



# Methodological Challenges in the Economic Evaluation of a Gene Therapy for RPE65-Mediated Inherited Retinal Disease: The Value of Vision

Simone A. Huygens<sup>1</sup> · Matthijs M. Versteegh<sup>1</sup> · Stefan Vegter<sup>2</sup> · L. Jan Schouten<sup>3</sup> · Tim A. Kanters<sup>1</sup>

Accepted: 23 January 2021 / Published online: 19 February 2021  
© The Author(s) 2021

## Abstract

The emergence of gene therapies challenge health economists to evaluate interventions that are often provided to a small patient population with a specific gene mutation in a single dose with high upfront costs and uncertain long-term benefits. The objective of this study was to illustrate the methodological challenges of evaluating gene therapies and their implications by discussing four economic evaluations of voretigene neparvovec (VN) for the treatment of RPE65-mediated inherited retinal disease. The checklist for economic evaluations of gene therapies of Drummond et al. was applied to the economic evaluations of VN performed by US Institute for Clinical and Economic Review, two country adaptations of the company model in the UK and the Netherlands, and another US publication. The main differences in methodological choices and their impact on cost-effectiveness results were assessed and further explored with sensitivity analyses using the Dutch model. To enable comparison between the economic evaluations, costs were converted to US dollars. Different methodological choices were made in the economic evaluations of VN resulting in large differences in the incremental cost-effectiveness ratio varying from US\$79,618 to US\$643,813 per QALY. The chosen duration of treatment effect, source of utility values, discount rate and model structure had the largest impact on the cost-effectiveness. This study underlines the findings from Drummond et al. that standard methods can be used to evaluate gene therapies. However, given uncertainty about (particularly long-term) outcomes of gene therapies, guidance is required on the acceptable extrapolation of treatment effect of gene therapies and on how to handle the uncertainty around this extrapolation in scenario and sensitivity analyses to aid health technology assessment research and align submissions of future gene therapies.

## 1 Introduction

Gene therapies have long been a promise, potentially bringing durable or even curative clinical benefit. Recently, several therapies have started delivering on this promise, although long-term effectiveness remains unclear [1]. Currently available gene therapies are targeted to small patient populations with a specific gene mutation and provided in

one single dose with high upfront costs, while the benefits theoretically should last a long time (sometimes even a lifetime). These characteristics have led to a discussion among health economists about whether a specific methodological reference case is required for economic evaluation of gene therapies [2]. Drummond et al. argued that a new methodological reference case for economic evaluation of gene therapies is not required because their characteristics are not unique to gene therapies [2]. However, they emphasized that “the confluence of various characteristics can lead to specific methodological challenges when evaluating the cost-effectiveness of a gene therapy” [2]. Therefore, they proposed a checklist for economic evaluations of gene therapies to inform audiences whether and to what extent key elements affecting gene therapies have been identified and considered in the analyses [2].

✉ Simone A. Huygens  
huygens@imta.eur.nl

<sup>1</sup> Institute for Medical Technology Assessment (iMTA),  
Erasmus University Rotterdam, Rotterdam, The Netherlands

<sup>2</sup> Novartis Pharma B.V., Amsterdam, The Netherlands

<sup>3</sup> Canisius Wilhelmina Hospital, Nijmegen, The Netherlands

### Key Points

Different methodological choices were made in the economic evaluations of a specific gene therapy (i.e. voretigene neparvovec) resulting in large differences in cost-effectiveness results.

Standard economic evaluation methods can be used to evaluate gene therapies.

Given uncertainty about (particularly long-term) outcomes of gene therapies, guidance is required on the acceptable extrapolation of treatment effect of gene therapies and on how to handle the uncertainty around this extrapolation in scenario and sensitivity analyses to aid health technology assessment research and align submissions of future gene therapies.

The objective of this study was to illustrate the methodological challenges when evaluating a gene therapy and the implications on cost-effectiveness results by discussing the most recent gene therapy for which economic evaluations have been conducted: voretigene neparvovec (VN) for the treatment of RPE65-mediated inherited retinal dystrophy (IRD). VN (brand name Luxturna) is one of the first gene therapies approved by the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA) (in 2017 and 2018, respectively) [3, 4]. VN is an adeno-associated viral vector-based gene therapy for the treatment of patients with vision loss due to IRD caused by mutations in the *RPE65* gene [3, 5, 6]. During a surgical procedure called vitrectomy, a healthy copy of the defective *RPE65* gene is introduced into the RPE cells in the retina. The healthy copy of the gene enables patients to produce functional RPE65 protein, which may prevent further deterioration of the patient's visual abilities that would eventually lead to blindness in untreated patients. Four of the first economic evaluations of VN were selected for review in this study. First, the economic evaluation of the US Institute for Clinical and Economic Review [7, 8], who were the first to evaluate the cost effectiveness of VN compared with best supportive care. They reported an incremental cost-effectiveness ratio (ICER) of US\$643,813 per QALY from the US health care system perspective [7]. Since then, the marketing authorization holder of VN has submitted health technology assessment (HTA) dossiers to national HTA agencies in several countries using a commissioned model (from now on referred to as the 'company model'). Two of the first HTA submissions of VN submitted in 2019 (to the 'Zorginstituut Nederland' [ZIN] in the Netherlands and

to the 'National Institute for Care and Health Excellence' [NICE] in the United Kingdom) will also be reviewed in this study [9–11]. Using the list price of VN, the ICER in the UK HTA submission from an NHS perspective was £86,635/QALY (US\$115,513/QALY) and in the Dutch HTA submission the ICER from a societal perspective was €80,994/QALY (US\$95,625/QALY). Finally, an economic evaluation applying a similar model as the company model to the United States' setting by Johnson et al. (2019) was included in this study [12]. Johnson and colleagues' results were vastly different from the US Institute for Clinical and Economic Review results; they found an ICER of US\$79,618/QALY when applying a healthcare payer perspective. These examples show that there can be substantial differences in the outcomes of economic evaluations of the same gene therapy.

In the following, Drummonds' checklist for assessing gene therapies will be applied to the four economic evaluations of VN discussed above [2]. Subsequently, the impact of methodological choices on cost-effectiveness outcomes will be explored. Finally, the necessity of special methods to assess the cost effectiveness of gene therapies will be discussed.

## 2 Methods

Four economic evaluations of VN were evaluated performed by US Institute for Clinical and Economic Review [7, 8], the HTA submission to NICE in the UK [9, 10], the HTA submission to ZIN in the Netherlands [11] and Johnson et al. [12]. The HTA submissions in the UK and the Netherlands both used the company model but made different methodological choices and are therefore reviewed separately. The cost-effectiveness model developed by US Institute for Clinical and Economic Review, the company model (used in the submission in the UK and the Netherlands) and the model used in Johnson et al. had comparable settings. All models used a lifetime time horizon, cycle length of 1 year and starting age of 15 in the base case. However, the US Institute for Clinical and Economic Review model structure differed from the company and Johnson et al. model. The US Institute for Clinical and Economic Review model used two health states (alive and dead). Within the alive state, visual acuity (VA, the ability of the eyes to discriminate detail or 'clarity of vision') and visual field (VF, the total area in which objects can be seen while focusing the eyes on a central point) were modelled using an exponential and linear functional form, respectively. Efficacy of VN was based on the change in VA and VF observed in the clinical trial and extrapolated over 10 years followed by a 10-year waning period (Table 1) [5]. In contrast, the company and Johnson et al. models used VA

and VF observed in the clinical trial to populate a Markov model with five health states (Fig. 1). The extrapolation of the treatment effect observed in the clinical trial varied between 20 and 70 years in these economic evaluations (Table 1). Furthermore, different list prices, assumptions for the duration and waning of treatment effect and sources for utilities were used in the four economic evaluations. The model characteristics and settings of the four studies are summarized in Table 1. Full descriptions of the economic evaluations are provided elsewhere [7–9, 11, 12]. The HTA submissions in the UK and the Netherlands were sponsored by Novartis Pharma and the study of Johnson et al. was sponsored by Spark Therapeutics, the economic evaluation of US Institute for Clinical and Economic Review did not have external sponsors.

## 2.1 Drummond's Checklist

The checklist for assessing the cost-effectiveness of gene therapies of Drummond et al. consists of three sections related to clinical effectiveness, elements of value and other considerations, respectively [2]. For every checklist item, the approach in the four economic evaluations of VN was described (Table 2).

## 2.2 Sensitivity and Scenario Analyses

To illustrate the impact of some methodological choices, additional analyses were performed using the Dutch adaptation of the company model performed for the HTA submission to ZIN (from now on referred to as the 'Dutch model'; a technical report of this model is available from the corresponding author upon request). These results are expressed in 2018 Euros.

## 2.3 Comparison of Cost-Effectiveness Results

To enable comparison between the results of the four economic evaluations, costs of the base-case ICER from a healthcare and societal perspective were converted to US dollars using the 2018 exchange rates [13].

## 3 Results

Table 2 shows the completed checklist for assessing gene therapies [2] for the four economic evaluations performed for VN. Individual items that varied between the four economic evaluations or required more explanation than provided in Table 2 are discussed in more detail below.

## 3.1 Drummond's Checklist

### 3.1.1 Clinical Effectiveness

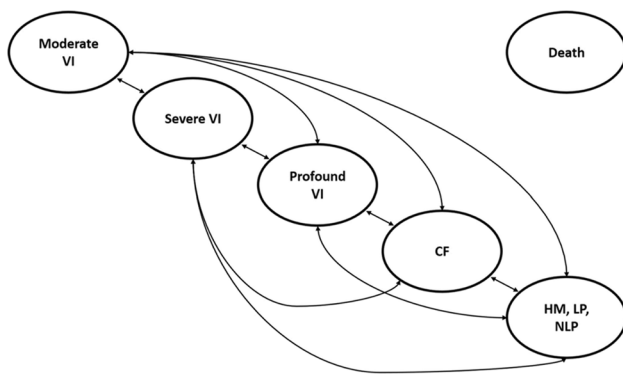
**3.1.1.1 Surrogate Endpoint Used** The primary outcome in the phase III randomized controlled trial (RCT) was change in performance on a standardized multi-luminance mobility test (MLMT) at 1 year relative to baseline [5]. The MLMT quantifies the patient's ability to navigate around a course of obstacles in varying environmental illuminations, including very low light levels, integrating aspects of VA, VF and light sensitivity [5]. Preferably, the model structure would be based on this primary outcome. However, due to the absence of data linking the MLMT to costs and utilities and long-term data on changes in MLMT, the health states in all four models were defined using VA and VF instead of MLMT. VA and VF were considered to be the best available alternative to MLMT, considering that MLMT is a functional endpoint that also captures changes in VA and VF [14, 15].

**3.1.1.2 Rare Disease** One of the characteristics of gene therapies that lead to methodological challenges in economic evaluations is that they often target rare conditions [2]. This not only raises challenges in the assessment of clinical effectiveness and safety (discussed later in Sect. 3.1.1.5), but also in collecting patient-reported health-related quality of life (HRQoL) data. Due to the rarity of RPE65-mediated IRD, patient-reported data of HRQoL in the target population of VN was not available and the feasibility of collecting this data outside the clinical trial setting was limited. As a consequence, alternative sources were explored in the four economic evaluations. Utilities in the US Institute for Clinical and Economic Review model were based on direct utility assessment in other retinal disease populations published by Brown et al. [16, 17], while the other three economic evaluations used utilities derived from clinical experts who completed EQ-5D-5L (Johnson et al. and Dutch HTA submission) or HUI-3 (health utility index, UK HTA submission) from the perspective of patients described in vignettes (Fig. 2) [18]. Both sources have limitations: the utility values in the US Institute for Clinical and Economic Review study were derived from a time trade-off (TTO) exercise in patients of a population that was not representative for patients with RPE65-mediated IRD (i.e. diabetic retinopathy patients, who are older and have diabetes-related comorbidities), while the utility values in the other three economic evaluations were based on proxy utility assessment by six US clinical experts. Figure 2 illustrates the range in utility val-

**Table 1** Model characteristics and settings

Item	US ICER	US Johnson et al.	UK	The Netherlands
Model structure	One alive state with varying visual impairment based on exponential function and one death state	Five alive states with varying visual impairment and one death state	Five alive states with varying visual impairment and one death state	Five alive states with varying visual impairment and one death state
Health states based on average or best-seeing eye	VA in 'best-seeing eye' and VF in 'average eye'	Worst of VA and VF in 'average eye'	Worst of VA and VF in 'average eye'	Worst of VA and VF in 'best-seeing eye'
Effectiveness in first years after VN treatment based on observed data	Adjusted change in VA and VF based on RCT (original intervention arm only)	Transitions in intervention and control arm based on RCT (original intervention arm only)	Transitions in intervention and control arm based on RCT (original intervention arm only)	Transitions in intervention and control arm based on RCT (original intervention arm and cross-over from control arm)
Long-term effectiveness based on extrapolation of observed data	10 years, followed by 10-year waning period from full treatment effect to natural history (i.e. no treatment effect)	Lifetime, i.e. > 70 years	40 years, followed by natural history (i.e. no treatment effect)	20 years, followed by natural history (i.e. no treatment effect)
Health-related quality-of-life source	TTO in patients from a different patient population (older patients with diabetic retinopathy) [14]	Proxy-based indirect utility assessment (EQ-5D-5L filled out by rehabilitation experts [n = 6] based on patient descriptions and scored using the van Hout et al. algorithm) [15, 16]	Proxy-based indirect utility assessment (HUI-3 filled out by rehabilitation experts [n = 6] based on patient descriptions)[15]	Proxy-based indirect utility assessment (EQ-5D-5L filled out by rehabilitation experts (n = 6) based on patient descriptions and scored using the Dutch EQ-5D-5L tariff) [15, 17]
Discounting	3.0% costs and effects	3.0% costs and effects	3.5% costs and effects	4.0% costs and 1.5% effects
Cycle length	1 year	1 year	1 year	1 year
Time horizon	Lifetime	Lifetime	Lifetime	Lifetime
List price of VN	US\$850,000	US\$850,000	US\$817,880 (€613,410)	US\$814,640 (€690,000)

*HUI* Health Utility Index, *TTO* time trade-off, *UK* United Kingdom, *US* United States, *US ICER* United States Institute for Cost-Effectiveness Research, *VA* visual acuity, *VF* visual field, *VN* voretigene neparovect



**Fig. 1** Structure of the company model [12]. Patients can transition to death from every health state. *VI* visual impairment, *CF* counting fingers, *HM* hand motion, *LP* light perception, *NLP* no light perception

ues for the different sources. The utility values used in the US Institute for Clinical and Economic Review model with one health state with varying visual impairment based on an exponential function were translated to the five health states used in the other models. The range in utility values between the best and worst health state is smaller in the US Institute for Clinical and Economic Review economic evaluation than in the other economic evaluations (Fig. 2), leading to smaller potential health gains and explaining the relatively high ICERs. Application of the utility source used in the US Institute for Clinical and Economic Review model in the Dutch model resulted in a 58% higher ICER (Table 3).

In the vignette study [18], HRQoL was measured with two questionnaires: HUI-3 and EQ-5D-5L. Use of the EQ-5D-5L for utilities is requested by health economic guidelines in the Netherlands and was chosen as the base case in Johnson et al. and the Dutch HTA submission. However, in contrast to the EQ-5D-5L, the HUI-3 does include a vision component. The HUI-3 showed better performance in patients with vision disorders than EQ-5D-3L, and the majority of studies on the convergent validity of the HUI-3 with clinical measures demonstrated good validity [19, 20]. The company base case in the UK HTA submission therefore used utility values based on the HUI-3. The impact of using EQ-5D-5L instead of the HUI-3 utilities from the vignette study is shown by Viriato et al. [21], who performed comparable analyses with the company model for the UK setting, as in the NICE submission, resulting in an ICER of £95,072/QALY (compared with £86,635/QALY calculated with the HUI-3 in the NICE submission). The UK ICER was higher in the NICE submission due to larger utility decrements for HS2 compared with HS1 (the most important transition according to the one-way sensitivity analyses [21] is from HS1 to HS2) when using the HUI-3 (− 0.16, Fig. 2)

compared with the UK EQ-5D-5L (− 0.09, Fig. 2). In contrast, the Dutch ICER was higher if utility values were based on the Dutch EQ-5D-5L tariff compared with the HUI-3 (Table 3) because of the smaller range in utility values of HS1 to HS5 when using the HUI-3 (0.56) instead of the Dutch EQ-5D-5L (0.60). Both NICE and ZIN considered the lack of patient-reported utility values for patients treated with VN to be a key limitation of the cost-effectiveness model [9, 11]. As utility values were not captured in the clinical trial, only arguments, not data, were available to identify an alternative source of utility values. The different researchers made different choices based on different arguments and these choices vastly impacted the ICERs.

**3.1.1.3 Single-Arm Trial** Often the effectiveness of gene therapies is based on single-arm trials. However, in the case of VN, a RCT was executed [5]. All four economic evaluations based the clinical efficacy of VN on this RCT. Therefore, the limitations of single-arm trials do not play a role in these economic evaluations.

**3.1.1.4 Reporting of Adverse Consequences and Risks** Johnson et al. did not include any adverse events because they considered that VN was well tolerated and adverse events were mostly mild to moderate in severity [12]. The other models used the RCT for the incidence of short-term adverse events after treatment with VN, but applied different inclusion criteria resulting in different adverse events considered in the economic evaluations. However, it is not likely that this explains large differences in the cost-effectiveness outcomes, due to the low healthcare costs and small impact on HRQoL of the adverse events. In fact, in a scenario analyses where short-term adverse events are excluded from the Dutch model, the ICER only decreases with 0.2% (Table 3).

The administration procedure of VN (i.e. vitrectomy) is associated with a risk of cataract [22]. In the UK and Dutch HTA submission, all patients with cataract were assumed to undergo cataract surgery. In addition to the short-term risk of cataract, cataract is commonly reported within 2 years following vitrectomy with the incidence rate varying from 12.5 to 80% [22]. These cataracts, and subsequent impact on utilities and costs, were not included in any of the economic evaluations because of lack of long-term data on the occurrence of cataracts. However, considering the small impact on utility (disutility of 0.14 during 1 month) and the relative low costs of cataract surgery (£913 in the UK, €1062 in the Netherlands [9, 23]) compared with the acquisition costs of VN, and that only a proportion of patients (12.5–80%) will experience this adverse event, the impact on cost-effectiveness results is expected to be small. Other long-term adverse consequences and risks are unknown due to the limited follow-up period of patients and could therefore

**Table 2** Drummond et al. checklist for assessing gene therapies completed for four economic evaluations of VN

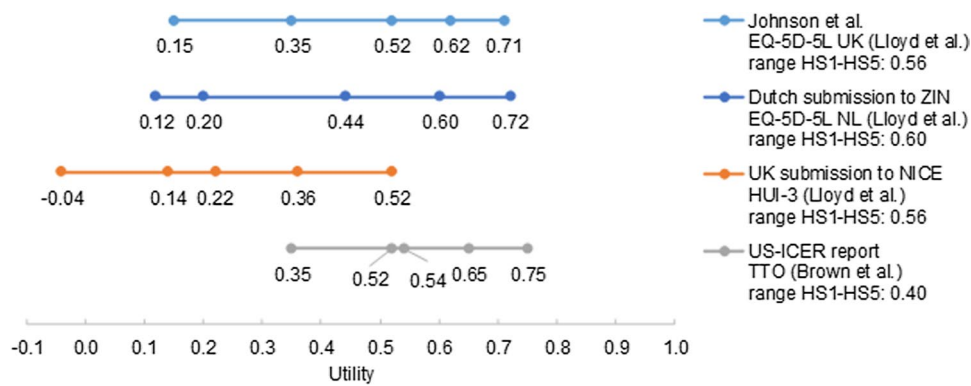
Item	US-ICER [7, 8]	Johnson et al. [12] (US)	UK [9]	The Netherlands [11]
Clinical effectiveness				
Surrogate endpoint used	VA and VF (secondary endpoints in phase III clinical trial) instead of multi-luminance mobility test (primary endpoint in RCT)			
Rare disease (number of eligible patients)	1000–3000 [8]	1000–2000 [12]	86 [9]	45 [11]
Serious condition	Severe visual impairment throughout childhood with deteriorating vision over time resulting in complete blindness by 30–40 years of age			
Single-arm trial	RCT with intervention and control arm			
Paediatric population	Age in RCT: mean 15, median 11, interquartile range 6–20 years			
Reporting of adverse consequences and risks	Only short-term adverse events included	No adverse events included	Only short-term adverse events included	Only short-term adverse events included
Size of clinical trial	Intervention arm $n = 29$ ; control arm $n = 9$			
Length of clinical trial	1-year data used in model; follow-up data until 3–4 years available			
Extrapolation to long-term outcomes, number of years extrapolation of treatment effect	10 years, followed by 10-year waning period from full treatment effect to natural history (i.e. no treatment effect)	Lifetime, i.e. > 70 years	40 years, followed by natural history (i.e. no treatment effect)	20 years, followed by natural history (i.e. no treatment effect)
Elements of value				
Severe disease	Not considered			Considered in cost-per-QALY threshold
Value to caregivers	Not considered			Considered in a scenario analysis
Insurance value	Not considered			
Scientific spillovers	Not considered			
Lack of alternatives	Before the introduction of VN, there were no interventions that alter the natural history of RPE65-mediated IRD. Patients were treated with BSC			
Substantial improvement in life expectancy	RPE65-mediated IRD does not affect mortality risk and therefore treatment with VN does not improve life expectancy directly			
Other considerations				
Discounting				
Different discount rates explored	Not varied in scenario analysis	Not varied in scenario analysis	1.5% and 0% in scenario analyses	0%, 1.5% and 5% for costs and outcomes in scenario analyses
Uncertainty				
Alternative payment models explored	Threshold analysis to estimate the maximum price of VN in order to be cost effective	Not considered	Confidential simple discount PAS	Pay for performance agreement ('no cure, no pay') considered in a scenario analysis

BSC best supportive care, IRD inherited retinal dystrophy, PAS patient access scheme, QALY quality-adjusted life-year, RCT randomized controlled trial, UK United Kingdom, US United States, US ICER United States Institute for Cost-Effectiveness Research, VA visual acuity, VF visual field, VN voretigene neparvovec

not be included in the economic evaluations until data from post-authorization safety studies are available.

**3.1.1.5 Size of the Trial** Although the efficacy of VN in the model could be based on an RCT, the rarity of RPE65-mediated IRD resulted in efficacy data that was based on a small number of patients (intervention  $n = 29$  including cross-over patients; control  $n = 9$ ) [5]. As a result, there is a lot of uncertainty around the transition probabilities based on this RCT. In particular, the number of patients in the control group

who received BSC is very low. The data from the RCT was not used to model the disease progression of BSC patients in the US Institute for Clinical and Economic Review model, instead they used natural history data [24]. In contrast, the other models used the RCT data of the first year to model disease progression in BSC patients. However, because this data is only used in the first year of patients receiving BSC (i.e. after the first year the transition probabilities are based on the natural history data in the BSC arm), the influence on the cost-effectiveness outcomes is limited. In a scenario



**Fig. 2** Utility values ranging from moderate visual impairment to hand motion/no light perception based on different sources (Lloyd et al. [18], Brown et al. [17]). The utility values used in the US-ICER model with one health state with varying visual impairment based on an exponential function were translated to the five health states used

in the other models. *HS* health state, *HUI* health utility index, *NICE* National Institute for Care and Excellence, *TTO* time trade-off, *UK* United Kingdom, *US ICER* United States Institute for Cost-Effectiveness Research, *ZIN* Zorginstituut Nederland

**Table 3** Scenario analyses using Dutch adaptation of company model

Scenario	Incremental costs (€)	Incremental QALYs	ICER (€)	% Change from base-case ICER
Base case (according to ‘The Netherlands’ model specification in Table 2)	672,736	8.31	80,994	
Utilities				
Utility values: Brown et al. [17] (used in US-ICER)	672,736	5.24	128,368	+ 58%
Utility values: Lloyd et al. [18] HUI-3 values	672,736	7.36	91,445	+ 13%
Reporting of adverse consequences and risks				
Exclusion of adverse events	672,551	8.32	80,836	0%
Size of the trial				
BSC effectiveness from natural history data	669,360	7.99	83,779	3%
Extrapolation to long-term outcomes				
Duration of treatment effect: 10 years +10 years waning effect (comparable to US-ICER assumptions)	680,045	5.88	115,673	+ 43%
Duration of treatment effect: 40 years (NICE submission assumption)	662,891	12.21	54,310	- 33%
Duration of treatment effect: 70 years (life-long; Johnson et al. assumption)	658,708	13.94	47,239	- 42%
Value to caregivers				
Include caregiver disutility	672,736	8.52	78,974	- 2%
Improvements in life expectancy				
Include excess mortality risk	672,791	8.27	81,364	0%
Discounting				
4% discount rate for costs and outcomes	672,736	4.60	146,395	+ 81%
List price VN US				
List price VN equal to US	702,686	8.31	84,600	+ 4%

*BSC* best supportive care, *ICER* incremental cost-effectiveness ratio, *NICE* National Institute for Health and Care Excellence, *QALY* quality adjusted life year, *US* United States, *US ICER* United States Institute for Cost-Effectiveness Research, *VN* voretigene neparovvec

using the Dutch model where the transition probabilities of patients receiving BSC were based on the natural history data [24] from baseline instead of after year 1, the ICER increased with 3% (Table 3).

The influence of uncertainty in efficacy data on cost-effectiveness results was also estimated in probabilistic sensitivity analyses (PSA) using the Dutch model. Because of

the small number of patients included in the RCT, there were no data available on transitions in some of the health states. To overcome the potential problem of zero observed counts in some of the cells of the matrices, a fixed value (referred to as ‘prior’) was added to all cells of the transition matrices. In the Netherlands, a prior of 1 was applied as recommended by Briggs et al. [25], while in the UK submission a prior of

0.1 was used. Our results showed that the chosen prior introduced uncertainty. Depending on the prior, the average ICER of 10,000 PSA iterations varied from €95,970 with a prior of 0.1 to €108,752 per QALY gained with a prior of 1 (Fig. 3).

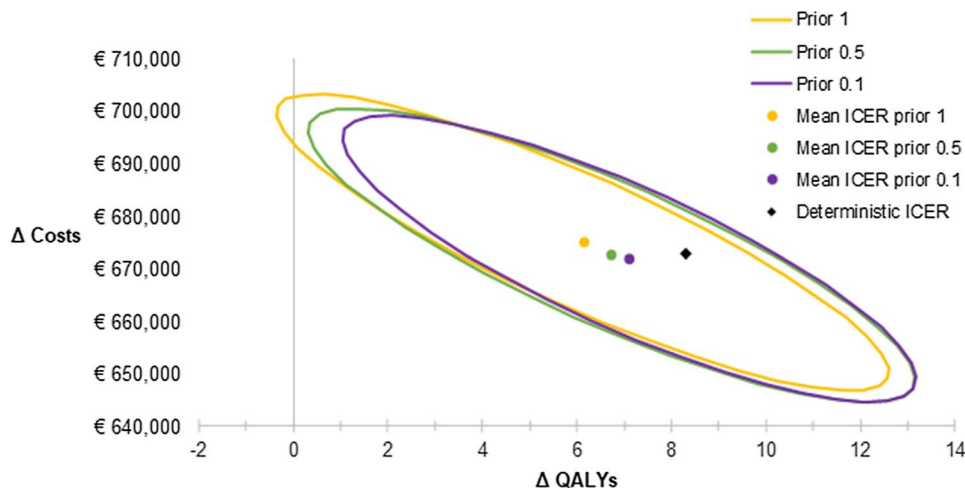
**3.1.1.6 Length of Clinical Trial** The efficacy data in the health economic model was based on the 1- or 2-year results of the RCT in all four economic evaluations [5]. At the time of the economic evaluations performed with the company model, 3–4 years follow-up data was available for the RCTs (Study 301/302) [26]. However, the available data from years 2 and 3 was not used as the modelers argued that the distribution of patients across health states would be essentially the same as the assumption of stabilization beyond year 1 because significant changes in outcomes beyond year 1 were not observed. In any case, the follow-up period is relatively short for the potential promise of lifetime duration of treatment effect of VN. Ideally, longer follow-up data would have been included in the model but waiting on additional long-term data would postpone the availability of treatment for patients for several years. Instead, the results of the trial were extrapolated to long-term outcomes.

**3.1.1.7 Extrapolation to Long-Term Outcomes** The efficacy data was based on the 1-year follow-up of patients in the RCT in all four models [5]. Due to the short timeframe of the clinical trial, an important part of the value hypothesis of VN (i.e. long-term benefit) is not substantiated by trial data [27]. However, the current data and biological knowledge on the (patho)physiology of the targeted cells supported the suggestion of maintained improvement of visual function after injection with VN beyond what has been observed in the RCT. Current available data shows that the majority of patients (86%) included in the RCT showed stable functional vision 1 year post-treatment through 3–4 years of follow-up [28]. Furthermore, 7.5- and 4-year follow-up data of the phase I clinical trial for the first and second

injected eye, respectively, showed sustained improvements on full-field light sensitivity threshold (FST) in the majority of the patients [5, 15, 29]. These FST improvements are relevant because a post-hoc analysis of the RCT described a good correlation between MLMT performance and FST [15]. Furthermore, it is anticipated that the *RPE65* gene will remain active during the lifetime of RPE cells, which normally undergo minimal proliferation (i.e. cell division) throughout life [21].

The four economic evaluations applied different assumptions on the duration of treatment effect. The treatment effect in the US Institute for Clinical and Economic Review model was assumed to be sustained for 10 years after which they modelled a waning period of 10 years in which the effect slowly decreased until the rate of decline in vision was the same as with BSC [7, 8]. A treatment effect duration of 20 years was assumed in the Dutch HTA submission. In the base-case analysis of the submission to NICE, a 40-year treatment effect duration was assumed [9]. Finally, the treatment duration effect was extrapolated to the entire patient's lifetime (i.e.  $\pm 70$  years) in Johnson et al. [12]. Scenario analyses showed that this parameter has the largest impact on the cost-effectiveness outcomes in all four economic evaluations. In the US Institute for Clinical and Economic Review report, the ICER decreased from US\$643,813 to US\$384,624 per QALY when applying a lifetime duration of the treatment effect instead of the 10 years with a 10-year waning period [7]. The ICER (without indirect costs) in the US economic evaluation of Johnson et al. increased from US\$79,618 to US\$380,185 when the treatment effect was assumed to last for only 3 years instead of a lifetime [12]. In the UK HTA submission, the ICER increased by 25% if a 20-year treatment duration was assumed and decreased by 5% if the treatment duration was assumed to be 50 years instead of 40 years in the base-case analysis [9]. The impact of the assumed treatment duration using the Dutch model is illustrated in Fig. 4 (blue line) and Table 3, ranging from an

**Fig. 3** Cost-effectiveness ellipses and mean ICERs with varying priors. *ICER* incremental cost-effectiveness ratio, *QALY* quality-adjusted life-year





ICER of €80,994/QALY when a 20-year duration of treatment effect was assumed to €47,239/QALY when a lifetime duration of treatment effect was assumed.

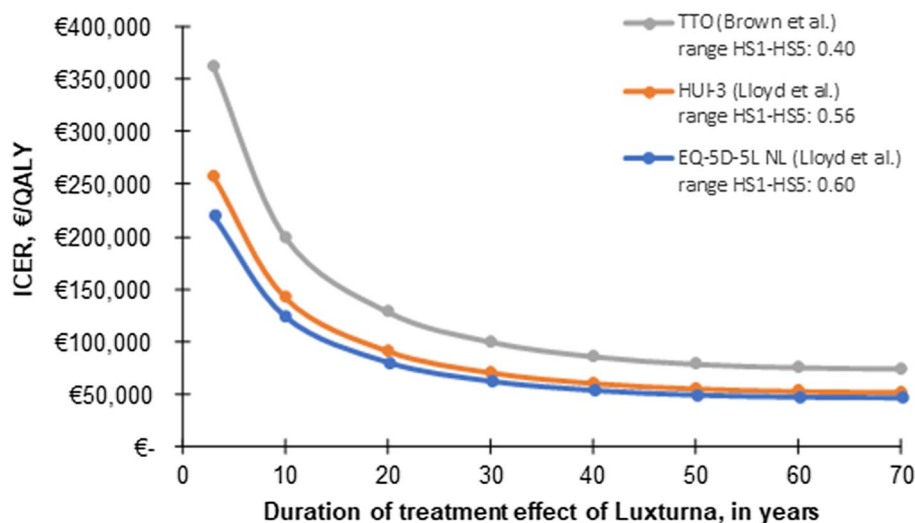
### 3.1.2 Elements of Value

**3.1.2.1 Severe Disease** Severity of disease is one of the additional elements of value beyond the QALY that may have the most support in the health economic field, but it has not been applied in many countries yet. The US Institute for Clinical and Economic Review, UK HTA submission and Johnson et al. did not consider the severity of disease in their analyses. In the Netherlands, however, the Dutch health economic guidelines state that the appropriate cost-per-QALY threshold depends on the severity of the disease expressed in proportional shortfall [30]. Proportional shortfall refers to QALYs lost due to disease as a proportion of quality-adjusted life expectancy of the age- and gender-matched general population. The severity-weighted thresholds are 20,000, 50,000 and 80,000 euro per QALY for proportional shortfalls (proportion of normal quality-adjusted life expectancy lost) under 41%, 71% and 100%, respectively [31]. Proportional shortfall is calculated with the iMTA disease burden calculator [32]. The proportional shortfall of RPE65-mediated IRD patients was estimated to be 69% based on the mean QALYs in BSC and VN derived from the deterministic analysis using the Dutch model, corresponding to a €50,000/QALY threshold. The proportional shortfall was also estimated for every PSA iteration using the Dutch model. The 95% confidence interval of these proportional shortfall estimates was 39–90%, indicating the high uncertainty around the proportional shortfall. Most of the PSA iterations (52%) using the Dutch model resulted in a proportional shortfall of 70% or higher corresponding to a cost-per-QALY threshold of €80,000/QALY. To consider the

uncertainty around the appropriate cost-per-QALY threshold, the severity-adjusted probability of being cost effective (SAPCE) [32] was calculated using the Dutch model. This procedure evaluates the ICER against the likelihood of each of the three severity-adjusted thresholds. SAPCE was 10%, compared with the probability of being cost effective of 0% when the ICER was evaluated against a single threshold of €50,000 per QALY or 19% when evaluated against a single threshold of €80,000/QALY. ZIN concluded that the severity of disease of RPE65-mediated IRD patients is difficult to estimate due to concerns about the validity of the utility values and uncertainty around the extrapolation of treatment effect of VN but is likely to be high due to the physiological impact of the disease [11].

**3.1.2.2 Value to Caregivers** The severity of RPE65-mediated IRD, just like many other diseases treated by gene therapy, means that the burden falling on caregivers is high. All four economic evaluations included caregiver costs in their base-case or scenario analyses, but only the NICE submission included the value of VN to caregivers on the effect side. Exclusion of the carer disutilities in a scenario analysis showed that the ICER increases by 9% if carer disutilities are not taken into account [9]. The US Institute for Clinical and Economic Review report states that VN may improve the independence of patients and thus reduce caregiver and family burden, but that there is no data available to substantiate this assumption at this time [7]. Besides limited data availability, there are methodological challenges to including the value to caregivers, such as the uncertainty about the proportion of informal caregivers who experience disutility, unknown number of informal caregivers or other family members or friends who experience disutility and unknown duration of disutility in the caregivers. In addition to the challenges of including disutilities of caregivers, it is debat-

**Fig. 4** Cost-effectiveness results for varying durations of treatment effect of VN and utility sources (Brown et al. [17] and Lloyd et al. [18]) using the Dutch health economic model. *HUI* health utility index, *HS* health state, *ICER* incremental cost-effectiveness ratio, *NL* the Netherlands, *QALY* quality-adjusted life-year, *TTO* time trade-off, *VN* voretigene neparovvec



able whether the standard cost-per-QALY threshold should be applied to outcomes of cost-effectiveness analyses including disutilities of caregivers. For example, a threshold based on an opportunity cost framework should take into account that displaced treatments also generate caregiver utilities.

Despite these challenges, an explorative scenario analysis was performed including the impact of carer disutility using the Dutch model to provide insight into the potential impact of considering the consequences of informal caregiving on the effect side. The results of the scenario analysis showed that including carer disutility resulted in a 2% lower ICER (Table 3).

**3.1.2.3 Lack of Alternatives** Before the introduction of VN, there were no interventions that could alter the natural history of RPE65-mediated IRD. Patients were treated with best supportive care, which may include correction of refractive error (i.e. glasses, contact lenses or refractive surgery), low-vision aids (e.g. strong magnifying reading glasses and loupes) and optimal access to educational and work-related opportunities. The lack of alternative treatment may result in a greater tendency by decision makers to accept high ICERs.

**3.1.2.4 Substantial Improvements in Life Expectancy** In all four economic evaluations, the base-case analysis assumed that RPE65-mediated IRD does not affect mortality risk and therefore treatment with VN does not improve life expectancy. In the company's base case in the submission to NICE, an increased risk of death was modelled based on the hazard ratios reported in Christ et al. [33]. However, this assumption was not included in the evidence review group's base case because no deaths occurred in any study included in VN's evidence base and the population in Christ et al. (i.e. patients aged 65–84 years) was substantially different from the target population of VN (i.e. patients aged 15 years at baseline) [9]. Including the excess mortality reported by Christ et al. (RR of 1.08 in health state 1 and 1.18 in health states 2–5) results in an increase of the ICER of 0.5% in the Dutch model (Table 3).

### 3.1.3 Other Considerations

**3.1.3.1 Discounting** Gene therapies often have large upfront costs because the therapy is delivered in a single dose, while the benefits stretch far into the future [34]. A lower discount rate would therefore be more favourable for gene therapies. However, Drummond et al. do not believe that deviating from the general methodological principle applied in most countries of discounting costs and benefits at the same rate is justified when evaluating gene therapies [2]. Nevertheless, they recommend that different discount rates for costs and benefits should be explored in scenario

analyses [2]. However, the discount rate was only varied in scenario analyses in two of the four economic evaluations reviewed in this study: the UK and Dutch HTA submissions. NICE stated that discount rates of 1.5% may be appropriate to consider because clinical evidence suggests benefits of VN may extend beyond 30 years and VN requires the NHS to commit significant, irrecoverable costs as a 'one-off' gene therapy. The ICER in this scenario analysis was 43% lower than the base-case ICER. ZIN also specifically requested additional scenario analyses with other discount rates. In the Netherlands, differential discounting for costs (4%) and effects (1.5%) is prescribed by the health economic guidelines [30]. A scenario analysis using the Dutch model in which the discount factor of effects was increased to 4% (i.e. equal to costs) showed that the impact of this assumption is substantial, that is, an 81% higher ICER compared with using the 1.5% discount rate for effects that is usually applied in the Netherlands (Table 3).

**3.1.3.2 Uncertainty—Alternative Payment Models** Payers are faced with the challenge to determine whether to reimburse the high one-time costs of VN that potentially, but not certainly, result in lifelong benefits. Managed entry agreements (MEA) between manufacturers and payers are a way to reduce the uncertainty related to the lack of long-term data at launch [27]. MEAs can take the form of discounts or be outcome-based (e.g. pay for performance). The US Institute for Clinical and Economic Review did not investigate a formal alternative payment model, but they performed threshold analyses to estimate the maximum price of VN in order to achieve an ICER that is below certain cost-per-QALY thresholds that could be used to determine discounts for VN. In the NICE submission, confidential simple discount patient access schemes were included [9]. In the Netherlands, ZIN advised to implement a pay-for-performance arrangement with the market authorization holder of VN (Novartis) as a condition for reimbursement, the details of such an agreement have not been disclosed following confidential negotiations between the Dutch Ministry of Health and the market authorization holder [11].

## 3.2 Sensitivity and Scenario Analyses

Table 3 summarizes the sensitivity and scenario analyses performed with the 'Dutch model' mentioned in the discussion of the Drummond checklist in paragraph 3.1. The assumption of the duration of treatment effect, discount rates and utility values had the largest impact on the cost-effectiveness results. As discount rates are prescribed by national health economic guidelines, only the duration of treatment effect and utility values were considered in two-way sensitivity analyses using the Dutch model. The results are presented in Fig. 4 and Table 4.

### 3.3 Comparison of Cost-Effectiveness Results

Table 5 summarizes the cost-effectiveness results of the four economic evaluations converted to 2018 US dollars. Due to the relatively short extrapolation of treatment effect, the use of utilities from a diabetic retinopathy population and the model structure with one health state with varying visual impairment, the ICER of US Institute for Clinical and Economic Review was the highest of the four economic evaluations [7]. In contrast, due to the lifetime extrapolation of treatment effect and the use of utilities based on the RPE65-mediated IRD vignette study, Johnson et al. reported the lowest ICER of the four studies [12].

Despite the longer duration of treatment effect and use of the same utility source as the Dutch HTA submission, the ICER of the UK HTA submission was higher than the ICER of the Dutch HTA submission from a healthcare perspective. This is not only due to the different extrapolation of treatment effect; if a 40-year treatment effect duration is applied in the Dutch model with a healthcare perspective, the ICER would be even lower (i.e. US\$67,371/QALY). This is mainly due to the difference in discount rates between the UK and the Netherlands. If the discount rates in the Dutch model are adjusted to the UK discount rates (i.e. 3.5% for both costs and effects), the ICER increases to US\$122,784/QALY. If the other model settings were also equal to the UK settings (including average eye instead of best-seeing eye and no use of cross-over data), the ICER in the Dutch model would be comparable to the UK ICER (i.e. US\$115,077/QALY).

In contrast to the healthcare perspective, the ICER of the UK is lower than the ICER of the Netherlands when applying a societal perspective. This is caused by the inclusion of more societal costs in the societal perspective of the UK than the Netherlands. In addition to informal care costs, the UK included social security benefits and caregiver productivity losses. Furthermore, patient productivity losses were estimated with the human capital method, while the friction cost method was applied in the Netherlands in which productivity losses are only included for a limited period of time (i.e. 3 months in the Netherlands).

## 4 Discussion

Four economic evaluations of VN were discussed in this paper. Properties of the clinical evidence, such as sample size, duration of the clinical trial and the use of surrogate outcomes, required a number of assumptions to model long-term cost-effectiveness. Different reimbursement authorities preferred different assumptions and this resulted in large differences in the cost-effectiveness estimates, mainly due to different assumptions on duration of treatment effect of VN, discount rates, sources of utility values and model structures. However, the difficulty in translating available clinical evidence to long-term cost-effectiveness estimates is not unique to gene therapies. Also, the impact of assumptions can be assessed with standard methods for sensitivity analyses that can also be applied to other types of therapies. Therefore, this study underlines the statement of Drummond et al. that gene therapies do not require a new methodological reference case for economic evaluation. However, as Angelis et al. rightly expressed, gene therapies are “pushing against the boundaries of the methodological and budgetary capacity available” [35]. Our study showed that guidelines on the acceptable extrapolation of treatment effect of gene therapies and on the reporting of results when important input parameters are highly uncertain are necessary to aid HTA research and align submissions of new gene therapies.

**Table 5** Incremental cost-effectiveness ratios (ICERs) in 2018 US dollars of the four economic evaluations

	Healthcare perspective (US\$)	Societal perspective (US\$)
US-ICER report	643,813	480,130
US Johnson et al.	79,618	VN dominant
UK HTA submission	115,513	86,247
NL HTA submission	99,104	95,625

*HTA* Health technology assessment, *NL* The Netherlands, *UK* United Kingdom, *US* United States, *US ICER* United States Institute for Cost-Effectiveness Research, *VN* voretigene neparovec

**Table 4** Two-way sensitivity analysis varying duration of treatment effect and utility values

Duration of treatment effect, in years	Proxy utility assessment using EQ-5D-5L in same patient population (source Lloyd et al. 2019 [18]) (€)	Proxy utility assessment using HUI-3 in same patient population (source Lloyd et al. 2019 [18]) (€)	Direct utility assessment with TTO in different patient population (source Brown et al. 1999 [17]) (€)
10 + 10 waning	115,673	131,147	184,417
20	80,994	91,445	128,368
40	54,310	61,002	85,873
70 (± life-long)	47,239	52,628	74,325

*HUI* Health Utility Index, *TTO* time trade-off

In addition, alternative payment models, such as MEAs between manufacturers and payers, can reduce the financial risk of the payer associated with the uncertainty of (cost-) effectiveness due to the lack of long-term data at launch.

In the case of VN, uncertainty was mostly related to the extrapolation of the treatment effect. According to Gonçalves, the willingness to consider economic models that include extrapolating longer-term benefit from limited existing data varies between HTA agencies [27]. In the case of VN, both NICE and ZIN did not have objections about the extrapolation of treatment effect proposed by the manufacturer for the base-case analysis, but in the UK, this was set to 40 years, while in the Netherlands a more conservative assumption of 20 years was applied. Uncertainty about the long-term benefits is not unique to VN but will likely be an important question in other gene therapies as well due to the unavoidable lack of long-term data supporting lifetime benefits at market entry [36]. One may even suggest that it is impossible to define a base case based on this limited data, but in practice HTA agencies will probably request or define the most plausible scenario anyway. Therefore, the methodological reference case may not need to be changed for this issue, but it does call for guidance on the acceptable extrapolation of treatment effect and the required evidence base (e.g. duration of trial or expert opinion) to support this assumption. Furthermore, the uncertainty around the assumed duration of treatment effect can be handled in different ways. In the economic evaluations of VN, the uncertainty around the treatment effect duration assumption was addressed by performing extensive scenario analyses assuming different treatment effect durations that can be presented in a table or plot (e.g. Fig. 4 in this paper). This is in line with the new approach of the US Institute for Clinical and Economic Review, who decided to perform ‘optimistic’ and ‘conservative’ scenario analyses in which the clinical benefit is varied in their analyses of high-impact “single and short-term therapies” [37], and recommendations by Angelis et al. [35]. Although these scenario analyses can show the importance of this assumption on the cost-effectiveness results, it is still up to the discretion of the decision maker how this uncertainty is weighted in the final reimbursement decision. One way to potentially improve the plausibility of the treatment effect duration assumption is to include feedback from clinical experts about the expected long-term benefits in a formal manner using the Sheffield Elicitation Framework (SHELF) [38]; for example, like Cope et al. did for CAR-T therapy [39]. Experts could be asked to estimate upper and lower plausible limits and then the most likely duration of the treatment effect of VN based on their clinical experience and knowledge, for example. Based on the uncertainty around the most likely estimates of the experts provided by the upper and lower plausible limits, a probability distribution of this input parameter could be included in the PSA.

This way the uncertainty around the treatment effect duration would be formally included in the cost-effectiveness plane and acceptability curve.

Another important issue was the small patient population, which, among other things, complicated collecting utility values. Utilities in the HTA submissions in the UK and the Netherlands as well as Johnson et al. were based on a vignette-based study where utilities were estimated based on expert valuation of health state descriptions given by patients and caregivers [18]. The main critique of this study by HTA agencies was that the resulting utilities are based on proxy utility assessment [9, 11]. Ideally, utility data should have been collected directly from (the parents of) patients using a quality-of-life questionnaire (i.e. HUI or EQ-5D-5L) during the RCT. The main limitation of a direct valuation study would be that it yields data from a small number of patients. However, the transition probabilities used in the health economic model are also based on this small number of patients and the patient-reported utilities could at least have been used as validation of the utilities derived by expert elicitation. In the specific case of vision problems, additional problems arise in using EQ-5D to collect utility data as it does not include a vision component [19]. In contrast, the HUI-3 does include a vision component and indeed showed better performance in patients with vision disorders and the majority of studies on the convergent validity of the HUI-3 demonstrated good validity [19, 20]. Another alternative is using a mapping algorithm to convert VA to EQ-5D-3L published recently by Pennington et al. [40]. However, an important limitation of using this algorithm in the case of VN would be that it is only based on VA and not on the primary outcome of the clinical trial (i.e. MLMT) or any of the other important secondary outcomes (i.e. VF and FST). In contrast to extrapolation of treatment effect, limited quality of available data on HRQoL is not necessarily an issue in all gene therapies. Therefore, this issue also does not call for changing the methodological reference case for gene therapies. The specific (and often small) patient populations with certain gene mutations that gene therapies are often targeted to, however, bring the same challenges as with evaluating interventions for orphan diseases. Therefore, guidance and regulations for rare diseases may also be applicable to gene therapies.

Although sensitivity analyses can provide more insight on the impact of certain uncertain parameters, the capability of these analyses to help deal with the decision uncertainty is limited [2]. In practice, MEAs are often arranged to share the financial risk between the healthcare payer and the manufacturer [41]. This often entails that funding and use of the intervention is restricted to certain centres, and real-world patient outcomes must be recorded [42]. Outcomes-based arrangements have their limitations, mostly related to the difficulty in defining appropriate outcomes and thresholds

to define response, and the significant clinical, administrative and patient burden. This has historically limited their implementation and favoured confidential discounts instead [42]. In the case of VN, however, the patient population is limited and durable outcomes that are sensitive to change are well defined [15, 43]. The feasibility of such an agreement has been shown in the US, where the manufacturer of VN (Spark Therapeutics) is offering outcome-based agreements that include rebates to payers if established efficacy goals (in terms of full-field light sensitivity threshold scores against baseline measurements) are not reached at 30–90 days and 30 months [44].

## 5 Conclusion

As the first gene therapy is approved by the FDA and EMA for retinal diseases, clinicians are gaining their first experiences with gene therapies that may increasingly become standard care in the future. At the same time, VN also provided a valuable learning experience for health economists by highlighting the challenges in evaluating the cost effectiveness of gene therapies, which can help in improving the economic evaluation of future gene therapies. But in general, this study showed that standard methods can be used for the economic evaluation of gene therapies, supporting Drummond et al. statement. However, due to uncertainty about the (long-term) outcomes of gene therapies, health economists evaluating the same intervention can end up making different methodological choices resulting in large variations in cost-effectiveness results. This study showed that the main challenge in evaluating gene therapies is handling the promise of long-term benefits that are not observed in clinical trials yet. This calls for guidance on the acceptable extrapolation of treatment effect of gene therapies in the base-case analysis and on how to handle the uncertainty around this extrapolation in scenario and sensitivity analyses to aid HTA research and align submissions of future gene therapies. As a single assumption on extrapolation of treatment effect is likely to be incorrect, exploration of the impact of the main assumptions that drive model results in a table or figure as presented in this paper (Fig. 4) is recommended. In addition, alternative payment models can reduce the financial risk caused by uncertainty of long-term outcomes due to lack of long-term data at launch.

## Declarations

**Funding** Simone Huygens, Matthijs Versteegh and Tim Kanters are employees of iMTA, which was sponsored by Novartis Pharma B.V. to independently review and alter the global cost-effectiveness model for voretigene neparvovec for RPE65-mediated inherited retinal disease to the national setting.

**Conflict of interest** Simone Huygens, Matthijs Versteegh and Tim Kanters are employees of iMTA, which was sponsored by Novartis Pharma B.V. to review and alter the global cost-effectiveness model for voretigene neparvovec for RPE65-mediated inherited retinal disease to the national setting. The sponsor contract contains a clause guaranteeing independent decisions to uphold the scientific quality and integrity of the study and the manuscript, including freedom to publish irrespective of the outcomes of the study. Stefan Vegter is an employee of Novartis Pharma B.V. Jan Schouten did not have any conflict of interest.

**Ethics approval** Not applicable.

**Consent to participate** Not applicable.

**Consent for publication** Not applicable.

**Data availability** The manuscripts of the economic evaluations that were reviewed in this manuscript are publicly available in the published literature (see references). The technical report of the Dutch health economic model is available from the corresponding author upon request.

**Code availability** The codes of the underlying health economic models used in the economic evaluations that were reviewed in this manuscript are not publicly available.

**Author contributions** Simone Huygens, Tim Kanters, Matthijs Versteegh and Stefan Vegter contributed to the study conception and design. Data analyses using the Dutch health economic model were performed by Simone Huygens. The first draft of the manuscript was written by Simone Huygens and all authors commented on subsequent versions of the manuscript. All authors read and approved the final manuscript.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

## References

1. Dunbar CE, High KA, Joung JK, Kohn DB, Ozawa K, Sadelain M. Gene therapy comes of age. *Science*. 2018;359(6372):eaan4672.
2. Drummond MF, Neumann PJ, Sullivan SD, et al. Analytic considerations in applying a general economic evaluation reference case to gene therapy. *Value Health*. 2019;22(6):661–8.
3. Food and Drug Administration (FDA). FDA Briefing Document—Advisory Committee Meeting October 12, 2017—BLA 125610 Voretigene Neparvovec of Spark Therapeutics, Inc.; 2017.
4. European Medicines Agency (EMA). CHMP summary of positive opinion for Luxturna. Published 2018. <https://www.ema.europa.eu/en/medicines/human/summaries-opinion/luxturna>. Accessed 21 Sept 2018.

5. 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):849–60.
6. European Medicines Agency (EMA). European Public Assessment Report (EPAR) Luxturna (Procedure No. EMEA/H/C/004451/0000). 2018. <https://www.ema.europa.eu/en/medicines/human/EPAR/luxturna>.
7. Institute for Clinical and Economic Review (ICER). Voretigene Neparvovec for Biallelic RPE65-Mediated Retinal Disease: Effectiveness and Value. Published online 2018.
8. Zimmermann M, Lubinga SJ, Banken R, et al. Cost utility of voretigene neparvovec for biallelic RPE65-mediated inherited retinal disease. *Value Health*. 2019;22:161–7.
9. National Institute for Health and Care Excellence (NICE). NICE Committee Papers—Voretigene Neparvovec for treating inherited retinal dystrophies caused by RPE65 Gene Mutations [ID1054]; 2018.
10. Farmer C, Bullement A, Packman D, et al. Voretigene neparvovec for treating inherited retinal dystrophies caused by RPE65 gene mutations: an evidence review group perspective of a NICE highly specialised technology appraisal. *PharmacoEconomics*. 2020;38(12):1309–18. <https://doi.org/10.1007/s40273-020-00953-z>.
11. Zorginstituut Nederland (ZIN). Pakketadvies Sluisgeneesmiddel Voretigene Neparvovec (Luxturna®) Voor de Behandeling van Een Specifieke Erfelijke Netvliesandoening.; 2019. <https://www.zorginstituutnederland.nl/publicaties/adviezen/2020/02/17/pakketadvies-voretigene-neparvovec-luxturna>. Accessed 17 Jan 2020.
12. Johnson S, Buessing M, O'Connell T, Pitluck S, Ciulla TA. Cost-effectiveness of voretigene neparvovec-rzyl vs standard care for RPE65-mediated inherited retinal disease. *JAMA Ophthalmol*. 2019;137(10):1115–23.
13. OECD Data. Exchange rates. Published 2020. <https://data.oecd.org/conversion/exchange-rates.htm>. Accessed 20 Mar 2020.
14. Chung DC, McCague S, Yu Z, et al. Novel mobility test to assess functional vision in patients with inherited retinal dystrophies. *Clin Exp Ophthalmol*. 2018;46(3):247–59.
15. Chung DL, Kelly RK, High KA, Lacey A, Viriato D. Long-term effect of voretigene neparvovec on the full-field light sensitivity threshold test of patients with rpe65 mutation-associated inherited retinal dystrophy: post-hoc analysis of phase I trial data. 2019.
16. Brown MM, Brown GC, Sharma S, Kistler J, Brown H. Utility values associated with blindness in an adult population. *Br J Ophthalmol*. 2001;85(3):327–31.
17. Brown GC. Vision and quality-of-life. *Trans Am Ophthalmol Soc*. 1999;97:473.
18. Lloyd A, Piglowska N, Ciulla T, et al. Estimation of impact of RPE65-mediated inherited retinal disease on quality of life and the potential benefits of gene therapy. *Br J Ophthalmol*. 2019.
19. Tosh J, Brazier J, Evans P, Longworth L. A review of generic preference-based measures of health-related quality of life in visual disorders. *Value Health*. 2012;15(1):118–27.
20. Brazier J, Muston D, Konwea H, et al. Evaluating the relationship between visual acuity and utilities in patients with diabetic macular edema enrolled in intravitreal aflibercept studies. *Investig Ophthalmol Vis Sci*. 2017;58(11):4818–25.
21. Viriato D, Bennett N, Sidhu R, et al. An economic evaluation of voretigene neparvovec for the treatment of biallelic RPE65-mediated inherited retinal dystrophies in the UK. *Adv Ther*. 2020;37(3):1233–47.
22. Feng H, Adelman RA. Cataract formation following vitreoretinal procedures. *Clin Ophthalmol (Auckland, NZ)*. 2014;8:1957.
23. Dutch Healthcare Authority (in Dutch: Nederlandse Zorgautoriteit). DBC-zorgproducten. Published 2018. Accessed May 3, 2018. <http://opendisdata.nl/msz/zorgproduct>. Accessed 03 May 2018.
24. 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.
25. Briggs AH, Ades A, Price MJ. Probabilistic sensitivity analysis for decision trees with multiple branches: use of the Dirichlet distribution in a Bayesian framework. *Med Decis Mak*. 2003;23(4):341–50.
26. Drack A, Bennett J, Russell S, High KA, Yu Z, Tillman A, Chung D, Reape KZ, Ciulla T, Maguire AM. How long does gene therapy last? 4 Year follow-up of phase 3 voretigene neparvovec trial in RPE65-associated LCA/inherited retinal disease. 2019.
27. Gonçalves E. Advanced therapy medicinal products: value judgement and ethical evaluation in health technology assessment. *Eur J Health Econ*. 2020;21(3):311–20. <https://doi.org/10.1007/s10198-019-01147-x>.
28. Viriato D, Thakker D, Aouadj C, Kali S. How long does gene therapy last? Overview of the Long-Term Efficacy of Voretigene Neparvovec: From Pre-Clinical to Clinical Evidence; 2018.
29. 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 I trial. *Lancet*. 2016;388(10045):661–72.
30. Zorginstituut Nederland (ZIN). Guideline for conducting economic evaluations in healthcare [in Dutch: Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg]. Published online 2016.
31. Zorginstituut Nederland. Kosteneffectiviteit in de Praktijk.; 2015.
32. Versteegh MM, Ramos IC, Buyukkaramikli NC, Ansari-pour A, Reckers-Droog VT, Brouwer WBF. Severity-adjusted probability of being cost effective. *Pharmacoeconomics*. 2019;37(9):1155–63. <https://doi.org/10.1007/s40273-019-00810-8>.
33. Christ SL, Zheng DD, Swenor BK, et al. Longitudinal relationships among visual acuity, daily functional status, and mortality: the Salisbury Eye Evaluation Study. *JAMA Ophthalmol*. 2014;132(12):1400–6.
34. Jönsson B, Hampson G, Michaels J, Towse A, von der Schulenburg J-MG, Wong O. Advanced therapy medicinal products and health technology assessment principles and practices for value-based and sustainable healthcare. *Eur J Health Econ*. 2019;20(3):427–38.
35. Angelis A, Naci H, Hackshaw A. Recalibrating health technology assessment methods for cell and gene therapies. *PharmacoEconomics*. 2020. <https://doi.org/10.1007/s40273-020-00956-w>.
36. Garrison LP, Jackson T, Paul D, Kenston M. Value-based pricing for emerging gene therapies: the economic case for a higher cost-effectiveness threshold. *J Manag Care Spec Pharm*. 2019;25(7):793–9.
37. Institute for Clinical and Economic Review (ICER). Adapted value assessment methods for high-impact “single and short-term therapies” (SSTs). 2019. <https://icer-review.org/material/valuing-a-cure-final-white-paper-and-methods-adaptations/>.
38. Dias LC, Morton A, Quigley J. Elicitation, vol. 1. Cham: Springer International Publishing; 2018.
39. Cope S, Ayers D, Zhang J, Batt K, Jansen JP. Integrating expert opinion with clinical trial data to extrapolate long-term survival: a case study of CAR-T therapy for children and young adults with relapsed or refractory acute lymphoblastic leukemia. *BMC Med Res Methodol*. 2019;19(1):182.
40. Pennington BM, Hernández-Alava M, Hykin P, et al. Mapping from visual acuity to EQ-5D, EQ-5D with vision bolt-on, and VFQ-UI in patients with macular edema in the LEAVO trial. *Value Health*. 2020. <https://doi.org/10.1016/j.jval.2020.03.008>.

41. Antonanzas F, Juárez-Castelló C, Lorente R, Rodríguez-Ibeas R. The use of risk-sharing contracts in healthcare: theoretical and empirical assessments. *Pharmacoeconomics*. 2019;1–15.
42. Jørgensen J, Kefalas P. Reimbursement of licensed cell and gene therapies across the major European healthcare markets. *J Mark Access Health Policy*. 2015;3(1):29321.
43. 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):1273–85.
44. Salzman R, Cook F, Hunt T, et al. Addressing the value of gene therapy and enhancing patient access to transformative treatments. *Mol Ther*. 2018;26(12):2717–26.