoor train station at night, or inside of illuminated office building stairwell
125 (100 and 150)[a]	2	30 minutes before cloudless sunrise, or interior of shopping mall, train, or bus at night
250 (200 and 250)[b]	1	Interior of elevator, library, or office hallway
400	0	Office environment, or food court
> 400[c]	−1	NA

MLMT = multi-luminance mobility testing; NA = not applicable.

[a] Analysis incorporates runs at 100 lux and 150 lux.

[b] For analysis purposes, 200 lux and 250 lux were combined; however, no patients were tested at either of these levels because these levels were not needed to identify any lowest-passing or highest-failing level.

[c] Does not pass at 400 lux.

Source: Chung et al.[41]

Validity and Reliability

Chung et al.[41] assessed the validity and reliability of the MLMT in collaboration with the sponsor in a prospective, observational study in 54 individuals who had 20/20 vision or were visually impaired with IRDs (of which 20 were diagnosed with LCA and 4 with RP). Discriminant validity was demonstrated by showing the MLMT could differentiate between individuals with 20/20 vision and impaired vision based on time and accuracy in completing the MLMT. Individuals with 20/20 vision passed on both time and accuracy at all light levels, whereas patients with visual impairment demonstrated a wide range of failing and passing performances. Among patients with impaired vision, higher performers could be differentiated from lower performers, depending on levels of illumination.[41]

Construct validity was evaluated by assessing whether the test components of the speed and accuracy score at defined levels of illumination correlated with aspects of VF, VA, and HRQoL (using the VFQ). Performance on the MLMT correlated with other conventional visual function tests, including VF (using Goldmann VFs and Humphrey VFs) and VA (measured in LogMAR units), in a way that showed a threshold effect of the MLMT, i.e., deterioration on the MLMT declined markedly when VA and VF fell below certain thresholds. Among patients with visual impairment, the correlation between the accuracy score and VA was strong (range 0.75 to 0.86), moderate to strong between the accuracy score and VF sum total degrees (range −0.37 to −0.53 for Goldmann VFs with V4e and III4e test stimuli, and −0.37 to −0.64 for Humphrey VFs with foveal sensitivity and macula threshold), and high between the accuracy score and parent or guardian assessment using VFQ (−0.6).[41]

The responsiveness of the test (i.e., ability to detect change over time) was assessed by monitoring the changes in the MLMT score with disease progression over the course of 1 year. However, the methodology was not appropriate for testing responsiveness. It was shown that the MLMT was able to track declines in performance during this time in those with RP or LCA, but not among those with other forms of visual impairment. However, the changes in performance over time were based on the MLMT score change as opposed to an external anchor to categorize disease progression or severity. [41]

In addition to the validity parameters mentioned previously, it was noted that no validation work had been done to demonstrate that the 12 configurations are equivalent in terms of difficulty.

The following reliability parameters were measured: inter-observer, test-retest, and intra-observer reproducibility using interclass coefficients for individual components of the MLMT and kappa statistics for MLMT completion and pass or fail outcomes. Using data from more than 3,500 video-recorded MLMT tests from this study as well as studies 101, 102, and 301, a high inter-observer, test-retest, and intra-observer reproducibility was shown in test scores (interclass coefficient > 0.9). Additionally, agreement between independent, masked graders across the many continuous and binary pass or fail components of the MLMT was high (kappa > 95%). Among patients with visual impairment, concordance between the baseline visits across all light levels and times was high for all pass and fail components: 88% for the bilateral MLMT pass or fail outcome and 86% for time.[41]

Minimal Important Difference

The sponsor considered a change of 1 light level in passing the MLMT as a clinically significant effect, given the progressive nature of the condition.[73] However, no supporting documentation was provided to report how this MID was derived. The rationale was that a

patient's safety and independence is tied closely to their ability to navigate more quickly and more accurately at lower light levels. In the article by Chung et al.,[41] it is stated that since the lighting levels selected and utilized for MLMT testing cover a broad range of everyday situations, the ability to safely navigate in dimmer conditions, in general, should provide patients with the independence and safety to carry out ADL more efficiently. However, the FDA reported that an MLMT score change of 1 may represent background fluctuation. In addition, the FDA noted the interval between some luminance levels is uneven. Based on these observations, the FDA considered a change in MLMT score of 2 or more as an acceptable MID.[37] Similarly, the EMA indicated that any clinically relevant change in MLMT with voretigene neparvovec would need to exceed 1 light level; however, the EMA did not indicate what change would be considered clinically relevant.[38]

FST Testing

The FST is a measure of light sensitivity of the entire VF and is aimed at detecting the lowest luminance of a flash detected by a patient. In this test, flashes of various luminance (range spanning approximately 80 dB) are presented to individuals, who are then asked to press a response button if they are able to see the visual stimulus.[42] Using a predefined algorithm described in Roman et al.,[74] a different luminance is presented based on the response from each preceding test run to identify a threshold for each eye. Three wavelengths of stimuli are tested (blue, red, and white) multiple times, and the threshold is calculated as an average of all trials for each stimulus. FST results are expressed in dB, which are then converted to $\log_{10}(cd.s/m^2)$ to accommodate different dB conversion rates. Smaller values of dB and $cd.s/m^2$ indicate better sensitivity, and negative $\log_{10}(cd.s/m^2)$ values indicate better sensitivity.[15]

Validity and Reliability

Developers of the FST protocol assessed the test-retest reliability based on the mean sensitivity difference between study visits spanning 6 months among patients with inherited retinal degenerations and individuals with 20/20 vision. Inter-visit variability was not significantly different between visits for either group and was not reportedly dependent on mean sensitivity level.[43]

Kiser et al. assessed the reproducibility of a number of psychophysical vision measures, including FST, among individuals considered legally blind (RP, macular disease, optic nerve disease, diabetic retinopathy, and other retinal diseases) and 20/20 vision. The coefficients of repeatability (based on the Bland-Altman method) for FST varied from 5 dB to 15 dB depending on the ocular condition, and the authors concluded that the FST produced repeatable results at all levels of vision loss and for all disease states.[44]

Minimal Important Difference

The CSR for Study 301 suggested a clinical significance threshold of 10 dB or 1 log change for the FST test, referencing a study by Bittner et al.,[14] which is a case series involving 12 adult patients with RP. Visual function was assessed using a number of tests, including dark-adapted FST. The authors reported an improvement in FST of 10.3 dB to 17.5 dB (13- to 53-fold) following acupuncture, which was reportedly outside the typical test-retest variability of 3 dB to 3.5 dB. Based on these findings, the lower end of the VF improvement was taken as an MID.[14]

Visual Acuity

VA is a measure of the ability of the eye to distinguish shapes and the details of objects, also knowns as optotypes (individual letters of standardized size and contrast presented to the test-taker for VA assessment), from a set viewing distance.[15] The ETDRS chart is a commonly used tool for testing VA in clinical studies and was used for most patients in Study 301. ETDRS charts present a series of 5 letters of equal difficulty on each row, with standardized spacing between letters and rows for a total of 14 lines (70 letters). ETDRS letter score can be calculated when 20 or more letters are read correctly at 4.0 m; the VA letter score is equal to the total number of letters read correctly at 4.0 m plus 30. If fewer than 20 letters are read correctly at 4.0 m, the VA letter score is equal to the total number of letters read correctly at 4.0 m (number recorded on line 1), plus the total number of letters read correctly at 1.0 m in the first 6 lines. Therefore, the ETDRS letter score can result in a maximum score of 100.[45,46] Standard chart-testing distance is 4 m; however, shorter distances may be used when vision is severely impaired.[47,48] Scoring for EDTRS charts is designed to produce a LogMAR suitable for statistical analysis in which individual letters score 0.02 log units.[49] The authors of the trial reported an improvement of 0.1 LogMAR corresponded to a 5-letter improvement (or equivalent of 1 line) on an ETDRS eye chart. Notably, a lower LogMAR score represents an improvement in VA.[15]

Among patients in Study 301 who were unsuccessful in correctly identifying the line of largest letters on the ETDRS chart, off-chart VA measurements were collected by counting fingers and evaluating hand-motion perception, light perception, and no light perception. For off-chart VA measurements, LogMAR values were assigned using the scale adapted from Holladay et al.[50] and Lange et al.[51] The authors of the trial reported the Lange method was implemented in response to the EMA's concern that a difference of 1 log-unit step between counting fingers and hand-motion perception in the Holladay method could present a biased estimate of a treatment effect (improvement or reduction in LogMAR). The Lange scale was therefore recommended as a sensitivity analysis for VA assessment in which the 1 log-unit step between counting fingers and hand motion is reduced to a 0.3 log-unit step.[15,51]

Validity and Reliability

CADTH's literature search did not find studies assessing the validity of the ETDRS; however, a number of studies assessing its reliability was found. One study reported the test-retest reliability of the electronic version of the ETDRS compared with the standard version (used in Study 301) in patients with various forms of vision loss. For both versions, the test-retest reliability was high (r = 0.99), with 98% of retests falling within 0.2 LogMAR of the initial test.[75]

Kiser et al.[44] assessed the repeatability of 2 psychophysical vision measures, including the Pelli-Robson letter contrast sensitivity test and ETDRS, in patients with severe vision loss (resulting from RP, macular disease, optic nerve disease, diabetic retinopathy, and other retinal diseases). The ETDRS was assessed under regular and dim illumination, and up to 5 test repetitions were performed at monthly intervals. For the ETDRS, the mean 95% coefficient of repeatability ($CR_{.95}$) for most ocular disease subgroups was 0.20 log units (range of between 0.13 log units and 0.36 log units for various ocular disease subgroups). Both measures showed comparable repeatability, but it is not clear if the repeatability can be considered high against established metrics (kappa statistics or interclass coefficient). Notably, among the major findings of the study, the authors suggested that reliability may be related to visual degradation or ocular disease, that the measures may not be as

effective when monitoring for small amounts of change in patients with severe light or dim-light adaptation problems, and that the documented reliability of these psychophysical tests will allow for monitoring vision changes in a clinical trial setting.[44]

Another study comparing the reliability of VA and contrast sensitivity was Bittner et al., where between-visit test-retest variability in mesopic measures of VA (ETDRS) and contrast sensitivity (Pelli-Robson chart) was assessed over 2 study visits (1 to 2 weeks apart) in patients with RP. The test-retest $CR_{.95}$ was not statistically significantly different when comparing between photopic and mesopic tests of VA (0.16 LogMAR and 0.12 LogMAR, respectively). The authors noted that ETDRS had good acceptable test-retest repeatability and is a suitable outcome measure to monitor mesopic visual function in clinical practice or trials.[76]

NICE reported that the measures of VA are often considered unreliable due to inter-test variability.[53]

Minimal Important Difference

An IRD-specific MID for ETDRS was not identified from the literature search done by CADTH. ETDRS charts may reliably identify changes in VA of 2 lines (10 letters) or more, but not changes of 1 line (5 letters) or fewer. A loss or gain of 3 lines (15 letters) is considered a moderate degree of change and is commonly used as an outcome in clinical trials.[77] The FDA considers a mean change of 15 letters or more on a standard ETDRS chart or a change of 0.3 LogMAR as clinically meaningful, based on expert opinion.[37]

The test-retest variability of a measure can help guide what would be considered a clinically meaningful change. Beck et al.[75] reported that a change in VA of 0.2 LogMAR (10 letters) from a baseline level is unlikely to be related to measurement variability using either the electronic or standard version of the ETDRS VA testing protocol, indicating this could be used as an MID.[75]

Visual Field

A VF test, also known as perimetry, is a measure of an individual's entire scope of vision, including central and peripheral vision. With this test, the VF of each eye is mapped individually and blind spots (scotomas), as well as more subtle areas of dim vision, can be detected. There are different types of VF tests, e.g., confrontation VF testing, Amsler grid, frequency doubling perimetry, static automated perimetry (such as the Humphrey field analyzer) and kinetic perimetry (such as the Goldmann perimetry). The Goldmann perimetry is generally used to capture the entire VF, whereas the Humphrey testing is typically reserved to specific domains within the VF. In Humphrey computerized testing for static fields, flashes of light of varying size and brightness are projected within a large white bowl. The patient is asked to look at the centre of the bowl and press a button each time a light is seen in their peripheral vision. An automated computer algorithm controls the whole test in order to determine the field of vision. In Goldmann perimetry for kinetic fields, moving targets of various light sizes and intensities are shown and the patient indicates when they become visible in their peripheral vision. The resulting data are used to map the full VF, which normally extends approximately 120° vertically and 160° horizontally.[15]

In Study 301, VF was assessed using both Goldmann perimetry for kinetic fields and Humphrey testing for static fields, allowing for evaluation of different regions of the retina. For Goldmann VF testing, the stimuli vary in size and are expressed using a Roman numeral (indicating the Goldmann size of the stimulus), an Arabic number, and a letter

(indicating the attenuation of the light). In Study 301, both the size III4e and size V4e test stimuli were used, the III4e target being 1/16th smaller in total area and one-quarter the diameter of the V4e target. The test stimulus at baseline was used to determine which test stimulus would be used for follow-up visits. For example, if the size III4e test stimulus resulted in individuals being able to see and reliably perform the VF testing, this size was used for each subsequent visit and vice versa. In contrast, the Humphrey testing for static fields (macula sensitivity and foveal sensitivity thresholds) always used a size V isopter. The Humphrey VF macula threshold testing targets both cone and rod photoreceptor cells, while the Humphrey foveal sensitivity targets the most central, cone-enriched region of the macula. Results for the Goldmann perimetry were presented as sum total degrees; higher sum totals indicate a greater area of functional, light-sensitive retina, corresponding to a greater field of vision for the patient. Results for the Humphrey VF testing was reported in dB of attenuation, or dimming, ranging from 0 dB (the brightest, unattenuated stimulus) to 51 dB (the dimmest, maximally attenuated stimulus); higher numbers (dB) corresponded to higher retinal sensitivity.[15]

Validity and Reliability

The literature reviewed by CADTH showed limited evidence of the validity of the Goldmann perimetry and Humphrey testing method. Indeed, 1 report noted the Goldmann method is the accepted gold standard for recording VF,[78] while others used Goldmann perimetry as a reference when comparing the diagnostic accuracy of other kinetic VF assessment tools.

In contrast to validation studies for VF measures, a number of studies reporting various reliability estimates were found in the literature. It should be noted that due to the heterogeneity and complexity of the study methodology, the instrument used, the study procedure, the choice and assessment of reliability estimates, and the variation in study population and context, the findings of these studies may have limited application to this review. Overall, the reliability of both Goldmann and Humphrey VF perimetry can be considered mixed and inconclusive. Notably, NICE considers measures of VF unreliable because of inter-test variability.[53]

Barnes et al.[79] compared the test-retest variability (among other assessments) of a semi-automated kinetic perimetry, with Goldmann manual kinetic perimetry as standard, among pediatric and adult patients (median age 20 years) with IRDs. Test-retest reliability was assessed over 2 visits about 1 week apart. Pertinent to this review, Goldmann VF perimetry showed high between-visit correlations ($r^2 > 0.9$), with a median test-retest variability of less than 10%. Overall variability was reportedly lower for children than for adults.[79]

Dedania et al.[80] assessed the change in VF using Goldmann perimetry with age and disease progression among young children (ages 5 years to 16 years) with retinal dystrophies. Data for at least 2 different visits within the 12-year time frame were collected retrospectively. The repeatability coefficients were high for the 3 tested isopters, 7,381 mm^2 (I4e), 9,379 mm^2 (III4e), and 10,346 mm^2 (IV4e). In addition, VF area increased with age regardless of progression or stability of disease, with greater increases seen at earlier ages (younger than 12 years). Based on the large variability and increased VF area with age, the authors concluded that Goldmann perimetry can be an unreliable measure of visual impairment and response to treatment in children with retinal dystrophies.[80]

One study by Patel et al.[81] compared the feasibility, reliability, and repeatability of 3 perimetry (Humphrey static, Goldmann, and Octopus kinetic perimeters) in children. Although the children had no ophthalmological disease that could cause a VF defect, they

had other reasons for undergoing an eye test. Additionally, the participants were assessed with the Humphrey static SITA 24–2 FAST algorithm, which may differ in protocol compared with the FASTPAC strategy used in Study 301. The authors reported high feasibility (test completion rate of 96.1% for Goldmann and 100% for Humphrey), and "good" examiner-rated reliability (using an Examiner Based Assessment of Reliability scoring system) in 125 (81.2%) and 98 (63.6%) participating children for Goldmann and Humphrey perimetry, respectively. The authors concluded that the feasibility and reliability of formal perimetry in children improved with age, with Goldmann perimetry being the most reliable method in children under 9 years of age.[81]

NICE reported that the measures of VF are often considered unreliable due to inter-test variability.[53]

Minimal Important Difference

CADTH's literature search found no studies reporting an MID for either Goldmann or Humphrey perimetry (including Goldmann V4e, Goldmann III4e, and Humphrey foveal sensitivity). The FDA reported a change of 7 dB to be clinically meaningful for Humphrey macula threshold, based on expert opinion.[37]

The 25-Item Visual Function Questionnaire

The VFQ-25 is a 25-item version of the original 51-item National Eye Institute VFQ that is available in both a self-administered and interviewer-administered format. This patient-reported survey measures the effect of visual disability and visual symptoms on generic health domains such as emotional well-being and social functioning, in addition to ADL. The VFQ-25 includes 25 vision-targeted questions that represent 11 vision-related constructs: near vision (3 items), distance vision (3 items), social functioning (2 items), vision-specific role difficulties (2 items), dependency (3 items), mental health symptoms (4 items), driving difficulties (3 items), limitations with peripheral (1 item) and colour vision (1 item), and ocular pain (2 items). In addition, the VFQ-25 includes a single item that assesses general health status. Patients (or their parent or guardian, where applicable) provide a response to each of the 25 questions using a Likert scale ranging from 0 (worst visual function) to 100 (best visual function). Subscale-specific items are averaged to create summary scores for each subscale, and subscale scores are averaged to estimate the overall composite score.[54]

It should be noted that the VFQ used in Study 301 was modified by the investigative team to evaluate the ADL that are dependent on vision, or have a vision element, while items related to HRQoL were removed. The adaptations were done to accommodate IRD-associated poor vision and to include a pediatric population. The scoring system for the individual domains was changed as well, with the perceived difficulty of these activities rated on a numerical scale from 0 to 10 (0 being the most difficult), before the average of the responses is taken to determine the total or overall score for each individual.[15]

Validity and Reliability

Both the 25-item and 51-item versions of the VFQ have been extensively validated in various ophthalmic conditions and in a variety of age groups (i.e., adult and pediatric populations);[55-58] however, the adapted version of the VFQ used in Study 301 has not been assessed psychometrically.

Minimal Important Difference

A number of studies assessed a clinically meaningful change for the standard VFQ-25;[49,59] however, an MID was not assessed for the adapted version of the scale, nor was any evidence found based on the literature search. Given the modifications made to the original VFQ, the MIDs identified in the literature for that measure were not considered directly generalizable to the version used in Study 301.

Appendix 5: Summary of *RPE65* Mutation Testing

The material considered in this section was provided as supporting information. The information has not been systematically reviewed.

Aim

To summarize the use of *RPE65* mutation testing in patients with IRDs.

Findings

Overview of Visual Cycle and *RPE65* Mutation

The visual cycle (or retinoid cycle) is the process in which retinal photoreceptors are activated by light, following which a series of biochemical reactions take place in which 11-*cis*-retinal goes through a chemical conversion into all-*trans*-retinal, then all-*trans*-retinol, before 11-*cis*-retinol is regenerated again in the retinal pigment epithelium cells. The regeneration of 11-*cis*-retinol requires an enzyme, all-*trans*-retinol isomerase, which is also known as RPE65.[82,83] Mutation in the *RPE65* gene encoding this protein results in damage to retinal pigment epithelium cells, which causes damage to the photoreceptors over time. *RPE65* mutation primarily affects rod photoreceptors that mediate peripheral vision and the ability to see in low light.[5] A decrease in light sensitivity leads to night blindness which, over time, affects daily activities under regular daytime lighting conditions.[5]

List of *RPE65* Mutation Testing Platforms

In preparation for this review, the sponsor provided a non-exhaustive list of genetic tests available in Canada for the diagnosis of *RPE65* mutation obtained through consultations with patient groups and experts. These testing platforms are listed below:

- LifeLabs Genetics: Retinitis pigmentosa type 20, autosomal recessive (2020)
- LifeLabs Genetics: Leber congenital amaurosis panel (2020)
- LifeLabs Genetics: Retinitis pigmentosa panel, autosomal dominant and recessive
- Blueprint Genetics: Retinal dystrophy panel
- Molecular Vision Lab: MVL Vision Panel
- Prevention Genetics: Retinitis pigmentosa
- Prevention Genetics: Leber congenital amaurosis
- Invitae: *RPE65* single gene
- Invitae: Leber congenital amaurosis panel
- Cen4Gen

These genetic tests include platforms specific for identifying single-gene mutation, or a panel of genetic mutations, and can be used to diagnose or assess the risk of a given condition. These tests generally have high accuracy, with sensitivity and specificity well over 95% in most cases.[84] The clinical experts consulted for this report shared the view that these genetic testing platforms have a high level of accuracy.

Interpretation of Genetic Test Results

This section provides a brief overview on how to interpret genetic test results, which applies to all forms of genetic tests, including tests to identify *RPE65* mutation. Results of genetic tests are not always presented in a straightforward binary (yes or no) manner; instead, 3 possible scenarios can arise. A positive test result is indicative of a change in a particular gene or chromosome. Testing for *RPE65* mutation is generally done to confirm clinical diagnosis once RP or LCA is suspected. A negative test result for *RPE65* mutation in patients with an IRD may indicate other genetic mutations in play, although the use of a genetic test that can identify a panel of gene mutations (as opposed to a single gene mutation test) can circumvent this problem. In addition to a positive or negative result, genetic test results can sometimes be uninformative or indeterminate, either because the genetic variation occurs naturally in the population (polymorphisms) without affecting health, or the diagnosed genetic variation is not assigned to a particular disease. This is unlikely in the case of RPE65-associated IRD since the role of RPE65 in causing IRD is well documented. The likelihood of a confirmatory result from a genetic test depends on an accurate diagnosis of the clinical condition, in addition to the accuracy of the test itself. This is because some conditions have a single or a few underlying genetic mutations that are well-known; therefore, a confirmatory genetic test result is more likely if an accurate clinical diagnosis is made. On the other hand, having multiple disease-causing mutations (such as in IRDs, which may be caused by any of more than 260 different gene mutations) create difficulty in accurately identifying the underlying mutation(s), unless all underlying mutations can be tested in 1 testing platform. A positive clinical diagnosis of the condition needs to be specific enough (e.g., RP and LCA, which have a higher prevalence of *RPE65* mutation than other forms of IRDs) such that the suspected underlying mutation is more likely to be captured by testing.[85]

Availability of Genetic Tests in Canada

In preparation for this review, a survey was sent to publicly funded drug plans in Canada to determine the availability of genetic tests for detecting the *RPE65* mutation. A response was received from 7 plans. The most common response was that genetic testing was not available locally, but that testing may be accessed by sending samples out of the province or country for analysis if requested by a geneticist or appropriate governing body on genetic testing. The clinical experts consulted for this review shared the same message regarding funding structure, and no concern was raised regarding the availability of genetic tests, although a lack of coordination between medical geneticists and physicians was noted as an implementation challenge.

Conclusion

The sponsor indicated there are a number of genetic tests available in Canada for diagnosing the *RPE65* mutation; some of these tests are targeted at a single gene mutation and some test a panel of genetic mutations. The accuracy of these tests is generally high, as noted by clinical experts and the testing companies. However, most Canadian jurisdictions indicated that genetic tests are not readily available. Instead, samples are sent out of the province or out of the country for analysis, upon request from a medical geneticist, with local or provincial bodies supporting the cost of tests.

References

1. Retinal Information Network (RetNet). Summaries of genes and loci causing retinal diseases. Gene and locus symbols by disease category (one or more diseases per gene/locus). 2020; https://sph.uth.edu/retnet/sum-dis.htm#B-diseases. Accessed 2020 Jul 2.
2. Kumaran N, Moore AT, Weleber RG, Michaelides M. Leber congenital amaurosis/early-onset severe retinal dystrophy: clinical features, molecular genetics and therapeutic interventions. *Br J Ophthalmol.* 2017;101(9):1147-1154.
3. Berger W, Kloeckener-Gruissem B, Neidhardt J. The molecular basis of human retinal and vitreoretinal diseases. *Prog Retin Eye Res.* 2010;29(5):335-375.
4. Russell S, Bennett J, Wellman JA, et al. Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in patients with RPE65-mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial. *Lancet.* 2017;390(10097):849-860.
5. Tang PH, Kono M, Koutalos Y, Ablonczy Z, Crouch RK. New insights into retinoid metabolism and cycling within the retina. *Prog Retin Eye Res.* 2013;32:48-63.
6. Ferrari S, Di Iorio E, Barbaro V, Ponzin D, Sorrentino FS, Parmeggiani F. Retinitis pigmentosa: genes and disease mechanisms. *Curr Genomics.* 2011;12(4):238-249.
7. Fighting Blindness Canada. Retinitis pigmentosa. 2018; https://www.fightingblindness.ca/eye-diseases-pathways/retinitis-pigmentosa/. Accessed 2020 Jul 2.
8. Understanding retinitis pigmentosa. London (GB): Royal National Institute of Blind People; 2017: https://www.rnib.org.uk/sites/default/files/Understanding_Retinitis_Pigmentosa_NV.pdf. Accessed 2020 Jul 2.
9. Chung DC, Bertelsen M, Lorenz B, et al. The natural history of inherited retinal dystrophy due to biallelic mutations in the RPE65 gene. *Am J Ophthalmol.* 2019;199:58-70.
10. American Academy of Opthalmology (AAO). Clinical statement: recommendations on clinical assessment of patients with inherited retinal degenerations 2016. 2016; https://www.aao.org/clinical-statement/recommendations-on-clinical-assessment-of-patients. Accessed 2020 Jul 2.
11. Ganesh A, Keep RB. Genetic testing in retinal dystrophies. *Oman J Ophthalmol.* 2011;4(3):105-107.
12. CDR submission: Luxturna (voretigene neparvovec), 5 x 10^{12} vector genomes/mL concentrate for solution for subretinal injection [**CONFIDENTIAL** sponsor's submission]. Dorval (QC): Novartis Pharmaceuticals Canada Inc.; 2020 Apr 30.
13. Luxturna (voretigene neparvovec): 5 x 10^{12} vector genomes/mL concentrate for solution for subretinal injection [product monograph]. Dorval (QC): Novartis Pharmaceuticals Canada Inc.; 2020 Oct 13: https://pdf.hres.ca/dpd_pm/00058324.PDF. Accessed 2020 Nov 10.
14. Bittner AK, Gould JM, Rosenfarb A, Rozanski C, Dagnelie G. A pilot study of an acupuncture protocol to improve visual function in retinitis pigmentosa patients. *Clin Exp Optom.* 2014;97(3):240-247.
15. Clinical study report addendum 2017: AAV2-hRPE65v2-301. A safety and efficacy study in subjects with Leber Congenital Amaurosis (LCA) using adeno-associated viral vector to deliver the gene for human RPE65 to the Retinal Pigment Epithelium (RPE) [**CONFIDENTIAL** internal sponsor's report]. Philadelphia (PA): Spark Therapeutics; 2018 Feb 15.
16. Stone EM, Andorf JL, Whitmore SS, et al. Clinically focused molecular investigation of 1000 consecutive families with inherited retinal disease. *Ophthalmology.* 2017;124(9):1314-1331.
17. Morimura H, Fishman GA, Grover SA, Fulton AB, Berson EL, Dryja TP. Mutations in the RPE65 gene in patients with autosomal recessive retinitis pigmentosa or Leber congenital amaurosis. *Proc Natl Acad Sci U S A.* 1998;95(6):3088-3093.
18. Thompson DA, Gyürüs P, Fleischer LL, et al. Genetics and phenotypes of RPE65 mutations in inherited retinal degeneration. *Invest Ophthalmol Vis Sci.* 2000;41(13):4293-4299.
19. Redmond TM, Yu S, Lee E, et al. Rpe65 is necessary for production of 11-cis-vitamin A in the retinal visual cycle. *Nat Genet.* 1998;20(4):344-351.
20. Redmond TM, Poliakov E, Yu S, Tsai JY, Lu Z, Gentleman S. Mutation of key residues of RPE65 abolishes its enzymatic role as isomerohydrolase in the visual cycle. *Proc Natl Acad Sci U S A.* 2005;102(38):13658-13663.
21. Kumaran N, Rubin GS, Kalitzeos A, et al. A cross-sectional and longitudinal study of retinal sensitivity in rpe65-associated Leber congenital amaurosis. *Invest Ophthalmol Vis Sci.* 2018;59(8):3330-3339.
22. Stone EM, Aldave AJ, Drack AV, et al. Recommendations for genetic testing of inherited eye diseases: report of the American Academy of Ophthalmology task force on genetic testing. *Ophthalmology.* 2012;119(11):2408-2410.
23. Lee K, Couser NL. Genetic testing for eye diseases: a comprehensive guide and review of ocular genetic manifestations from anterior segment malformation to retinal dystrophy. *Curr Genet Med Rep.* 2016;4(2):41-48.
24. Fighting Blindness Canada. Genetic testing for inherited retinal diseases. 2016; https://www.fightingblindness.ca/resources/genetic-testing-for-inherited-retinal-diseases/. Accessed 2020 Jul 2.
25. Sugawara T, Hagiwara A, Hiramatsu A, Ogata K, Mitamura Y, Yamamoto S. Relationship between peripheral visual field loss and vision-related quality of life in patients with retinitis pigmentosa. *Eye (Lond).* 2010;24(4):535-539.

26. Knauer C, Pfeiffer N. The value of vision. *Graefes Arch Clin Exp Ophthalmol.* 2008;246(4):477-482.

27. Wang F, Wang H, Tuan HF, et al. Next generation sequencing-based molecular diagnosis of retinitis pigmentosa: identification of a novel genotype-phenotype correlation and clinical refinements. *Hum Genet.* 2014;133(3):331-345.

28. Fighting Blindness Canada. Leber congential amaurosis. 2018; https://www.fightingblindness.ca/eye-diseases-pathways/leber-congenital-amaurosis/. Accessed 2020 Jul 2.

29. Lotery AJ, Namperumalsamy P, Jacobson SG, et al. Mutation analysis of 3 genes in patients with Leber congenital amaurosis. *Arch Ophthalmol.* 2000;118(4):538-543.

30. National Eye Institute. Retinitis pigmentosa. 2019; https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/retinitis-pigmentosa. Accessed 2020 Jul 2.

31. Spark Therapeutics Inc. Prescribing information: Luxturna (voretigene neparvovec) for subretinal injection. Silver Spring (MD): U.S. Food and Drug Administration; 2017: https://www.fda.gov/media/109906/download. Accessed 2020 Jul 8.

32. Product information: Luxturna (voretigene neparvovec). *(European public assessment report)*. Amsterdam (NL): European Medicines Agency; 2020 https://www.ema.europa.eu/en/documents/product-information/luxturna-epar-product-information_en.pdf. Accessed 2020 Jul 8.

33. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17(5):405-424.

34. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 guideline statement. *J Clin Epidemiol.* 2016;75:40-46.

35. Grey matters: a practical tool for searching health-related grey literature. Ottawa (ON): CADTH; 2019: https://www.cadth.ca/grey-matters. Accessed 2020 May 8.

36. Center for Drug Evaluation Research. Statistical review(s). *Luxturna (voretigene neparvovec) subretinal injection. Company: Spark Therapeutics, Inc.. Application No.: BL 125610/0 Approval date: 12/19/2017 (FDA approval package)*. Rockville (MD): U. S. Food and Drug Administration (FDA); 2017 Dec 4: https://www.fda.gov/media/110105/download Accessed 2020 Jun 23.

37. Center for Drug Evaluation Research. Medical review(s). *Luxturna (voretigene neparvovec) subretinal injection. Company: Spark Therapeutics, Inc.. Application No.: BL 125610/0 Approval date: 12/19/2017 (FDA approval package)*. Rockville (MD): U. S. Food and Drug Administration (FDA); 2017: https://www.fda.gov/media/110105/download Accessed 2020 Jun 23.

38. Committee for Medicinal Products for Human Use. Assessment report: Luxturna (voretigene neparvovec). *(European public assessment report)*. Amsterdam (NL): European Medicines Agency; 2018 Sep 20: https://www.ema.europa.eu/en/documents/assessment-report/luxturna-epar-public-assessment-report_en.pdf. Accessed 2020 Jun 23.

39. Maguire AM, Russell S, Wellman JA, et al. Efficacy, safety, and durability of voretigene neparvovec-rzyl in RPE65 mutation-associated inherited retinal dystrophy: results of phase 1 and 3 trials. *Ophthalmology.* 2019;126(9):1273-1285.

40. Vimont C. What does 20/20 vision mean? San Francisco, (CA): American Academy of Ophthalmology 2020: https://www.aao.org/eye-health/tips-prevention/what-does-20-20-vision-mean?gclid = EAIaIQobChMI0fm6vuHh6gIVqdSzCh1N_wj1EAAYASAAEgJV3vD_BwE. Accessed 2020 Aug 20.

41. Chung DC, McCague S, Yu ZF, et al. Novel mobility test to assess functional vision in patients with inherited retinal dystrophies. *Clin Experiment Ophthalmol.* 2018;46(3):247-259.

42. Collison FT, Park JC, Fishman GA, McAnany JJ, Stone EM. Full-field pupillary light responses, luminance thresholds, and light discomfort thresholds in CEP290 Leber congenital amaurosis patients. *Invest Ophthalmol Vis Sci.* 2015;56(12):7130-7136.

43. Roman AJ, Schwartz SB, Aleman TS, et al. Quantifying rod photoreceptor-mediated vision in retinal degenerations: dark-adapted thresholds as outcome measures. *Exp Eye Res.* 2005;80(2):259-272.

44. Kiser AK, Mladenovich D, Eshraghi F, Bourdeau D, Dagnelie G. Reliability and consistency of visual acuity and contrast sensitivity measures in advanced eye disease. *Optom Vis Sci.* 2005;82(11):946-954.

45. CATT: comparison of age-related macular degeneration treatment trials. ETDRS chart worksheet. Philadelphia (PA): University of Pennsylvania; 2014.

46. Beck RW, Moke PS, Turpin AH, et al. A computerized method of visual acuity testing: adaptation of the early treatment of diabetic retinopathy study testing protocol. *Am J Ophthalmol.* 2003;135(2):194-205.

47. Recommended stardard procedures for the clinical measurement and specification of visual acuity. Report of working group 39. Committee on vision. Assembly of Behavioral and Social Sciences, National Research Council, National Academy of Sciences, Washington, D.C. *Adv Ophthalmol.* 1980;41:103-148.

48. Kniestedt C, Stamper RL. Visual acuity and its measurement. *Ophthalmol Clin North Am.* 2003;16(2):155-170.

49. Ranibizumab (Lucentis): visual impairment due to choroidal neovascularization secondary to pathologic myopia. Appendix 4, validity of outcome measures. Ottawa (ON): CADTH; 2015: https://www.ncbi.nlm.nih.gov/books/NBK349531/. Accessed 2020 Aug 20.

50. Holladay JT. Visual acuity measurements. *J Cataract Refract Surg.* 2004;30(2):287-290.

51. Lange C, Feltgen N, Junker B, Schulze-Bonsel K, Bach M. Resolving the clinical acuity categories "hand motion" and "counting fingers" using the Freiburg Visual Acuity Test (FrACT). *Graefes Arch Clin Exp Ophthalmol.* 2009;247(1):137-142.

52. Aflibercept (Eylea): treatment of neovascular (Wet) Age-Related Macular Degeneration (wAMD). Appendix 5, validity of outcome measures. Ottawa (ON): CADTH; 2015: https://www.ncbi.nlm.nih.gov/books/NBK349601/. Accessed 2020 Aug 20.

53. National Institute for Health and Care Excellence. Voretigene neparvovec for treating inherited retinal dystrophies caused by RPE65 gene mutations. *(Highly specialised technologies guidance HST11)* 2019; https://www.nice.org.uk/guidance/hst11. Accessed 2020 Jun 23.

54. Mangione CM. The National Eye Institute 25-item Visual Function Questionnaire (VFQ-25). Bethesda (MD): National Eye Institute; 2000: https://www.nei.nih.gov/sites/default/files/2019-06/manual_cm2000.pdf. Accessed 2020 Aug 20.

55. Revicki DA, Rentz AM, Harnam N, Thomas VS, Lanzetta P. Reliability and validity of the National Eye Institute Visual Function Questionnaire-25 in patients with age-related macular degeneration. *Invest Ophthalmol Vis Sci.* 2010;51(2):712-717.

56. Mangione CM, Lee PP, Gutierrez PR, et al. Development of the 25-list-item National Eye Institute visual function questionnaire. *Arch Ophthalmol.* 2001;119(7):1050-1058.

57. Suner IJ, Kokame GT, Yu E, Ward J, Dolan C, Bressler NM. Responsiveness of NEI VFQ-25 to changes in visual acuity in neovascular AMD: validation studies from two phase 3 clinical trials. *Invest Ophthalmol Vis Sci.* 2009;50(8):3629-3635.

58. Orr P, Rentz AM, Margolis MK, et al. Validation of the National Eye Institute Visual Function Questionnaire-25 (NEI VFQ-25) in age-related macular degeneration. *Invest Ophthalmol Vis Sci.* 2011;52(6):3354-3359.

59. Hirneiss C. The impact of a better-seeing eye and a worse-seeing eye on vision-related quality of life. *Clin Ophthalmol.* 2014;8:1703-1709.

60. Maguire AM, Simonelli F, Pierce EA, et al. Safety and efficacy of gene transfer for Leber's congenital amaurosis. *N Engl J Med.* 2008;358(21):2240-2248.

61. Bennett J, Ashtari M, Wellman J, et al. AAV2 gene therapy readministration in three adults with congenital blindness. *Sci Transl Med.* 2012;4(120):120ra115.

62. Maguire AM, High KA, Auricchio A, et al. Age-dependent effects of RPE65 gene therapy for Leber's congenital amaurosis: a phase 1 dose-escalation trial. *Lancet.* 2009;374(9701):1597-1605.

63. Simonelli F, Maguire AM, Testa F, et al. Gene therapy for Leber's congenital amaurosis is safe and effective through 1.5 years after vector administration. *Mol Ther.* 2010;18(3):643-650.

64. Testa F, Maguire AM, Rossi S, et al. Three-year follow-up after unilateral subretinal delivery of adeno-associated virus in patients with Leber congenital amaurosis type 2. *Ophthalmology.* 2013;120(6):1283-1291.

65. Bennett J, Wellman J, Marshall KA, et al. Safety and durability of effect of contralateral-eye administration of AAV2 gene therapy in patients with childhood-onset blindness caused by RPE65 mutations: a follow-on phase 1 trial. *Lancet.* 2016;388(10045):661-672.

66. Clinical study report amendment 002, EU: AAV2-hRPE65v2-101. A phase 1 safety study in subjects with Leber Congenital Amaurosis (LCA) using adeno-associated viral vector to deliver the gene for human RPE65 into the Retinal Pigment Epithelium (RPE) [**CONFIDENTIAL** internal sponsor's report]. Philadelphia (PA): Spark Therapeutics; 2017 Jun 8.

67. Clinical study report addendum 2015, amendment 001, EU: AAV2-hRPE65v2-102. A follow-on study to evaluate the safety of readministration of adeno-associated viral vector containing the gene for human RPE65 [AAV2-hRPE65v2] to the contralateral eye in subjects with Leber Congenital Amaurosis (LCA) previously enrolled in a phase 1 study [**CONFIDENTIAL** internal sponsor's report]. Philadelphia (PA): Spark Therapeutics; 2017 Jun 9.

68. Koenekoop RK, Lopez I, den Hollander AI, Allikmets R, Cremers FP. Genetic testing for retinal dystrophies and dysfunctions: benefits, dilemmas and solutions. *Clin Experiment Ophthalmol.* 2007;35(5):473-485.

69. Pierce EA, Bennett J. The status of RPE65 gene therapy trials: safety and efficacy. *Cold Spring Harb Perspect Med.* 2015;5(9):a017285.

70. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174.

71. Reeve BB, Wyrwich KW, Wu AW, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res.* 2013;22(8):1889-1905.

72. Cohen J. A power primer. *Psychol Bull.* 1992;112(1):155-159.

73. Center for Drug Evaluation Research. FDA advisory committee briefing document. *Spark Therapeutics, Inc Luxturna (voretigene neparvovec)*. Rockville (MD): U. S. Food and Drug Administration (FDA); 2017: https://www.fda.gov/media/108385/download. Accessed 2020 Aug 21.

74. Roman AJ, Cideciyan AV, Aleman TS, Jacobson SG. Full-field stimulus testing (FST) to quantify visual perception in severely blind candidates for treatment trials. *Physiol Meas.* 2007;28(8):N51-56.

75. Beck RW, Moke PS, Turpin AH, et al. A computerized method of visual acuity testing: adaptation of the early treatment of diabetic retinopathy study testing protocol. *Am J Ophthalmol.* 2003;135(2):194-205.

76. Bittner AK, Ferraz MC. Reliability of mesopic measures of visual acuity and contrast sensitivity and their correlation with rod and cone function in retinitis pigmentosa. *Ophthalmic Res.* 2020;63(2):133-140.

77. Beck RW, Maguire MG, Bressler NM, Glassman AR, Lindblad AS, Ferris FL. Visual acuity as an outcome measure in clinical trials of retinal diseases. *Ophthalmology.* 2007;114(10):1804-1809.

78. Langelaan M, Wouters B, Moll AC, de Boer MR, van Rens GHMB. Functional field score: the effect of using a Goldmann V-4e isopter instead of a Goldmann III-4e isopter. *Invest Ophthalmol Vis Sci.* 2006;47(5):1817-1823.

79. Barnes CS, Schuchard RA, Birch DG, et al. Reliability of Semiautomated Kinetic Perimetry (SKP) and Goldmann kinetic perimetry in children and adults with retinal dystrophies. *Transl Vis Sci Technol.* 2019;8(3):36.

80. Dedania VS, Liu JY, Schlegel D, et al. Reliability of kinetic visual field testing in children with mutation-proven retinal dystrophies: implications for therapeutic clinical trials. *Ophthalmic Genet.* 2018;39(1):22-28.

81. Patel DE, Cumberland PM, Walters BC, Russell-Eggitt I, Rahi JS. Study of Optimal Perimetric Testing in Children (OPTIC): feasibility, reliability and repeatability of perimetry in children. *PLoS One.* 2015;10(6):e0130895.

82. Ridge KD, Palczewski K. Visual rhodopsin sees the light: structure and mechanism of G protein signaling. *J Biol Chem.* 2007;282(13):9297-9301.

83. Travis GH, Golczak M, Moise AR, Palczewski K. Diseases caused by defects in the visual cycle: retinoids as potential therapeutic agents. *Annu Rev Pharmacol Toxicol.* 2007;47:469-512.

84. National Center for Biotechnology Information. Genetic testing registry. https://www.ncbi.nlm.nih.gov/gtr/. Accessed 2020 Aug 21.

85. National Institutes of Health. What do the results of genetic tests mean? 2020; https://ghr.nlm.nih.gov/primer/testing/interpretingresults. Accessed 2020 Aug 21.