Idebenone for treating visual impairment in Leber's hereditary optic neuropathy in people 12 years and over [ID547]

STA Report

Source of funding

This report was commissioned by the NIHR Evidence Synthesis Programme as project number 136145.

Title:	Idebenone for treating visual impairment in Leber's hereditary optic neuropathy in people 12 years and over
Produced by:	BMJ Technology Assessment Group (BMJ-TAG)
Authors:	Steve Edwards, Director of Health Technology Assessment, BMJ-TAG, London Melina Vasileiou, Clinical Evidence Analyst, BMJ-TAG, London Archie Walters, Health Economist, BMJ-TAG, London Benjamin Farrar, Senior Clinical Evidence Analyst, BMJ-TAG, London Kate Ennis, Senior Health Economist, BMJ-TAG, London
Correspondence to:	Steve Edwards, BMJ-TAG, BMJ, BMA House, Tavistock Square, London, WC1H 9JR.
Date completed:	26/01/2024
Source of funding:	This report was commissioned by the NIHR Evidence Synthesis Programme as project number 136145.
Declared competing interests of the authors	No competing interests were declared which affect the impartiality of this report. BMJ Technology Assessment Group (BMJ-TAG) and the editorial team of The BMJ work independently to one another. The views and opinions expressed in this report are those of the BMJ-TAG.
Acknowledgments:	The EAG would like to thank Professor Andrew Lotery (Professor of Ophthalmology, Faculty of Medicine, University of Southampton) and Dr Denize Atan (Associate Professor in Neuro-ophthalmology, Neuroscience & Genetics, University of Bristol and Honorary Consultant in Neuro-ophthalmology, Bristol Eye Hospital) for providing clinical advice throughout the project, and for providing feedback on the clinical sections of the report. The EAG would also like to thank Professor Augusto Azura-Blanco (Clinical Professor of Ophthalmology, Centre for Public Health, Queen's University Belfast for providing feedback on the clinical sections of the report).
Rider on responsibility for report:	The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.
Report reference:	Edwards SJ, Vasileiou M, Walters A, Farrar B, Ennis K. Idebenone for treating visual impairment in Leber's hereditary optic neuropathy in people 12 years and over [ID547]: A Single Technology Appraisal. BMJ Technology Assessment Group, 2024.

Contribution of authors:

Steve Edwards	Critical appraisal of the company's submission; validated the statistical analyses; provided feedback on all versions of the report. Guarantor of the report
Melina Vasileiou	Critical appraisal of the company's submission; critical appraisal of the clinical evidence; cross checking of company's search strategies; and drafted the summary, background and clinical results sections
Benjamin Farrar	Critical appraisal of the company's submission; critical appraisal of the clinical evidence; cross checking of company's search strategies; and assisted with drafting the clinical results sections
Archie Walters	Critical appraisal of the company's submission; critical appraisal of the economic model; cross checking of company's search strategies; critical appraisal of the economic evidence; carried out the economic analyses; and drafted the economic sections
Kate Ennis	Critical appraisal of the company's submission; critical appraisal of the economic model; cross checking of company's search strategies; critical appraisal of the economic evidence; carried out the economic analyses; and drafted the economic sections

All authors read and commented on draft versions of the EAG report.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AE	Adverse events
ARMD	Age-related macular degeneration
BCVA	Best-corrected visual acuity
CaRS	Case Record Survey
CE	Cost-effectiveness
CEA	Cost-effectiveness analysis
CEAC	Cost-effectiveness acceptability curve
CF	Count fingers
CGIC	Clinical Global Impression of Change
CRB	Clinically relevant benefit
CRR	Clinically relevant recovery
CRS	Clinically relevant stabilisation
CS	Company Submission
CSR	Clinical Study Report
EAG	External Assessment Group
EAP	Expanded Access Program
EBMR	Evidence-based Medicine Reviews
EP	Efficacy population
ETDRS	Early Treatment Diabetic Retinopathy Study
EUCTR	EU Clinical Trials Register
HRQoL	Health related quality of life
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio
IQR	Interquartile range
ITT	Intent-to-treat
KOL	Key opinion leader
LHON	Leber's hereditary optic neuropathy
LOCF	Last observation carried forward
logMAR	Logarithm of the minimum angle of resolution
LP	Light perception
LS	Least squares
LY	Life years
MA	Marketing authorisation
mITT	Modified intent-to-treat
MMRM	Mixed-Model for Repeated Measures
NH	Natural history
NICE	National Institute for Health and Care Excellence
NR	Not reported

OWSA	One-way sensitivity analysis
PAS	Patient access scheme
PSA	Probability sensitivity analysis
PSM	Propensity-score matching
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
QALY	Quality-adjusted life years
QoL	Quality of life
RCT	Randomised controlled trial
RGC	Retinal ganglion cells
ROBINS-I	Risk of Bias In Non-Randomised Studies-of Interventions
ROI	Republic of Ireland
RWE	Real-world evidence
SAE	Serious adverse events
SD	Standard deviation
SLR	Systematic literature review
SmPC	Summary of Product Characteristics
SoC	Standard of care
STA	Single Technology Appraisal
TTO	Time-trade off
VA	Visual acuity
VAS	Visual Analog Scale
VF	Visual function
WHO ITCRP	World Health Organisation International Clinical Trials Registry Platform
WTP	Willingness-to-pay

1 Executive summary

This summary provides a brief overview of the key issues identified by the External Assessment Group (EAG) as being potentially important for decision making. It also includes the EAG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 provides an overview of key model outcomes and the modelling assumptions that have the greatest effect on the ICER. Sections 1.3 to 1.4 explain the key issues in more detail. Background information on the condition, technology and evidence and information on non-key issues are in the main EAG report.

All issues identified represent the EAG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

1.1 Overview of the EAG's key issues

Table 1. Summary of key issues

ID	Summary of issue	Report sections
1	There is no precise effectiveness estimate for treatment with idebenone beyond six months to draw robust conclusions about its long-term clinical and cost-effectiveness.	3.2.4, 3.4
2	The benefit of treatment with idebenone may be larger in subgroups of patients but the limited sample sizes available in the current evidence leads to a high degree of uncertainty.	2.3.5, 3.3.4
3	The company model structure is inappropriate given the insufficient evidence to support the high number of health states in the economic model. Additionally, there are limited data to provide robust transition probabilities for the company's model and, given the modest differences in HRQoL between the health states, the justification for a high number of health states is weak.	4.2.2
4	The model fails to accurately replicate the SoC treatment effects as measured in studies and clinical trials, with the company failing to derive a treatment effect using all appropriate available data.	4.2.4
5	The model lacks the functionality to allow idebenone and SoC transition probabilities to vary according to treatment effectiveness uncertainty. The PSA therefore fails to account for treatment effectiveness uncertainty.	4.2.4

Abbreviations: EAG, External Assessment Group; HRQoL, Health related quality of life; PSA, probabilistic sensitivity analysis, SoC, Standard of care

1.2 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival) and quality of life in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost for every QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Improving logMAR recovery;

Overall, the technology is modelled to affect costs by:

- Being more costly than the current standard of care (SoC);
- Reducing the requirement for additional health care resources;

The modelling assumptions that have the greatest effect on the ICER are:

- The SoC treatment effect;
- Off chart blindness health related quality of life utilities.

1.3 Summary of the EAG's key issues

Table 2. Issue 1: Lack of robust long-term treatment effect estimates for idebenone and standard of care

Report section	3.2.4; 3.4
Description of issue and why the EAG has identified it as important	Randomised controlled trial (RCT) evidence on the efficacy of idebenone compared to standard of care (SoC) is available for up to six months, while evidence on the long-term treatment is limited to observational data. The company uses two unmatched populations from the real-world observational Expanded Access program (EAP) and the Case record survey (CaRS) natural history studies to model the long-term treatment effects of idebenone and SoC, respectively, resulting in an estimate at high risk of bias due to imbalances in prognostic factors between patients from the data sources. Although no matched control analyses are provided in the original CS, following a request by the EAG at the clarification stage, the company provided a propensity-score matching (PSM) analysis of changes in patient's best visual acuity between LEROS and CaRS-I and CaRS-II at Month 24. The EAG notes there were limitations to the PSM analysis and the EAG is mostly concerned that only a limited amount of CaRS follow-up data were included in the analyses by choosing to only analyse a single visit pair, rather than all available data, for SoC patients. Matching resulted in a very limited sample and the baseline characteristics suggest issues with the matching persisted. As a result, the EAG considers there to be a lack of a precise estimate for long-term treatment benefit with idebenone. The EAG considers this to be a fundamental issue impacting the technology appraisal, as with no long-term RCT data or an alternative approach involving adequate matching, a robust conclusion on clinical and cost-effectiveness of long-term treatment with idebenone cannot be drawn.
What alternative approach has the EAG suggested?	The EAG considers that matching the idebenone and SoC cohorts would provide a less biased method to model the long-term treatment effect of idebenone and SoC compared to the company's original approach using unmatched populations and requested that the company conduct a matched-controlled analysis using the LEROS trial with a CaRS matched controlled analysis. In regard to the company's PSM analyses provided following the EAG's request at the clarification stage, the EAG considers matching patients between LEROS and CaRS at baseline and then including all available data from subsequent follow-up visits in the analysis would be a preferable approach, instead of using a single baseline and 24-month visit window only.
What is the expected effect on the cost-effectiveness estimates?	The EAG considers that if a more appropriate matching methodology had been used the ICER would likely increase given that in the matched control analysis provided using LEROS and CaRS patients demonstrates no significant difference in treatment effects.
What additional evidence or analyses might help to resolve this key issue?	The EAG considers that the PSM analysis presented in the company's response to clarification does not provide strong evidence of a clinically meaningful long-term treatment benefit of idebenone over SoC for the treatment of LHON, but it is the only available long-term matched-control comparison of changes in patients' best logMAR VA over time. Thus, this is considered by the EAG to be the most appropriate analysis of the comparative effectiveness of long-term idebenone treatment compared to SoC that is currently available. However, considering the limited sample/amount of data resulting from the matching, the EAG considers further analyses making use of all available data might help resolve remaining uncertainties regarding long-term effectiveness of treatment with idebenone.

Abbreviations: CaRS, Case-record survey; CS, Company submission; EAG, External Assessment Group; EAP, Expanded Access Program; PSM, Propensity-score matching; RCT, Randomised controlled trial; SoC, Standard of care.

Table 3. Issue 2 Subgroup effects

Report section	2.3.5; 3.3.4
Description of issue and why the EAG has identified it as important	The EAG's clinical experts note that the benefit a patient may receive from idebenone treatment may be larger in subgroups of patients treated prior to nadir (i.e. <1 year since symptom onset), with baseline logMAR <1 or in subgroups of patients with a particular genotype. The EAG notes that the clinical trials were not powered to detect subgroup effects with subgroup sample sizes being too small to support meaningful conclusions about a difference in the magnitude of treatment effect between different subgroups of patients.
What alternative approach has the EAG suggested?	The EAG asked the company to comment on whether they believe the clinical and cost effectiveness of idebenone may be larger in a subgroup of patients treated either early on in the disease course or with a baseline logMAR <1 and to provide relevant scatterplots and regression analyses. The company do not believe that results will differ and consider current evidence from the RHODOS trial, the LEROS trial and the EAP show a benefit in patients regardless of disease stage but did not provide relevant scatterplots or analyses for different subgroups of patients across trials.
What is the expected effect on the cost-effectiveness estimates?	It's anticipated that idebenone may be more cost effective in specific subgroups.
What additional evidence or analyses might help to resolve this key issue?	Future trials including larger datasets, sufficiently powered to detect subgroup effects would be useful to resolve uncertainties regarding treatment effectiveness. However, the EAG recognises that LHON is a rare disease, and this may present a challenge.

Abbreviations: EAG, External Assessment Group; EAP, Expanded Access Program; logMAR, logarithm of the minimum angle of resolution standard of care.

Table 4. Issue 3 Cost effectiveness model structure

Report section	4.2.2
Description of issue and why the EAG has identified it as important	The EAG considers that the company's model is flawed and potentially inappropriate for decision making as there is insufficient evidence to support the high number of health states in the economic model. Additionally, there are limited data to provide robust transition probabilities for the company's model and, given the modest differences in HRQoL between the health states, the justification for a high number of health states is weak.
What alternative approach has the EAG suggested?	The EAG has suggested an alternative model structure, which the company has used in a scenario analysis.
What is the expected effect on the cost-effectiveness estimates?	Using the EAG's preferred model led to an increase in the ICER between idebenone and SoC treatments.
What additional evidence or analyses might help to resolve this key issue?	No additional analysis required

Abbreviations: EAG, External Assessment Group; HRQoL, health related quality of life; ICER, incremental cost effectiveness ratio; SoC, standard of care.

Table 5. Issue 4 Model standard of care treatment effects

Report section	4.2.4
Description of issue and why the EAG has identified it as important	The modelled SoC treatment effects do not replicate the RHODOS, RHODOS-OFU or the matched analysis study findings with mean change in logMAR from baseline being substantially greater in the company base case than in the RHODOS trial at 6 months. The EAG additionally considers that a robust SoC treatment effect, which replicates the trial results, may be derived from the available CaRS -I and -II data. However, limited patient observations are used from these studies to inform the SoC treatment effect.
What alternative approach has the EAG suggested?	The EAG has suggested informing the SoC transition probabilities using the patient observations from the CaRS studies matched to LEROS patient population or alternatively the RHODOS-OFU study as the EAG considers these data sources the most appropriate as described in key issue 1 .
What is the expected effect on the cost-effectiveness estimates?	The EAG expects that aligning the modelled SoC treatment effects to that of either the matched CaRS population or RHODOS-OFU will lead to an increase in the ICER, as can be seen in the illustrative scenario conducted by the EAG.
What additional evidence or analyses might help to resolve this key issue?	The EAG considers that if the company validates the modelled SoC treatment effect using appropriate study or trial data this would resolve the issue.

Abbreviations: EAG, External Assessment Group; ICER, incremental cost effectiveness ratio; SoC, standard of care.

Table 6. Issue 5 Failure of the PSA to account for treatment effect uncertainty

Report section	4.2.4
Description of issue and why the EAG has identified it as important	As the model lacks the functionality to allow idebenone and SoC transition probabilities to vary according to treatment effectiveness uncertainty the PSA fails to account for treatment effectiveness uncertainty. As such, the EAG is concerned that the probabilistic results are unfit for decision making given the high degree of uncertainty in the treatment effects that are not captured in the PSA results. Additionally, the EAG considers the company's justification for not including transition probabilities in the PSA, namely that this would lead to additionally uncertainty in the PSA results, is unfounded given that the aim of the PSA is to account for parameter uncertainty.
What alternative approach has the EAG suggested?	The EAG has suggested that transition probabilities be made probabilistic when calculating the PSA results.
What is the expected effect on the cost-effectiveness estimates?	Cost effectiveness estimates will be more robust and reliable by account for the treatment effectiveness uncertainty.
What additional evidence or analyses might help to resolve this key issue?	The EAG considers that only allowing transition probabilities to made probabilistic would resolve this issue.

Abbreviations: EAG, External Assessment Group; PSA, probabilistic sensitivity analysis; SoC, standard of care.

1.4 Summary of EAG's preferred assumptions and resulting ICER

Table 7. EAG's preferred model assumptions

Scenario	Incremental costs	Incremental QALYs	ICER (change from company base case
Company base case	[redacted]	[redacted]	18,758
EAG preferred model structure	[redacted]	[redacted]	27,053 (+8,295)
Using the LEROS study data to derive the idebenone long term treatment effect	[redacted]	[redacted]	28,459 (+9,701)
Applying the LEROS idebenone transition probabilities to SoC patients after RHODOS	[redacted]	[redacted]	59,061 (+40,303)
Adjusting the idebenone treatment discontinuation weighted average calculation and increasing the proportion who discontinue treatment to 10.8%	[redacted]	[redacted]	21,022 (+2,264)
Using the utilities calculated from Lawrence *et al.* that include patients from the Republic of Ireland*[2]	[redacted]	[redacted]	27,780 (+9,022)
No carer disutility applied	[redacted]	[redacted]	21,019 (+2,261)
Applying additional healthcare resource costs according to Meads *et al.**[3]	[redacted]	[redacted]	31,631 (+12,873)
Applying supportive living cost as a one-off cost	[redacted]	[redacted]	25,899 (+7,141)
Applying outpatient care cost as a one-off cost	[redacted]	[redacted]	19,595 (+837)

Abbreviations: EAG, External Assessment Group; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year

*EAG preferred model assumption also required

Table 8. EAG base case results

Interventions	Total Costs (£)	Total LY	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
Deterministic results							
SoC	[redacted]	[redacted]	[redacted]	-	-	-	-
Idebenone	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	130,269
Probabilistic results*							
SoC	[redacted]	-	[redacted]	-	-	-	-
Idebenone	[redacted]	-	[redacted]	[redacted]	-	[redacted]	126,422

Abbreviations: ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality-adjusted life-year; SoC, standard of care.

*Probabilistic results do not account for idebenone and SoC treatment effectiveness uncertainty

2 Introduction and background

2.1 Introduction

This report contains the External Assessment Group's (EAG's) critique of the clinical and cost effectiveness evidence submitted for the Single Technology Appraisal (STA) of idebenone (brand name: Raxone®; Chiesi Farmaceutici, Parma, Italy) for treating visual impairment in Leber's Hereditary Optic Neuropathy (LHON) in people 12 years and over [ID547].

2.2 Background

Section B.1.3 of the Company Submission (CS) provides an overview of LHON. LHON is a rare mitochondrial genetic disease that most often affects young adult males.[4] LHON causes degeneration of the optic nerve, and people with LHON experience a sudden and rapid loss of central vision, usually within weeks of symptom onset.[5] Approximately 95% of people with LHON have one of three mitochondrial DNA (mtDNA) mutations: m.11778G>A; m.14484T>C; and m.3460G>A.[4] Such mutations lead to the dysfunction of complex I of the electron transport chain, causing oxidative stress and the eventual apoptosis of retinal ganglion cells.[6, 7]

Over 1 in 1,000 individuals in the UK Biobank carry a mutation with the potential to cause LHON,[8] but the prevalence of LHON in the UK is rare, i.e., the disease has low penetrance.[8] The topic selection oversight panel for the current appraisal estimated that 471 patients would be eligible for idebenone treatment in clinical practice in England, should the technology be approved for routine commissioning.[9] This figure was calculated based on prevalence estimates for LHON in England from Yu-Wai-Man *et al.* 2003 of around 1 in 31,000 individuals.[10]

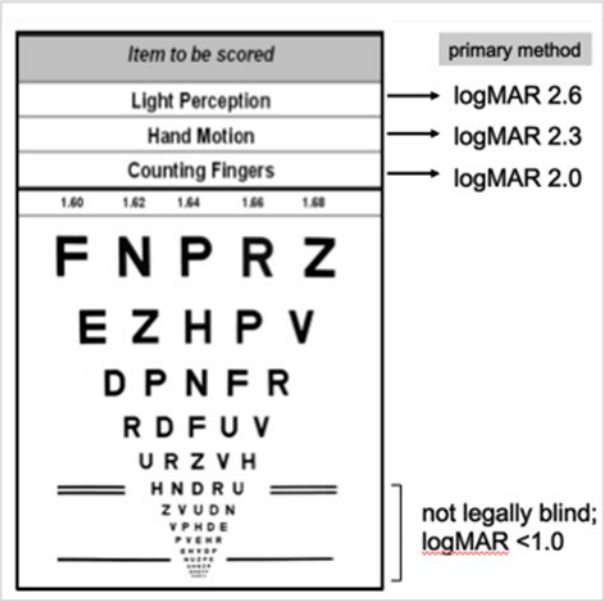
Initially, people with LHON present with a rapid loss of central vision.[5] After presentation, these individuals may be tested and treated for other causes of vision loss before a diagnosis of LHON is suspected and/or confirmed through genetic testing.[5] The typical disease course of LHON can be separated into three phases: subacute/acute, dynamic and chronic. In the acute phase of LHON, central vision is lost in both eyes (25% to 50% of the time) or sequentially (50% to 75% of the time), with the second eye usually being affected to a similar degree as the first eye weeks or months later.[11] The point at which an eye's visual acuity (VA) is lowest is termed nadir, which is usually reached a few months after the onset of symptoms.[12] Following nadir, a patient's VA usually stabilises during a dynamic phase around 6 to 12 months after symptom onset, before the disease enters a chronic phase >12 months after symptom onset.[13] In chronic LHON, a patient's VA is usually stable, but the EAG's clinical experts noted that some further decline is possible.[14] The EAG's clinical

experts also highlighted that the disease course of LHON is heterogeneous, and the natural history of LHON outlined above and in the CS reflects a "textbook" case of LHON. For example, the initial rate of vision loss may be slow and progressive for some patients, and nadir may not be reached until a year or more after initial diagnosis.

2.2.1 Disease progression and disease burden

In clinical trials, vision loss in LHON is usually measured using the Logarithm of the Minimum Angle of Resolution (logMAR) chart of VA (Figure 1).

Figure 1. The logMAR scale including 'off-chart VA' categories (Reproduced from CS Figure 2)



LogMAR values are assessed using the ETDRS charts.
Source: CS Figure 2
Abbreviation: logMAR, Logarithm of the minimum angle of resolution

In the CS, the Company distinguish between the following categories of vision loss based on logMAR values:

- LogMAR <1.0: Not legally blind;
- LogMAR ≥1.0: Legally blind;
- LogMAR ≥1.6: Off-chart.

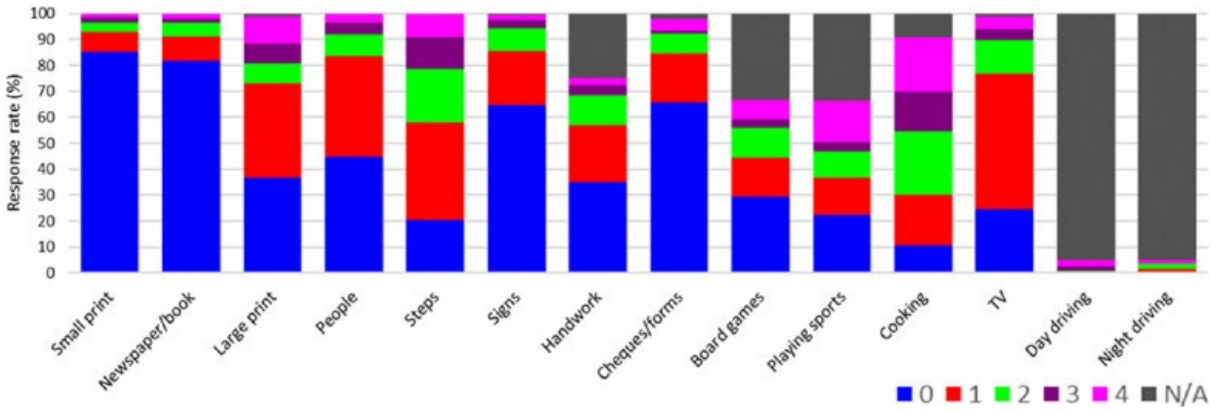
The EAG's clinical experts highlighted that the term "legally blind" is no longer used in clinical practice, and instead people with vision loss may be registered as sight impaired or severely sight impaired. Based on this, the EAG considers the following categories of vision loss based on logMAR values to reflect clinically meaningful categories recognised in English clinical practice:

- LogMAR <0.3: Not sight impaired, able to drive;
- 0.3≤ LogMAR <1.0: Not sight impaired, unable to drive;
- LogMAR ≥1.0: Sight impaired, on-chart sight;
- LogMAR ≥1.6: Sight impaired, off-chart sight.

The EAG notes that the exact criteria for registering an individual as sight impaired or severely sight impaired depends not only on measures of visual acuity, but also on the degree of visual field loss.[15]

LHON affects many dimensions of a person's life. Interview studies of LHON patients and caregivers highlight how LHON can severely impair a person's day to day activities and independence and likelihood of employment.[16, 17] The studies highlight LHON can have a large negative influence on the quality of a person's social, physical and emotional life.[16, 17] Depression, anxiety, and suicidal thoughts are reported for some patients, and people with LHON, and sight loss more broadly, often report stigmatisation following sight loss.[17] The Visual Function Index (VF-14) Questionnaire, originally developed as an index of functional impairment following cataract surgery,[18] is a disease-specific questionnaire designed to assess the level of visual impairment that has been used on a range of ophthalmologic conditions. Questions include: "Do you have any difficulty, even with glasses, reading small print, such as labels on medicine bottles, a telephone book, or food labels?" and "Do you have any difficulty, even with glasses, recognizing people when they are close to you?". Responses are measured on a 5-point scale from "No" to "Unable to do this activity". In a recent VF-14 survey of 196 LHON patients in the UK, Netherlands and Germany, most patients responded either 0 ("Unable to do this activity), 1 ("a great deal of difficulty") or NA to most questions.[1] These data are displayed in Figure 2 to provide an overview of the visual symptom burden of LHON. While the VF-14 can describe some of the symptoms of LHON, the psychometric validity of the VF-14 as a clinical trial endpoint in LHON has been criticised on several measurement grounds including disordered response thresholds and its multidimensionality.[1]

Figure 2. Chen *et al.* distribution of responses to items of the VF-14 by 196 people with LHON. Reproduced from Chen *et al.* Figure 1.[1]



2.2.2 LHON genotypes and spontaneous recovery

Three genotypes, m.11778G>A, m.14484T>C, and m.3460G>A, comprise around 95% of the LHON population.[4] A variety of other LHON genotypes make up the remaining ~5% of the population; however, the EAG's clinical experts highlighted that these mutations are harder to identify as tests for these are not routinely available. Even with complete mtDNA sequencing, interpreting the results of genetic tests for LHON when an individual is negative for one of the three primary mutations is difficult due to the rarity of presentation and characterisation of any suspected disease-causing allele.

LHON genotype is a key prognostic factor for individuals with LHON and affects the likelihood of spontaneous recovery of visual acuity.[10] Spontaneous recovery is inconsistently defined in the literature, but most definitions involve a clinically significant improvement in the number of letter rows (1 or 2 rows) a patient can read on the logMAR chart, i.e., a logMAR improvement of ≥ 0.1. A review of LHON collated the following estimates of the proportion of patients who experience spontaneous visual recovery:[5]

- m.11778G>A: 14% (all age groups), 11% (aged 15 and over);
- m.14484T>C: 37% to 64%;
- m.3460G>A: 15% to 25%.

The review authors highlighted clinical consensus that the m.3460G>A is the genotype with the lowest long-term probability of recovery, and the m.14484T>C genotype is associated with a milder disease and highest probability of spontaneous recovery.[5] This is supported by a VF-14 survey, which reported median VF-14 scores of <20 for people with either a m.11778G>A or m.3460G>A mutation, but a median VF-14 score of >40 for people with an m.14484T>C genotype.

Figure 3. Characteristics associated with LHON genotypes m.11778G>A, m.3460G>A and m.14484T>C

Genotype	Estimated prevalence, international	Estimated prevalence, England	Reported rates of spontaneous recovery	VF-14, median (IQR)
Source	Poincenot *et al.* 2023 (N=1512)[4]	Poincenot *et al.* 2023 (N=139)[4]	Yu-Wai-Man and Chinnery 2021[5]	Kirkman *et al.* 2009[19]*
m.11778G>A	69%	64%	14%	16.7 (9.1 to 29.0)
m.3460G>A	13%	27%	15% to 25%	15.1 (8.1 to 29.3)
m.14484T>C	17%	8%	37% to 64%	43.8 (23.1 to 59.1)

Abbreviations: IQR, interquartile range
*Data digitised by the EAG

The EAG's clinical experts highlighted that other key prognostic factors include:

- Age at symptom onset, with children having a higher rate of spontaneous recovery than adults;[13]
- VA at baseline or nadir, with a less severe early reduction in VA being associated with better long-term prognosis.[20]

2.2.3 Current treatment pathway for LHON

Currently, there are no therapies with marketing authorisation that treat the underlying cause of LHON. Established clinical management for LHON in the National Health Service (NHS) is limited to supportive measures, which include:

- Lifestyle management guidance, including avoiding behaviours that may trigger or exacerbate LHON, such as excessive drinking or smoking;[21]
- Genetic counselling;
- Low vision aids such as magnifiers;
- Occupational and low vision rehabilitation, including optimising features of the home to facilitate use by individuals who are sight impaired.

The EAG's clinical experts noted that further support, such as assistance dogs and technology such as tablets may be provided through the support of charities but are not routinely provided by the NHS.

The EAG notes that there is a large unmet need for people with LHON, and the EAG's clinical experts agreed with the company's clinical experts that treating an individual with confirmed LHON as soon as possible is desirable. This is in-line with the 2017 International Consensus Statement on the Clinical and Therapeutic Management of Leber Hereditary Optic Neuropathy, which stated that: "Idebenone should be started as soon as possible at 900 mg/day in patients with disease less than 1 year."[13] The EAG's clinical experts also noted that they would consider treating with idebenone in the prevalent population many years after diagnosis, should idebenone be available through routine commissioning.

2.3 Critique of the company's definition of the decision problem

A summary of the final scope issued by NICE, together with the company's rationale for any deviation from this, is provided in Table 9. The EAG considers the CS to generally be in-line with the final scope issued by NICE. However, the EAG notes that the clinical efficacy and effectiveness data in the evidence submission comes from patients who had < 5 years since LHON symptom onset. The EAG is concerned that the data from such patients may have limited generalisability to the proportion of the prevalent LHON population in England who have disease duration > 5 years.

Table 9. Summary of decision problem

	Final scope issued by NICE	Decision problem addressed in the submission	Rationale if different from the scope	EAG comment
Population	People aged 12 years and older with Leber's hereditary optic neuropathy (LHON)	As per NICE scope	N/A	The EAG considers the population included in the company Submission (CS) to be largely in line with the NICE final scope. The EAG notes that the clinical evidence available is for people with onset of visual loss of ≤5 years before baseline. The EAG is concerned that the population in the current evidence may overestimate the treatment effectiveness of idebenone in individuals whose symptom onset is >5 years ago. See Section 2.3.1 below for further discussion.
Intervention	Idebenone	As per NICE scope	N/A	The treatment regimen for idebenone in the economic model and the main sources of clinical evidence are consistent with the marketing authorisation for idebenone.[22] See Section 2.3.2 below for further discussion.
Comparator	Established clinical management without idebenone including: • Visual aids. • Occupational and low vision rehabilitation.	As per NICE scope	As per NICE scope	The EAG's clinical experts confirmed that established clinical management without idebenone matches the NICE final scope and established clinical management as described in the CS.

	• Lifestyle management (no smoking, reduced alcohol consumption, diet that includes fresh fruit and vegetables).			The EAG notes that current established clinical management for LHON: does not include any active treatment; does not address the underlying cause of LHON; and does not prevent vision loss or facilitate the recovery of visual functioning. See Section 2.3.3 below for further discussion.
Outcomes	The outcome measures to be considered include: • Visual acuity (VA) • Contrast sensitivity • Retinal nerve fibre layer • Visual field assessment • Adverse effects of treatment • Health-related quality of life	The outcome measures included are: • VA • Contrast sensitivity • Retinal nerve fibre layer • Visual field assessment • Adverse effects of treatment • Health-related quality of life	As per NICE scope	The EAG notes that the company has presented clinical evidence relevant each of the outcomes specified in the NICE final scope. The outcomes used in the economic model are: • Visual acuity (change in best VA/logMAR measurements); and • Health-related quality of life (HRQoL). The EAG agrees that change in best VA is the most relevant clinical effectiveness outcome to include in the economic model. See Section 2.3.4 below for further discussion.
Economic analysis	The reference case stipulates that the cost-effectiveness of treatments should be expressed in	The company is broadly aligned with the overview of the economic analysis outlined in the final scope, except for the cost-	Brown et al. (1999) demonstrated that a patient's quality of life is attributed more by the better-seeing eye than	The EAG notes that results of the economic analysis are expressed in term of an incremental cost per quality-adjusted life years and with the

	terms of incremental cost per quality-adjusted life-year. The reference case stipulates that the time horizon for estimating clinical and cost-effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective. The availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account. The availability of any managed access arrangement for the intervention will be taken into account. The cost-effectiveness analysis should include consideration of the benefit in the best- and worst-seeing eye.	effectiveness analysis, which includes consideration of the benefits in the best- and worst-seeing eye. The cost-effectiveness analysis will only include consideration of the benefit in the best-seeing eye as logMAR VA is measured in the better-seeing eye rather than the worst-seeing eye.	the worst-seeing eye.[23] The better-seeing eye has a higher predictability and consistency when measuring quality of life compared to the worst-seeing eye.[23] Furthermore, change in best VA was the main secondary endpoint in the RHODOS trial. It was considered to be the endpoint most relevant to clinical practice and the one that best reflects the impact of the disease on a patient, being the closest related to visual function in daily life.[24, 25] Furthermore, during protocol assistance the CHMP agreed with the rational for including this endpoint and that it may be more clinically relevant than the primary endpoint analysis (best recovery of logMAR VA between baseline and Week 24). This also aligns with the health technology assessments of idebenone in Wales and Scotland, both of which focused on change in best VA and were granted national reimbursement for patients with LHON.[26, 27]	treatment effect being informed with better seeing eye logMAR VA as described in the NICE final scope and decision problem.

Subgroups to be considered	If the evidence allows the subgroups of people with recent vision loss will be considered.	Within B.2 of the CS, clinical data is presented split by logarithmic minimum angle of resolution (logMAR) score, disease mutation or by acute and chronic patients.	As per NICE scope.	In addition to the subgroup of people with recent vision loss included in the NICE final scope, the EAG considers the following subgroups to potentially impact clinical effectiveness: • Baseline VA (logMAR <1 at baseline vs logMAR ≥1); and • LHON genotype. See Section 2.3.5 below for further discussion; the results of subgroup analyses from the primary sources of clinical evidence are presented in Section 3.3.
Special considerations, including issues related to equity or equality	Guidance will only be issued in accordance with the marketing authorisation. Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued only in the context of the evidence that has underpinned the marketing authorisation granted by the regulator.	There are no special considerations relating to issues of equity or equality.	N/A	The EAG notes that idebenone has been available via routine commissioning in Wales since March 2021.[26]

Abbreviations: CHMP, Committee for Medicinal Products for Human Use; CS. Company Submission; EAG, External Assessment Group; HRQoL, Health related quality of life; LHON, Leber's hereditary optic neuropathy; logMAR, logarithm of the minimum angle of resolution; NICE, National Institute for Health and Care Excellence; VA, visual acuity

2.3.1 Population

The EAG considers the population considered in the CS to be in-line with the NICE final scope: people with LHON aged 12 years and over. The clinical efficacy and effectiveness used in the company's economic model came from one RCT (RHODOS) and two real-world evidence studies (Expanded access program [EAP] and the Case record survey [CaRS-I]). Further clinical evidence was presented from the LEROS clinical trial and Case record survey II (CaRS-II).

The EAG considers populations from each of the clinical trials and real-world evidence sources to be largely consistent with the NICE final scope (See Section 3.2 for further discussion). However, the EAG notes that while idebenone is positioned for all individuals with LHON aged ≥12 years, the studies providing clinical evidence only included individuals for who the onset of visual loss was ≤5 years at baseline. Specifically, the following inclusion criteria were used in each study:

- RHODOS: Age ≥ 14 years and <65 years, impaired VA in at least one eye due to LHON, onset of visual loss due to LHON was 5 years or less prior to baseline, confirmation of either m.11778G>A; m.14484T>C; or m.3460G>A mtDNA mutations at >60% in blood, no explanation for the visual failure besides LHON;
- EAP: confirmation of any of the three major LHON-causative mtDNA mutations and onset of vision loss in the most recently affected eye less than 12 months prior to the date of the Baseline visit;
- LEROS: patients with a diagnosis of LHON, aged ≥12 years and with onset of symptoms within ≤5 years prior to Baseline;
- CaRS-I: historical case record data from LHON patients (with molecular diagnosis), from 11 participating clinical centres; included all patients with no record of idebenone use, whose case records were not previously included in the RHODOS or EAP datasets, where one of the three major LHON-causative mutations was carried, where the date of onset of symptoms in the first affected eye was known and where Presentation was ≤24 months of Onset.
- CaRS-II: historical case record data from LHON patients from 20 sites located in 7 countries; included patients aged ≥12 years, whose onset of symptoms dated after 1999 and was 'well documented' (at least the month of the onset of symptoms was known for each eye), with at least two VA assessments available within 5 years of onset of symptoms and prior to idebenone use, with a genetic diagnosis for LHON for one of the following mtDNA mutations

m.11778G>A; m.14484T>C; and m.3460G>A, with no participation in an interventional clinical trial after the onset of symptoms.

The EAG notes that time since symptom onset is an important prognostic factor for people with LHON. The likelihood of spontaneous recovery is greater early on, i.e. in the dynamic phase in the disease, and the EAG notes irreversible damage to the optic nerve may be established over time, limiting the potential for recovery for patients with longer disease durations.[5] Hence, the EAG has concerns that the population in the current evidence may overestimate the treatment effect of idebenone in the prevalent population with LHON in England, a large part of which is expected to have disease onset > 5 years ago.

EAG clinical experts advised the EAG that genotype is also an important prognostic factor for people with LHON. The EAG notes that the populations included in the clinical evidence were limited to people with the 3 most prevalent genotypes m.11778G>A, m.14484T>C or m.3460G>A. However, the EAG considers that these are representative of the vast majority of patients in England and thus has no concerns that the treatment effect of idebenone is likely to differ.

The EAG also notes that while the population in the NICE final scope is people with LHON aged 12 years and over, patients younger than 14 years were excluded in the RHODOS RCT. However, the EAG considers it reasonable to assume that the safety and efficacy of idebenone observed in the clinical trials would generalise to these patients.

In the economic model, the baseline characteristics of the patient cohort were based on the RHODOS trial. As such, a patient mean of age of 34 at baseline was assumed, with 14% of the population being female.

2.3.2 Intervention

Idebenone (Raxone®), a short-chain benzoquinone, is an antioxidant that as outlined in Table 2 of the CS is thought to re-activate viable-but-inactive retinal ganglion cells (RGCs) in LHON patients by restoring cellular energy (ATP) generation.[28] Idebenone has a marketing authorisation for the treatment of visual impairment in adults and adolescents aged 12 years and over with LHON.[29]

This indication is consistent with the company Submission for this NICE single technology appraisal (ID547). Idebenone is available as 150 mg film-coated tablets, and the recommended dose is 900

mg/day (two tablets, 3 times a day), to be taken with food.[28] No additional tests or investigations are required. Patients should be regularly monitored according to local clinical practice.[28]

Although the duration of treatment is not specified in the SmPC, the company has assumed that patients would continue treatment for up to a maximum of three years. This aligns with the length of follow-up data available from the clinical evidence. EAG clinical experts have confirmed that this is a reasonable assumption but highlighted that patients could be reluctant to stop taking idebenone if they have experienced a benefit from treatment and there is currently very limited evidence demonstrating what happens after treatment discontinuation.

The EAG considers that the dosing regimen of idebenone in the RHODOS trial, the EAP and the LEROS trial to be consistent with its marketing authorisation, with idebenone administered orally at a dose of 300mg (2 x 150mg) three times a day (total daily dose 900mg).

2.3.3 Comparators

The comparator listed in the final scope issued by NICE was established clinical management without idebenone, which includes:

- Visual aids;
- Occupational and low vision rehabilitation;
- Lifestyle management (no smoking, reduced alcohol consumption, diet that includes fresh fruit and vegetables).

As discussed in Section 2.2, there is currently no active treatment tackling the underlying genetic condition of LHON, and patients are currently managed with standard of care (SoC). Within this framework, idebenone potentially presents a step change in the management of LHON. The EAG's clinical experts confirmed established clinical management options without idebenone match the options listed in the NICE final scope and described in the CS; current supportive options included in the SoC do not prevent vision loss or allow recovery of visual functioning. As such, SoC was the only comparator to idebenone in the cost effectiveness model.

In the clinical trials and real-world evidence used to inform the CS, idebenone was compared to placebo or no treatment but established clinical management without idebenone as described in the NICE final scope is available to all LHON patients by default.

2.3.4 Outcomes

The EAG notes that the company has submitted evidence relevant to each of the outcomes specified in the NICE final scope. The clinical outcome used in the economic model is the change in a patient's best VA, which is applied as transition probabilities between logMAR categories for every three months of treatment.

The EAG considers a key difference between the decision problem specified by the company and the NICE final scope is in the endpoint considered in the cost-effectiveness analysis as the company only included consideration of the benefit in the best-seeing eye as logMAR VA is measured in the better-seeing eye rather than also including consideration of the benefit in the worst seeing eye as specified in the NICE final scope. While change in best VA was not the primary outcome of the RHODOS trial, the EAG agrees that change in best VA is the most clinically relevant outcome: the EAG's clinical experts highlighted that a patient's quality of life is primarily driven by the VA of their best seeing eye.

The EAG notes that the company has provided analyses for many outcomes both at the level of the individual eye (e.g. change in logMAR VA of individual eyes) and at the level of the patient (e.g. change logMAR VA of a patient's best eye), and the EAG notes that the treatment effect of idebenone is consistent between each level of analysis.

2.3.5 Subgroups/special considerations

The EAG notes that several baseline characteristics of people with LHON are meaningful prognostic factors and/or treatment effect modifiers. These include:

- LHON genotype:
    - As outlined in Section 2.2.2, the m.14484T>C genotype is associated with a milder disease and higher probability of spontaneous recovery than the m.11778G>A and m.3460G>A genotypes;
- Time since symptom onset:
    - The EAG's clinical experts noted that the treatment effect of idebenone could plausibly be greater for incident patients treated before reaching nadir, although published evidence has mostly been on patients in the dynamic/early chronic phase of the disease;
- VA at baseline or nadir:

- The EAG's clinical experts noted that the long-term possibility of recovering sight will be related to a patient's baseline or worst VA, with meaningful recovery being less likely for patients with worse VA at nadir. The EAG notes this subgroup was used as the basis for a restricted recommendation for idebenone for use within NHS Scotland: patients with LHON who are not yet blind i.e., who do not meet the UK criteria to be registered as severely sight impaired.[27]

The EAG notes that while outcome data are available for each of these subgroups (presented in Section 3.3.4), the limited sample size within each subgroup leads to a high degree of uncertainty in the comparisons.

3 Clinical effectiveness

3.1 Critique of the methods of the review

The company conducted two systematic literature reviews (SLRs) which were presented in Appendix D of the company Submission (CS):

- A clinical SLR that aimed to identify all randomised controlled trials (RCTs) and interventional studies reporting on the clinical efficacy and safety of idebenone and other treatments for LHON;
- A real-world evidence SLR aiming to identify any real-world evidence reporting on the clinical effectiveness and safety of idebenone and other treatments for LHON.

The EAG notes the SLRs were not limited by intervention or comparator, and studies of comparators not relevant to the current appraisal were included throughout. The EAG considered all other eligibility criteria of the SLR to appropriately reflect the final scope as issued by NICE, although studies of no pharmacological intervention (i.e., current SoC) were excluded from the real-world evidence SLR. The EAG is concerned that non-interventional studies were excluded from the real-world evidence review as the comparator for idebenone in the current appraisal is no intervention. This includes the company's own preferred source of long-term data for the comparator cohort in the economic model, the Case Record Survey (CaRS-I) and Case Record Survey II (CaRS-II), which "were excluded from the SLR due to their non-interventional nature, which falls outside the SLR criteria." (CS, page 31). Table 10 provides an overview of the EAG's critique of the company SLRs.

Table 10. Summary of EAG's critique of the methods implemented by the company to identify evidence relevant this appraisal

Systematic review step	Section of CS in which methods are reported	EAG's assessment of robustness of methods
Search Dates	Appendix D1.1.3	**Appropriate** • The primary database searches were conducted on 25 February 2022 and updated on 10 March 2023.
Data sources	Appendix D1.1.3	**Appropriate** • A range of electronic databases were searched, including Embase and MEDLINE, Econlit and a comprehensive search of EBMR; • An appropriate range of conference proceedings were searched between January 2019 and February 2023, detailed in Table 7 of the CS;

		• Nine HTA body websites were searched, and presented in Table 8 of Appendix D; • Ongoing trials were identified through a search of clinicaltrials.gov. The EAG notes that separate searches of EUCTR or WHO ITPR were not reported, but the EAG considers it unlikely any key ongoing trails would have been missed considering the pivotal trials of idebenone was first published in 2011. To verify this, the EAG conducted a search of EUCTR on 20 November 2023 using the key words "LHON" and "Optic neuropathy". These two searches did not identify any relevant data beyond that already identified by the company's SLR.
Search strategies	Appendix D1.1.3 Table 5 and Table 6	**Appropriate** • The EAG considers the search strategies reported in Appendix D to be likely to detect all studies relevant to the current appraisal.
Inclusion criteria	Appendix D1.1.2	**Clinical SLR: Appropriate** • The EAG considers the eligibility criteria to be broader than necessary to identify all clinical trials relevant to final scope issued by NICE. **Real world evidence SLR: Large concerns** • The EAG is concerned that non-interventional studies were excluded from the real-world evidence SLR, but the comparator in the current appraisal is no intervention. The EAG notes that studies not reported in the English language were excluded from both reviews.
Screening	Appendix D1.1.4	**Appropriate** • Screening was performed by two independent reviews at both the title and abstract, and full text, appraisal stages.
Data extraction	Appendix D.1.1.5	**Appropriate** • Data were extracted by a single reviewer and checked for accuracy by a second reviewer.
Tool for quality assessment of included study or studies	Appendix D1.1.5	**Some concerns** • The company used the NICE checklist for RCTs to assess the quality of included RCTs, and the ROBINS-I checklist to assess the quality of included non-randomised studies. The EAG considered these checklists to be appropriate; • Free-text justifications for each quality assessment decision were not reported, which made it difficult to assess the quality and validity of the risk of bias assessments for each study.

Abbreviations: CS, company submission; EAG, External Assessment Group; EBMR, Evidence-Based Medicine Reviews; EUCTR, EU Clinical Trials Register; HTA, health technology assessment; ITT, intent-to-treat; NICE, National Institute for Health and Care Excellence; RCT, randomised controlled trial; ROBINS-I, Risk Of Bias In Non-Randomised Studies - of Interventions; SLR, systematic literature review; WHO ICTRP, World Health Organisation International Clinical Trials Registry Platform.

In the clinical SLR, a total of 1,408 records were identified in the database searches. Following de-duplication (n=268), exclusions of non-human (N=132) and non-English language (N=77) records, 931

records entered the title and abstract appraisal. Of these, 162 records were selected for full-text appraisal, and 35 records were included. A further 23 records were identified through conference and bibliography searches, leading to 58 records from 16 clinical studies being included in the clinical SLR. The clinical trials investigated the following interventions: idebenone (N=3); rAAV2/2-ND4 gene therapy (N=6); EPI-743 (N=2); cyclosporine (N=1); brimonidine purite 0.15% (N=1); skin electrical stimulation (N=1); elamipretide 1% (N=1); and visomitin (N=1). The three idebenone clinical trials subsequently included in the CS were:

- RHODOS: a Phase 2 RCT comparing idebenone (n=55) with placebo (n=30) over 24 weeks of treatment. An observational follow-up visit (RHODOS OFU) was available for N=58 patients, a median of 30 months after the RHODOS Week 24 visit. RHODOS was conducted at sites in England, The Netherlands and Germany;[25]
- LEROS: a Phase 4 single arm study of idebenone (n=181) over a 24-month treatment period. An observational natural history cohort (n=372) was constructed for comparison. LEROS was conducted across 11 countries, including England, Wales, the USA and eight EU nations;[30]
- UMIN000017939: a single arm clinical trial of idebenone (n=57) over 24 weeks of treatment. UMIN000017939 was conducted in Japan.[31]

In the real-world evidence SLR a total of 1,490 records were identified in the database searches. Following de-duplication (n=286), exclusions of non-human (N=81) and non-English language (N=78) records, 1,045 records entered the title and abstract appraisal. Of these, 249 records were selected for full-text appraisal, and 28 records were included. A further 8 records were identified through conference and bibliography searches, leading to 36 records from 22 real world evidence studies being included in the real-world evidence SLR. Twenty of these studies were studies of idebenone alone (N=18) or in combination with vitamin therapy (N=2), one study examined rAAV2 ND4 gene therapy, and one study was of low vision devices. Following clarification, two further studies that were originally excluded from the SLR were re-included, but were not deemed relevant to the CS. The company's updated PRISMA diagram is presented in Figure 3 of Appendix 1 of the company response to clarification.

At clarification, the company stated they considered studies from the real-world evidence SLR for inclusion in the economic modelling based on "various factors such as geographical population, gender proportion, study design, intervention type, and sample size". From this, the Expanded Access Programme (EAP) was identified as "the most robust, being the only multicentre study with

UK patients and one of the largest sample sizes." Reasons for the exclusion of other real-world evidence studies of interventions for LHON were presented in Table 2 of the clarification response. The EAG agrees that the EAP is the most relevant real-world data source of the identified studies. While the EAG considered it plausible to exclude studies such as Van Everdingen 2022,[32] a retrospective multicentre study of idebenone in the Netherlands, could contain relevant data, the study did not report the individual participant transition probabilities between logMAR states that would be required for inclusion in the economic model.

While only one study was identified in the real-world evidence SLR, the company used data from three real world evidence sources in the economic modelling:

- Expanded Access Program (EAP): a retrospective analysis of 111 patients treated with idebenone. Records associated with the EAP were included in the real-world evidence SLR. The EAP was conducted in sites from the UK, Germany, Australia, New Zealand, Poland, Sweden, Spain, Turkey, Switzerland and the USA, and included patients with an onset of vision loss less than 12 months prior to initiating with idebenone only;[33]
- Case Record Survey (CaRS-I) and Case Record Survey II (CaRS-II): retrospective, observational studies of medical records of patients with a genetically confirmed diagnosis of LHON. Both were international studies, and the company explained that only results from the CaRS-I study were available at the time of the submission. The company provided the CSR for CaRS-II following the EAG's clarification questions. CaRS-I reported natural history data for 106 LHON patients;[34] CaRS-II reported natural history data for 219 patients.[35]
- PAROS: a post-authorisation safety study with idebenone due to be published in Q2 2024. Upon request from the EAG, the company provided the clinical study report (CSR) of PAROS, although the EAG notes that data from PAROS are not included in the CS.[36]

3.2 Critique of trials of the technology of interest

In the CS, three studies were presented containing evidence of the clinical efficacy and effectiveness of idebenone. The RHODOS RCT (N=85) comprised the main source of clinical evidence for the efficacy of idebenone up to 6 months (24 weeks),[37] whereas the LEROS Phase IV clinical trial (N=199) and EAP (N=111) provided data on the long-term effectiveness of idebenone for LHON.[38, 39] Data informing the disease course of LHON under established clinical management were presented up to 6 months from the placebo arm of RHODOS, and longer-term data were presented from CaRS-I and CaRS-II: retrospective, observational natural history studies of patients with LHON.[34, 35] The EAG now

presents a critique of the design and conduct of RHODOS, LEROS, the EAP and the CaRS-I and CaRS-II natural history studies.

3.2.1 RHODOS

RHODOS (NCT00747487), was a randomised, double-blind, placebo-controlled multicentre phase II trial, evaluating the efficacy and safety of idebenone in adolescent and adult patients aged ≥14 to <65 years with impaired VA in at least one eye due to LHON with onset of visual loss ≤5 years and a confirmation of diagnosis by identification of either m.11778G>A, m.14484T>C or m.3460G>A LHON mtDNA mutations. A single observational follow-up visit at median time of 30 months (range: 20.9 to 42.5 months; 131 weeks) was performed providing further follow-up data from N=58 participants of the original RHODOS trial. This included patients who previously participated in the RHODOS trial in both the idebenone and placebo arms, but who were not expected to receive idebenone treatment following the completion of RHODOS.

The EAG considered RHODOS to be a high quality RCT with appropriate randomisation, and blinding procedures. However, the EAG notes it was a phase II design with a relatively small population of people at various stages of disease progression and with a short follow-up providing limited evidence on the long-term effect of idebenone therapy. Thus, results should be treated with caution.

The EAG's assessment of the design, conduct, internal validity of the RHODOS trial is presented in Table 11 below.

Table 11. EAG's summary of the design, conduct and analysis of RHODOS

Aspect of trial design or conduct	Section of CS in which information is reported	EAG's critique
Randomisation	Section B.2.3.1 in CS and RHODOS CSR	**Appropriate** Patients were randomised in a 2:1 ratio to idebenone 900mg/day (n=55) or placebo (n=30). Randomisation was stratified by disease history (disease onset more or less than one year prior to randomisation) and by mutation type (m.11778G>A, m.3460G>A and m.14484T>C).
Concealment of treatment allocation	RHODOS CSR	**Appropriate** In the CSR is specified: the randomisation procedure was centralised [redacted]

Eligibility criteria	Section B.2.3.1.1 in CS	**Appropriate but limited to people with onset of vision loss ≤5 years prior to baseline.** Full details of the eligibility criteria for the RHODOS trial population are available in the CS Table 6. Key inclusion criteria were: • Age ≥ 14 years and <65 years; • Impaired VA in at least one eye due to LHON; • Onset of visual loss due to LHON was 5 years or less prior to baseline; • Confirmation of either m.11778G>A, m.14484T>C or m.3460G>A LHON mtDNA mutations at >60% in blood; • no explanation for the visual failure besides LHON. The EAG notes that a considerable proportion of the prevalent LHON population in England will have LHON onset > 5 years ago. Hence, RHODOS trial population may not be representative of the whole spectrum of LHON patients likely to be eligible for idebenone in UK clinical practice.
Blinding	Section B.2.3.1 in CS	**Appropriate** RHODOS was a double-blind, placebo controlled RCT with patients and any people involved in the study (including investigators, site staff, sponsor, and care provider) blinded to study treatment.
Baseline characteristics	Section B.2.3.2.2 in CS	**Appropriate** The EAG's clinical experts noted that participants' length of time since symptom onset, their baseline logMAR as well as the proportion of patients with onset of symptoms >1 year, suggest the population of the RHODOS trial was most likely representative of prevalent LHON patients at the chronic phase of the disease and less likely the earlier subacute/acute and dynamic phases.
Dropouts	RHODOS CSR	**Appropriate** In the CSR, it is reported that "of the 85 patients randomised and treated, 7 patients discontinued the study prematurely, 3 patients (5.5%) treated with idebenone, and 4 patients (13.3%) treated with placebo." The most commonly reported reason for premature discontinuation was withdrawal of consent (2 patients treated with idebenone and 1 patient treated with placebo). One patient in each treatment group was withdrawn due to adverse events." Considering the number of discontinuations and the reasons for discontinuation in each group, the EAG is not concerned about the potential impact of discontinuation in the RHODOS trial upon the results.
Outcome assessment	Section B.2.3.3.2	**Efficacy and safety: Appropriate** **Health related quality of life: Some concerns** Changes in VA were measured using the Early Treatment Diabetic Retinopathy Study (ETDRS) logMAR chart following the prespecified clinical trial protocol. The EAG notes that [redacted], but the EAG considers logMAR measured by ETDRS charts to be a valid endpoint.

		The primary endpoint of RHODOS was the best recovery of logMAR visual acuity in either right or left eye; however, the EAG considers the secondary endpoint, the change in best VA, to be the most clinically relevant endpoint. The EAG notes that these analyses are different analyses of the same fundamental measurement – logMAR score. Quality of life was assessed in RHODOS using the VF-14 questionnaire, which may have poor psychometric validity in LHON patients (see Section 2.2).
Statistical analysis		
Sample size and power	Section B.2.4.1.2.2 in CS	**Some concerns** The company reported that based on VA change of -0.05±0.3 logMAR in the placebo group and -0.25 ±0.3 logMAR in the idebenone group in the ITT population and with the proportion of patients receiving idebenone and placebo of 2:1, 84 patients were estimated to provide 80% statistical power to reject the null hypothesis of no difference in VA change between the two groups. No justification was provided for the VA change by Week 24 in either the placebo or idebenone arms assumed in the power calculation. The EAG agrees with the company's conclusion that 24 weeks "may not have been long enough to fully assess the benefit of idebenone" and considers it likely that a larger sample size with a longer follow-up would be required to allow for a minimum clinically important change in VA to be detected.
Handling of missing data	Section B.2.4.1.2.3. in CS	**Reasonable** Missing data were handled using a Mixed-Model for Repeated Measures (MMRM), assuming data are missing at random. This utilised the observed data to make inferences based on the multivariate normal distribution, with parameters estimated from the available data.
Analysis sets	Section B.2.4.1 in CS	**Some concerns** The ITT population (n=82) included all randomised patients who received at least one dose of the study medication, [redacted] The mITT population (n=81) was the same as the ITT population, but for VA and colour contrast analyses, one patient randomised to placebo, who was identified as a natural history confounder due to ongoing spontaneous recovery of vision at the time of randomisation was excluded. The EAG is concerned that the exclusion of the patient that was considered a natural history confounder in the mITT population biases the efficacy results in favour of idebenone.

Abbreviations: CS, company submission; CSR, clinical study report; EAG, External Assessment Group; ETDRS, Early Treatment Diabetic Retinopathy Study; ITT, intent-to-treat; LHON, Leber's hereditary optic neuropathy; logMAR, logarithm of the minimum angle of resolution; mITT, modified intent-to-treat; MMRM, Mixed-Model for Repeated Measures; RCT, randomised controlled trial; VA, visual acuity VF, visual function.

3.2.2 RHODOS-OFU

In RHODOS-OFU,[40] the long-term follow-up study of RHODOS, patients (n=58) previously randomised to idebenone (n=39) or placebo (n=19) in RHODOS (as described in Table 11), received no treatment. However, there were five patients from the total efficacy population (three from the idebenone group and two from the placebo group) who reported use of idebenone between Week 24 of RHODOS and the RHODOS-OFU single visit (median 30 months, range: 20.9 to 42.5 months). The dose used was not provided in all cases, although three patients reported the use of 900 mg/day. The sub-population recruited to RHODOS-OFU was representative of the RHODOS study population and there were no significant differences with the original RHODOS cohort. However, the EAG notes that the number of patients included in the RHODOS-OFU visit was lower than that included in the original RHODOS trial and the proportion of patients from the original sample included in RHODOS-OFU also differed between the idebenone (73.6%) and the SoC (65.5%) groups. Thus, the EAG has concerns this may indicate selection bias in the inclusion of patients in the RHODOS-OFU trial, considering that patients responding to treatment in the RHODOS trial would be more likely to complete the trial and be willing to participate in the RHODOS-OFU. As a result, the EAG notes that potential selection bias in the RHODOS-OFU visit data favouring idebenone, may overestimate the long-term treatment effect of idebenone compared to SoC (See Section 3.4.4.).

3.2.3 EAP and LEROS

The EAP and LEROS studies provide data on the clinical effectiveness of long-term treatment of LHON with idebenone.

The EAP (N=111) was an open-label, multicentre retrospective, non-controlled analysis of long-term VA and safety in LHON patients treated with idebenone (treatment duration up to 36 months) with onset of vision loss in the second eye less than 12 months prior to the date of the baseline visit. Follow-up time in the EAP ranged between 2.4 and 70.4 months. Patients were seen and followed up after initiating of treatment with idebenone, according to local practice. VA assessments were conducted at regular (generally 3-monthly clinical visits).

LEROS (N=199) was an external natural history controlled open-label, phase IV intervention study assessing the efficacy and safety of long-term treatment with idebenone in adolescent and adult

patients with LHON. LEROS had a 24-month treatment period with visits taking “place at Month 1, Month 3, Month 6, Month 9, Month 12, Month 18 and Month 24”.

The EAG presents a critique of the design and conduct of the EAP and LEROS trial in Table 12, for the idebenone treated patients. The EAG provides a separate critique of the natural history matched-controlled analyses from LEROS, the only statistical analyses presented by the company comparing long-term idebenone treatment with SoC, in Section 3.4.

Table 12. EAG's summary of the design, conduct and analysis of EAP and the LEROS trial

Aspect of trial design or conduct	EAP	LEROS
Randomisation, blinding and concealment of treatment allocation	**N/A** Given the EAP was a real-world, open-label non-controlled analysis, there was no randomisation procedure and blinding, and concealment of allocation were not applicable.	**N/A** Given LEROS was a natural history controlled study of patients treated with idebenone with no enrolled comparator group, there was no randomisation procedure and blinding, and concealment of allocation were not applicable.
Eligibility criteria	**Appropriate but not representative of prevalent population eligible for idebenone in UK clinical practice.** Full details of the eligibility criteria for the EAP population are available in B.2.3.5.2 in the CS and the EAP CSR. Key inclusion criteria were: • A confirmed mtDNA LHON mutation; • Onset of symptoms in the most recently affected eye within 1 year before enrolment. Since the EAP was restricted to patients with onset of vision loss of less than 12 months in the most recently affected eye, the EAG notes it included a population at an earlier stage of disease progression, than RHODOS, LEROS and the prevalent population in England. Thus, the EAG considers EAP patients to be more representative of the incident population of patients with LHON but not the prevalent population forming a large part of clinical practice in the UK.	**Appropriate** Full details of the eligibility criteria for the LEROS trial population are available in Appendix M, Table 2. Key inclusion criteria were: • Impaired VA in affected eyes due to LHON; • No explanation for visual loss besides LHON; • Age ≥ 12 years; • Onset of symptoms ≤ 5 years prior to baseline. Confirmation of either m.11778G>A, m.14484T>C or m.3460G>A LHON mtDNA (for the Intent-to-treat population, not required for enrolment).
Baseline characteristics	**Appropriate but reflective of an incident population with LHON** The EAG considers the baseline characteristics of the EAP to reflect a LHON population in the acute and dynamic phase of LHON, but not the chronic phase. Further discussion of the baseline characteristics of each trial are provided in Section 3.2.5.	**Appropriate** The EAG considers LEROS to contain a mixture of patients in the acute, dynamic and chronic phase of LHON. Further discussion of the baseline characteristics of each trial are provided in Section 3.2.5.

Sample size, dropouts, and long-term data availability	**Large concerns** The company reports that [redacted] In the CSR it is reported that: "… [redacted] Considering the length of follow-up was up to 36 months, the EAG notes a considerable proportion of patients discontinued or were lost to follow-up with data from a limited number of patients being available >24 months and the number of patients progressively decreasing as treatment duration increases (see Table 13 below). The limited number of patients with data available >24 months may limit the robustness of any conclusions about long-term effectiveness of idebenone.	**Appropriate/small concerns** In the CSR it is reported that: "… [redacted] Of the 199 patients enrolled in the LEROS trial, 57 had discontinued at 24 months (CSR, Figure 1). In the CSR it is reported that for the Safety population: "[redacted] The EAG notes that this indicates a considerably larger proportion of patients with data available overtime compared to the EAP. This has been confirmed in the company's response to clarification questions, with data availability for the EAP and LEROS trial displayed in Table 13 below.
Handling of missing data	**Unclear**	**Unclear** A sensitivity analysis assessing the impact of incomplete data was performed with a generalized linear mixed model. In the CSR it is reported that: "[redacted]
Outcome assessment	**Some concerns** Best-corrected visual acuity (BCVA) was "generally assessed using ETDRS logMAR charts or converted from standard Snellen notation to logMAR for analysis purposes" at 3 monthly intervals or according to the treating physician's normal clinical practice.	**Appropriate** BCVA was assessed at every visit using ETDRS logMAR charts, following detailed standardised procedures outlined in the clinical study protocol.

	The EAG notes that the study protocol only specified that "[redacted]" and the mixed recording of BCVA through ETDRS and Snellen charts likely increases the error associated with BCVA measurements in the EAP compared to RHODOS and LEROS. The EAG also notes that as visits could occur "according to the treating physician's normal clinical practice", data missingness is at a higher risk of introducing bias than when all visits are pre-specified.	
Analysis sets	The Safety population (n=111), including all patients enrolled who received at least one dose of idebenone, was used for analysis of safety information. The Efficacy population (n=87) was a sub-population of the Safety population, who carried one of the three major LHON-causative mtDNA mutations, who had time since onset at baseline of less than 12 months in the most recently affected eye, and for whom post-baseline VA efficacy data was available. All analyses for efficacy were carried out on the Efficacy population.	The Safety population (n=198), including patients who received treatment with idebenone was used for the analysis of adverse reactions. The mITT population (n=181) included all patients enrolled in LEROS who: were carriers of one of the three major LHON mtDNA mutations (m.11778G>A; m.3460G>A or m.14484T>C), had received at least one dose of the study medication and provided at least one post-baseline VA assessment. Apart from the Safety population, data were summarised by onset of symptoms (≤1 year or >1 year after onset of symptoms).

Abbreviations: AE, adverse events; BCVA, Best-corrected visual acuity; CSR, clinical study report; EAG, External Assessment Group; EAP, Expanded Access Program; ETDRS, Early Treatment Diabetic Retinopathy Study; LHON, Leber's hereditary optic neuropathy; logMAR, logarithm of the minimum angle of resolution; mITT, modified intent-to-treat; VA, visual acuity.

Table 13. Data availability in EAP and LEROS trial (adapted from Table 9 in company's clarification response)

	Months			
Patient populations	**>0**	**>6**	**>12**	**>24**
EAP EP: Patients with outcome data available N (%)	87 (100.0%)	81 (93.1%)	63 (72.4%)	42 (48.3%)
LEROS ITT: Patients with outcome data available N (%)	196 (100.0%)	171 (87.2%)	151 (77.0%)	125 (63.7%)

Abbreviations: EP, efficacy population; ITT, intent-to-treat.

The EAG notes that, although long-term follow up data spanning 36 months are available from the EAP, the number of patients for which data were available decreased with each clinic visit at a greater rate compared to the LEROS trial, with a considerable difference in the proportion of data available >24 months in favour of the LEROS trial. Thus, the EAG has concerns over the company's choice of the EAP as the preferred source of long-term effectiveness in the economic model as despite the overall length of follow-up for the EAP being longer, the availability of data was considerably lower. See Section 3.3.7 and Section 4.2.4 for further details of the EAG's critique of the company's choice of long-term effectiveness data source.

3.2.4 CaRS-I and CaRS-II

CaRS-I and CaRS-II were multi-centre, retrospective, observational, historical case record surveys of untreated patients with genetically confirmed diagnosis of LHON, providing clinical data on the natural progression of LHON. The studies were collecting historically documented VA data from existing medical records from patients who were not receiving idebenone with no comparison group, thus randomisation was not applicable. This was considered a limitation as similarly to the EAP and LEROS trials, CaRS do not provide direct comparative evidence on long-term treatment with idebenone compared to SoC. Comparative evidence had to be indirectly obtained through a matched controlled analysis of a subgroup of patients from the LEROS trial matched with a natural history group of idebenone naïve patients from data from CaRS-I and CaRS-II. See Section 3.4 for the EAG's critique of the company's matched controlled analysis.

CaRS-I (n=383) collected historical case record data from LHON patients (with genetically confirmed diagnosis), from 11 participating clinical centres; no exclusion criteria were specified, and data were collected without pre-selection, based on participating clinical centres record-keeping practices.

CaRS-II (n=219) collected data from patients with a genetically confirmed diagnosis of LHON who fulfilled the following prospectively defined inclusion criteria:

- Age≥12 years;
- The onset of symptoms was dated after 1999 and was well documented (at least the month of the onset of symptoms was known for each eye);
- At least two VA assessments were available within 5 years of onset of symptoms and prior to idebenone use;
- Have a genetic diagnosis for LHON for one of the following mtDNA mutations: m.11778G>A; m.3460G>A or m.14484T>C.

[redacted]In the CSR of CARS-I, it is noted that the studies [redacted]The EAG has concerns about the robustness of data from CaRS, considering it was a retrospective review of medical records with a large proportion of missing data and a high degree of variability in the availability of data from different patients at different time points. Thus, reliable conclusions about the natural course of VA changes in LHON cannot be drawn.

3.2.5 Trial baseline characteristics

The EAG noted that the baseline characteristics reported across studies included in the CS differed, making it difficult to assess the similarity between the populations. Thus, the EAG requested that the company provide baseline characteristics for each study consistently in a single table. In response to the EAG's request the company provided the following.

Table 14. Baseline characteristics across studies (adapted from Table 7 in the company's initial clarification response)

	RHODOS		EAP		LEROS		CaRS I		CaRS II	
Characteristic	Idebenone N=55 (N=53 ITT population)	Placebo N=30 (N=29 ITT population)	LHON population N=105	Efficacy population N=87	ITT N=196	NH matched comparator N=106	Natural history population N=106	Natural history outcomes population N=74	Natural history population N=219	Natural history outcomes population N=219
Age, mean ± SD [median] (range) (years)	33.8 ± 14.8 [30.0] (14–63)	33.6 ± 14.6 [28.5] (14–66)	31.7±18.5 [23.6] (6.9–80.1)	31.9±17.4 [24.6] (6.9–80.1)	34.1 ± 15.2 [31.9] (12.1–79.2)	32.1 ± 14.5 [28.0] (13.0–75.0)	32.4 (15.5) [29.5] (6 – 79)	31.1 ± 14.6 (7 – 75)	30.0±15.0 [26.0] (6-68)	30.0±15.0 [26.0] (6-68)
Male, n (%)	47 (85.5)	26 (86.7)	82 (78.1%)	71 (81.6%)	[redacted]	88 (83.0)	85 (80.2)	61 (82.4)	175 (79.9)	175 (79.9)
Age at symptom onset mean ± SD [median] (range) (years)	NR	NR	30.8±18.5 [23.0] (6.6 - 78.9)	31.4±17.3 [24.2] (6.6 - 78.9)	32.5 ± 15.2 [30.4] (8.8 – 78.2)	31.7 ± 14.5 [28] (13.0 – 75.0)	32.1 ± 15.4 [29.5] (6 – 78)	30.9 ± 14.6 (7 – 75)	29.8±15.0 [26.0] (6-68)	29.8±15.0 [26.0] (6-68)
Age at diagnosis mean ± SD [median]	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Race, n (%)										

	RHODOS		EAP		LEROS		CaRS I		CaRS II	
Caucasian/white	53 (96.4)	30 (100)	NR	NR	54 (27.6)	NR	NR	NR	NR	NR
Black	1 (1.8)	0	NR	NR	8 (4.1)	NR	NR	NR	NR	NR
Other	1 (1.8)	0	NR	NR	134 (68.4)	NR	NR	NR	NR	NR
Mutations, n (%)										
m.11778G>A	37 (67.3)	20 (66.7)	61 (58.1)	54 (62.1)	[redacted]	77 (72.6)	78 (73.6)	55 (74.3)	157 (71.7)	157 (71.7)
m.14484T>C	11 (20.0)	6 (20.0)	17 (16.2)	16 (18.4)	[redacted]	12 (11.3)	11 (10.4)	7 (9.5)	32 (14.6)	32 (14.6)
m.3460G>A	7 (12.7)	4 (13.3)	18 (17.1)	17 (19.5)	[redacted]	17 (16.0)	17 (16.0)	12 (16.2)	30 (13.7)	30 (13.7)
Other	-	-	2 (1.9)	-	[redacted]	-	-	-	-	-
Negative	-	-	-	-	[redacted]	-	-	-	-	-
Months since onset of vision loss, mean ± SD [median] (range)	22.8 ± 16.2 [17.8] (3–62)	23.7 ± 16.4 [19.2] (2–57)	10.6±18.7 [5.6] (0.9 - 133.7)	6.2±3.7 [5.0] (0.9 - 16.7)	18.4±15.8 [12.3] (0.3-58.3)	NR	Years: 0.3±0.4 [0.2] (0.0– 1.9)	Years: 0.3±0.4 [0.1] (0.0– 1.9)	3.4±5.6 [1.7] (0.7-3.9)	3.4±5.6 [1.7] (0.7-3.9)
Proportion of patients with nadir prior to baselines, n (%)	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

	RHODOS		EAP		LEROS		CaRS I		CaRS II	
Months since nadir at baseline, mean ± SD [median] (range)	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Patients with onset of symptoms >1 year, n (%)	36 (65.5)	19 (63.3)	NR	NR	87 (44.4)	NR	8 (7.5)	2 (2.7)	10 (4.6	10 (4.6)
Onset of vision loss within 1 year, n (%)	19 (34.5)	11 (36.7)	NR	NR	109 (55.6)	NR	98 (92.5)	72 (97.3)	209 (95.4)	209 (95.4)
Baseline logMAR distribution, n (%)										
One eye logMAR ≥1.0	5 (9.4)	2 (6.9)	Best VA: 70 (66.7)	Best VA: 63 (72.4)	NR	NR	NR	NR	NR	NR
Both eyes logMAR ≥1.0 (legally blind)	45 (84.9)	25 (86.2)	NR	NR	NR	NR	50 (47.1)	27 (36.5)	82 (37.7)	82 (37.7)
Both eyes logMAR <1.0	3 (5.7)	2 (6.9)	NR	NR	NR	NR	NR	NR	NR	NR
One eye off-chart	11 (20.8)	3 (10.3)	Best VA: 18 (17.1)	Best VA: 17 (19.5)	NR	NR	NR	NR	NR	NR
Both eyes off-chart	25 (47.2)	13 (44.8)	NR	NR	NR	NR	12 (11.3)	7 (9.5)	19 (8.8)	19 (8.8)

	RHODOS		EAP		LEROS		CaRS I		CaRS II	
Both eyes on-chart	17 (32.1)	13 (44.8)	NR	NR	NR	NR	NR	NR	NR	NR
Patients with both eyes off-chart,* n (%)	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Patients with discordant visual acuities,† n (%)	20 (37.7)	10 (34.5)	NR	NR	NR	NR	NR	NR	NR	NR
logMAR: mean ± SD,‡ (n)										
Best eye	1.61 ± 0.64 (53)	1.57 ± 0.61 (29)	1.16 ± 0.55	1.23 ± 0.52	1.15 ± 0.60	NR	0.75 ± 0.61	0.62 ± 0.61	0.94 ± 0.64 (438)	0.94 ± 0.64 (438)
Worst eye	1.89 ± 0.49 (53)	1.79 ± 0.44 (29)	NR	NR	NR	NR	NR	NR	NR	NR
Both eyes	1.75 ± 0.58 (106)	1.68 ± 0.54 (58)	NR	NR	1.26 ± 0.55	NR	1.03 ± 0.60	0.97 ± 0.63	NR	NR

*Off-chart defined as >logMAR 1.68 (patients unable to read any letter on the chart).

†Defined as patients with difference in logMAR>0.2 between both eyes

‡Applying logMAR 2.0 for counting fingers; logMAR 2.3 for hand motion; logMAR 2.6 for light perception

Abbreviations: EAP, Expanded Access Program; ITT, intent-to-treat; logMAR, logarithm of the minimum angle of resolution; NR – Not Reported; SD – Standard deviation;.

In terms of baseline characteristics, the EAG's clinical experts considered the population from studies included in the CS to be broadly representative of patients seen in clinical practice. However, they noted that patients across trials presented slightly older than the age at which patients tend to present in clinical practice, but the EAG does not consider this likely to impact the results. A further discrepancy was noted in the proportion of male participants in the LEROS trial, where it was [redacted] compared to the RHODOS, the EAP and CaRS-I and CaRS-II patients where the proportion of male patients better reflected UK clinical practice. However, the EAG notes that males still comprised the [redacted] of patients in the LEROS trial and that sex has not been highlighted as a prognostic factor for LHON by clinical experts or indicated in the submitted clinical evidence. Thus, the EAG has no concerns about any potential implication of this discrepancy on the results.

The EAG notes that the range in length of time since onset of vision loss in the RHODOS (2 to 62 months) and the LEROS trial (0.3 to 58.3 months) was wide, suggesting both trials included patients that were representative of both the incident and prevalent population of LHON. Given the range of time since onset of vision loss and the proportion of patients with onset of symptoms >1 year (44.4%) being close to 50%, the EAG notes the LEROS trial was representative of a mixture of patients in the acute, dynamic, and chronic phase of LHON. However, considering most patients in the RHODOS trial (~65%) had onset of symptoms >1 year and their baseline logMAR (>80% with logMAR≥1.0 in both eyes), the EAG considered the RHODOS patients to be more representative of patients in the chronic phase of the disease and less likely to be reflective of patients in the subacute or acute phase of the disease. Contrarily, based on their time since onset of vision loss, the EAG considers the EAP and CaRS patients (within 1 year for >90% of patients) to be representative of LHON patients in the acute and dynamic phase of the disease but not of the chronic phase. Thus, to include patients at an earlier stage of disease progression compared to RHODOS, LEROS and the prevalent population in England. Clinical experts advised the EAG that time since onset in the EAP and CaRS were more reflective of time to diagnosis seen in clinical practice compared to the RHODOS trial where participants' time since onset indicated they received idebenone much later than they would if it was to become available in clinical practice.

Considering the eligibility criteria for the EAP that was restricted to patients with onset of vision loss of less than 12 months (in the most recent eye) in addition to the time since onset of the included patients at baseline, the EAG considers the EAP and the CaRS study patients represent an incident population with LHON and has concerns over its applicability to the prevalent population with LHON

in England. Similarly, the eligibility criteria of the RHODOS trial limited the inclusion of participants to people with onset of vision loss ≤5 years. The EAG has concerns that the population for which data was available is of limited representativeness of the overall prevalent population in England, a considerable proportion of which will have LHON onset >5 years.

The EAG notes the distribution of mutations was largely in line with what is seen in clinical practice in England. However, it was noted that in the LEROS trial, the proportion of people with the m.11778G>A mutation, was [redacted] compared to the RHODOS trial, the EAP and CaRS. EAG clinical experts advised this mutation has the worse prognosis and a lower probability of spontaneous recovery compared to other mutations. Thus, the EAG has some concerns about the potential impact of a difference in the prevalence of mutations in the LEROS trial on the results. The EAG also noted that CaRS included a considerably larger proportion of patients with m.11778G>A mtDNA mutation compared to RHODOS, the EAP and the LEROS trial. EAG clinical experts have emphasised this mutation has a poorer prognosis, thus the EAG is concerned about the impact of this difference on the results and conclusions drawn about treatment with idebenone compared to the SoC using this retrospective review of medical records.

3.3 Critique of the clinical effectiveness analysis and interpretation

In Section B.2.6 of the company submission (CS), the company outlines results for primary and secondary outcomes of RHODOS, RHODOS-OFU and the Expanded Access Program (EAP). While the LEROS trial and matched natural history cohort from CaRS-I and CaRS-II were included in the submission, these are not focused on in the CS and were not included in the economic model due to "heterogeneity between patient populations" (see Section 4.2.4 on the EAGs critique). However, results from LEROS and its matched analysis were included in Appendix M to provide further evidence of the long-term efficacy of idebenone compared to SoC (see Section 3.4 on the EAG's critique of the matched-controlled analysis).

While the EAG agrees that the RHODOS trial is most relevant to the decision problem population given it was the only available RCT, the EAG raised concerns over the company's choice to present results for the mITT population over the ITT population, where possible for the primary efficacy analysis. The exclusion of one patient from the placebo group, that was considered a natural history confounder from the mITT, biases the results in favour of idebenone compared to the results from the ITT population. Therefore, the EAG requested that the company provide results from the ITT population. The request was fulfilled by the company and results are discussed below. While the

EAG considers that bias is likely to be associated with results of the mITT population from the RHODOS trial, the EAG considers it useful that these results are discussed alongside the ITT in the present report for comparative purposes.

All outcomes specified in the NICE final scope were presented in the CS. Changes in logMAR from the RHODOS trial, the EAP and CaRS-I and CaRS-II (natural history cohort) were used in the economic model by the company to inform transition probabilities. The company suggested that change in best VA was the most important outcome to consider, being the outcome that best reflects the impact of the disease on a patient and being the closest related to visual function in daily life. EAG clinical experts agreed change in best VA would be the most relevant outcome from a patients' perspective. Thus the EAG had no concerns over the choice of change in best VA over time as the outcome used to inform transition probabilities in the economic model. Specifically, data from the RHODOS trial informed transitions up to 6 months, while EAP data from patients in the efficacy population (N=87) informed transitions for over 6 months for up to 36 months. Although some patients in the EAP did provide follow-up visits post 36 months, with follow up ranging from 2.4 to 70.4 months, these occurred at variable time points and therefore could not be used to inform transition probabilities. Also, the number of patients on treatment >24 months was moderate (e.g. N=42; 48.3% at 24 months) in the efficacy population, and the number on treatment at 24 months was substantially reduced to nearly half by 36 months (N=23; 26%) with only 12 patients still receiving treatment at month 42.

3.3.1 Change in logMAR/ Change in best VA

3.3.1.1 RHODOS

In the RHODOS trial population, two analyses of changes in logMAR were presented:

- The best recovery of logMAR visual acuity in either right or left eye (primary efficacy endpoint);
- The change from baseline in patients' best VA.

As mentioned in Section 3.3, the EAG considers the change in patients' overall best VA to be the most clinically relevant endpoint, and therefore focuses on these analyses here. The results of the primary efficacy endpoint are summarised later in Table 16.

For the outcome of change in best VA, best VA at week 24 (best eye at Week 24) compared to best VA at baseline (best eye at baseline). Best recovery of logMAR VA in either right or left eye between baseline and Week 24 was reported for people with improving VA. In patients with neither eye improving in VA between baseline and Week 24, the change in VA representing the 'least worsening' was evaluated as 'best recovery'.

In the RHODOS ITT population, the difference between idebenone and SoC in the change in best VA from baseline to 24 weeks did not reach statistical significance. In people receiving idebenone, logMAR slightly improved with a change in logMAR of –0.035 (95% CI: –0.126 to 0.055), which equated to an improvement of only one letter on the Early Treatment Diabetic Retinopathy Study (ETDRS) chart. For people receiving placebo there was a worsening of logMAR +0.085; 95% CI: –0.032 to 0.203, which equated to worsening of 4 letters on the ETDRS chart. The between group difference was not statistically significant (logMAR –0.120, 95% CI: –0.255 to 0.014); equating to a 6-letter change (p = 0.078).

Best recovery of logMAR between baseline and 24 weeks for people receiving idebenone improved with a mean logMAR value of –0.135 (95% CI: –0.216 to –0.054). This equated to an improvement of 6 letters on the ETDRS chart. For people receiving placebo, the mean change from baseline also improved, with a logMAR value –0.071 (95% CI: –0.176 to 0.034), equating to an improvement of 3 letters on the ETDRS chart. The estimated mean difference between groups was not statistically significant (logMAR –0.064, 95% CI: –0.184 to 0.055); equating to a 3-letter change (p = 0.291).

Instead of the ITT population, the company presented the results of the RHODOS mITT population as the primary efficacy results in the CS. The RHODOS mITT population used the same population as the ITT but excluded one patient for VA data who had been randomised to placebo and was considered a natural history confounder due to an ongoing spontaneous recovery of vision at the time of randomisation. The EAG notes that the exclusion of one patient from the placebo group in the mITT analysis resulted in a considerable increase of the between group difference. However, when the analysis was based on the mITT population, the difference between treatment groups for all patients was still not statistically significant for the outcome of best recovery of logMAR VA (difference between groups –0.100, 95% CI –0.214 to –0.014; p = 0.0862), corresponding to a 5-letter difference on the ETDRS chart. Although, in the result for the change from baseline of best VA there was a statistically significant difference between groups in favour of idebenone (logMAR –0.160, 95% CI: –0.289 to –0.031; p = 0.015) that corresponded to an 8-letter difference on the ETDRS chart.

The EAG considers the results from the ITT population of RHODOS, which did not exclude any patients, are likely to be less biased compared to the mITT. The EAG notes that the patient excluded from the ITT population was identified as a natural history confounder due to on-going spontaneous recovery of vision at the time of randomisation to the study and their trajectory of VA being considered unusual compared to other patients, showing a marked improvement immediately prior to enrolment into the RHODOS trial. However, EAG clinical experts advised the EAG that spontaneous recovery of vision can reflect the natural progression of LHON in some patients. Thus, the EAG considers that the patient identified as a confounder should be included in the analysis.

The EAG also notes that this patient was excluded retrospectively and any criteria for exclusion from analysis had not been specified prospectively. Thus, the EAG considers the definition of the mITT population to be at high risk of bias.

3.3.1.2 RHODOS-OFU

The observational, single visit, follow-up study of RHODOS, RHODOS-OFU examined change in VA in 58 of the 85 patients originally included in the RHODOS trial for a median time of 30 months (range 20.9 to 42.5 months; 131 weeks). The mean change in best VA compared the results of the current VA with the observed VA at the original baseline and after 24 weeks of treatment in RHODOS. In patients in the placebo group, best VA at the RHODOS–OFU visit was slightly worse than at baseline (mean change in logMAR +0.039, corresponding to a worsening of 1 letter), whereas in the idebenone group best VA improved (mean change in logMAR –0.134, corresponding to an improvement of 6 letters). There was a benefit of treatment with idebenone that was maintained during the off–treatment period of the RHODOS–OFU follow-up but the difference between idebenone and placebo groups was not statistically significant (between group difference logMAR –0.173, 95% CI: –0.370 to 0.024; 8 letters; p = 0.0845). No statistical differences between groups were observed for baseline to week 24 of RHODOS (logMAR –0.175, 95% CI: –0.375 to 0.024; 8 letters; p=0.0844) or week 24 of RHODOS to the OFU visit (logMAR +0.002, 95% CI: –0.190 to 0.195; 0 letters; p=0.9819).

3.3.1.3 EAP and LEROS

The amount of data available at each timepoint from the key analysis set used from the EAP, LEROS and CaRS differed. Of the 87 patients with outcome data available at baseline in the EAP efficacy population, N=81 (93.1%) had data available >6 months, N=63 (72.4%) had data available>12 months

and 42 (48.3%) had data available >24 months. The EAG notes that the availability of outcome data over time from the LEROS ITT population was greater than the EAP. Of the 196 patients with outcome data available at baseline, N=171 (87.2%) had data available >6 months and N=151 (77%) had data available >12 months, and >24 months N=125 (63.7%) had data available, which indicated a significantly larger proportion compared to the EAP efficacy population at this time point.

In the EAP, there was a slight improvement in best VA from baseline to the last visit in the efficacy population (people who carried one of the three major LHON-causative mtDNA mutations with <12-month onset in the most recent eye), with logMAR decreasing from 1.23 (95% CI: –0.18 to 1.80) at baseline to 1.19 (95% CI: –0.16 to 1.80) at last visit.

In the LEROS trial, there was a slight improvement in best VA from baseline to 24 months in the ITT population with a mean (SD) change in logMAR of –0.09 (0.72) in people with disease onset in the second eye of ≤1 year and a mean (SD) change in logMAR of –0.19 (0.31) in people with disease onset in the second eye of >1 year.

3.3.1.4 CaRS-I and CaRS-II

The number of patients with outcome data available overtime from the CaRS-I natural history outcomes population was unclear, while the availability of outcome data from CaRS-II natural history population reduced substantially overtime, with N=203 (92.7%) of the total 219 people with data >6 months, N=58 (26.5%) with outcome data available >12 months and N=26 (11.9%) with outcome data available >24 months.

[redacted]

Data is mean (and 95% CI) of VA data from the Natural History Population over time since Onset of symptoms. Note: logMAR VA means and CIs calculated using logMAR 1.7 for all off-chart VA categories.

The distribution of all eyes of patients within the Natural History Outcomes Population between VA categories of logMAR <1.0, logMAR 1.0-1.68 or logMAR >1.68 ('off-chart' VA) at presentation, nadir and outcome assessment is presented in Figure 5 below.

Figure 4. Analysis by VA Category for Eyes at Presentation, Nadir and Outcome in the Natural History Outcomes population (reproduced from Case Record Survey CSR)

[redacted]

In CaRS-II, best VA was assessed during the periods of time indicated in the Table 15 below.

[redacted]

Table 15. Best VA at follow-up (reproduced from Table 11.3.1 in CaRS-II CSR)

	≤ 1 year (N=203)	1 to 2 years (N=58)	2 to 3 years (N=26)	3 to 4 years (N=25)	4 to 5 years (N=18)	> 5 years (N=37)
Best VA within 1 year follow-up (logMAR)						
Mean ± SD	[redacted]					
Median (Q1 – Q3)	[redacted]					
Min – Max	[redacted]					
Best VA within 1 year follow-up (blindness category)						
off-chart	[redacted]					
1.0 to 1.68 logMAR	[redacted]					
< 1.0 logMAR	[redacted]					
Difference in Best logMAR from visit and Baseline						
Mean ± SD	[redacted]					
Median (Q1 – Q3)	[redacted]					
Min – Max	[redacted]					

The data used from the CaRS studies are discussed in greater detail in Section 3.4.

3.3.2 Other Outcomes

The EAG notes that to evaluate the effectiveness of idebenone to prevent further vision loss (stabilisation) and recover lost vision (recovery), the company defined a clinically relevant benefit (CRB) to include clinically relevant recovery (CRR) or clinically relevant stabilisation (CRS) of visual acuity (VA). Across trials, CRR is defined as improvement of at least logMAR 0.2 (equal to two lines of readable letters on a logMAR chart) for patients with "on-chart" VA at baseline, or an improvement from "off-chart" VA to at least logMAR 1.6 (equal to one line on-chart) for patients with off-chart VA at baseline.

In the RHODOS trial, a higher proportion of patients in the idebenone group (ITT: 30.2%; n=16) than in the placebo group (ITT: 10.3%, n=3) showed CRR from baseline, but the difference between groups was not statistically significant (p=0.056).

In the EAP, of the 87 patients included in the efficacy population, 40 patients (46.0%) (by eyes, 67/173; 38.7%) had CRR from nadir to the last observation visit. The average magnitude of recovery, defined based on a patients' best recovering eye, corresponded to 22 letters (0.45 logMAR) on the ETDRS chart at the initial observation of CRR and increased with prolonged treatment to 36 letters (0.72 logMAR) at the last observation.

In LEROS, the proportion of eyes that achieved CRR of VA from baseline at 12 months was reported for patients who started treatment with idebenone ≤1 year after the onset of symptoms compared to eyes in the matching external natural history control group.

[redacted]

[redacted]

[redacted]

[redacted].

The EAG has concerns about CRR relating to the extent to which it reflects the effect of treatment with idebenone. The EAG notes the considerable proportion with CRR from nadir in CaRS-I was achieved without receiving treatment and could therefore have been a result of spontaneous recovery. Thus, the EAG has concerns about the extent to which CRR constitutes a good indication of treatment effectiveness of idebenone. See Section 4.2.4 for further discussion of CRR.

3.3.3 Comparison across studies

Outcomes demonstrating change in logMAR (change in best VA, best recovery of logMAR VA in either right or left eye, CRR) from the RHODOS trial, EAP, and LEROS discussed previously are presented in Table 16 below. Where change from baseline scores are reported in the studies, a positive logMAR value (showing an increasing logMAR) indicated worsening and negative logMAR value (showing a decreasing logMAR) indicated improvement.[41]

Table 16. Visual acuity outcome data (adapted from Table 10 from company's initial clarification response)

Outcome (95% CI) [equivalent EDTRS letters]		RHODOS		EAP		LEROS	
		Idebenone	Placebo	LHON population	Efficacy population	ITT	NH matched comparator
	N	53	29	105	87	196	106
	Timepoint	Week 24		Final analysis time-point		Month 24	
Best logMAR at baseline, mean (95% CI)	Baseline	1.61±0.64	1.57±0.61	1.16 (–0.18 to 1.80)	1.23±0.52 (–0.18 to 1.80)	[redacted]	—
Best logMAR at final visit	Final analysis time-point	—	—	1.09±0.66 (–0.18 to 1.80)	1.19±0.63 (–0.16 to 1.80)	[redacted]	—
Change in best VA (from baseline)	Final analysis time-point	–0.035 (–0.126 to 0.055) [+1 letter]	0.085 (–0.032 to 0.203) [–4 letters]	—	—	N= 70 2nd eye onset ≤1 year: –0.09 min –1.78, max 1.84 N=55 2nd eye onset >1 year: –0.19 min –1.24, max 0.12	Data only reported for individual eyes
Best recovery of logMAR visual acuity in either right or left eye (from baseline)	Final analysis time-point	–0.135 (–0.216 to –0.054)	–0.071 (–0.176 to 0.034)	Not an outcome measure	Not an outcome measure	Not an outcome measure	Not an outcome measure

		[+6 letters]	[+3 letters]				
CRR (from baseline)	Final analysis time-point	Patients:16, 30.2%	Patients: 3, 10.3%	Patients: 42, 40.00%	Patients: 31, 35.63%	Eye onset ≤ 1 year: N=44, 40.4% Eye onset>1 year N=33 (32.4%)	—
CRR (from nadir)	Final analysis time-point	Not an outcome measure	Not an outcome measure	53, 50.5%	40, 46.0%	Eye onset ≤ 1 year: N=53, 48.6% Eye onset >1 years: N=37, 36.3%	—

Abbreviations: CRR, clinically relevant recovery; ETDRS, Early Treatment Diabetic Retinopathy Study; ITT, intent-to-treat; logMAR, logarithm of the minimum angle of resolution; NH, natural history; VA, visual acuity

3.3.4 Subgroup analyses

Various subgroup analyses of data had been carried out to provide additional information on the effect of idebenone on VA: the subgroup of patients with logMAR <1 at baseline compared to patients with logMAR ≥1 from the RHODOS trial; patients with disease duration ≥1 year and disease duration <1 year; patients with different mtDNA mutations; and patients with discordant VA at baseline. These were presented in Sections B.2.6 and B.2.7 in the CS. Overall, the EAG notes that due to the rarity of LHON and the limited number of patients for each subgroup, it is difficult to draw robust conclusions about the effect of idebenone from any of the presented analyses.

In addition, the EAG notes that disease duration (≥1 year vs <1 year) is a potentially problematic dichotomisation as according to input from EAG clinical experts, it is noted that the majority of patients may have already hit nadir within the first year of onset and a 'nadir' health state may also be a prognostic factor impacting disease severity and confounding with the treatment effect.

3.3.5 Quality of life

Health-related quality of life (HRQoL) data were available from the RHODOS trial and RHODOS-OFU. These were obtained using the Visual Function (VF)-14 tool, the Clinician's Global Impression of Change (CGIC) score and the visual analogue scale (VAS). As discussed further in Section 4.2.6, HRQoL data derived from the RHODOS and RHODOS-OFU trial were not used in the economic model.

Over the 24-week follow-up of RHODOS, only small changes were observed in VF-14 and the difference between treatment groups in change of VF-14 score was not statistically significant (estimated mean treatment difference – 1.37; 95% CI: –6.25 to 3.51; p = 0.577). VF-14 data were available from 57 patients taking part in the RHODOS-OFU. The overall changes between VF-14 score recorded during RHODOS and RHODOS-OFU were small and differences between idebenone and placebo groups were not statistically significant. There was a small worsening in HRQoL in the idebenone group (–1.7%), whereas there was a small improvement in the placebo group (2.4%) for the entire period between RHODOS baseline to RHODOS-OFU (the between group difference was not statistically significant, p = 0.205).

Although statistical analysis on CGIC scores was not reported, at week 24 of RHODOS, 12 patients (22.6%) in the idebenone group and 7 patients (24.1%) in the placebo group from the ITT population had an improvement in overall CGIC scores. A total of 43 patients (81.1%) in the idebenone group and 24 patients (82.8%) in the placebo group reported experiencing less fatigue or no change in fatigue levels. At week 24, patients in both treatment groups reported minimally elevated energy levels assessed by the VAS (0.37 mm for idebenone and 2.17 mm for placebo) with no statistically significant difference between the treatment groups (estimated mean treatment difference –1.80; 95% CI: –11.37 to 7.77; p = 0.709).

The EAG and its clinical experts partially agree with the company's conclusion that the duration of the RHODOS trial (24 weeks) may not have been long enough to show the treatment benefit of idebenone. In addition, the EAG considers it likely that a larger sample size with a longer follow-up would be required to allow for a minimum clinically important change in VA to be detected.

3.3.6 Safety

Adverse event (AE) data are available from RHODOS, LEROS, the EAP and PAROS. Few safety data were available for placebo or untreated patients, as RHODOS was the only RCT, and safety data were not collected in the CaRS natural history studies.

The frequency of AEs reported in RHODOS and LEROS are presented in Table 17. In RHODOS, the proportion of patients experiencing AEs was similar between the idebenone and placebo groups. A slightly higher proportion of participants in the idebenone arm reported nasopharyngitis (idebenone: 25.5%; placebo: 16.7%); cough (idebenone: 10.9%; placebo: 0%); dizziness (idebenone: 5.5%; placebo: 0%); and left ventricular hypertrophy (idebenone: 7.3%; placebo: 0%). The number of idebenone treated individuals experiencing AEs in LEROS was similar in LEROS compared to RHODOS, with a small but expected increase in the number of investigations in LEROS, given the longer duration of follow up.

Table 17. Number of people experiencing at least one adverse event in RHODOS and LEROS

	RHODOS			LEROS
N (%) subjects	Idebenone 900 mg/day (N=55)	Placebo (N=30)	All Subjects (N=85)	Idebenone 900 mg/day (Safety Population) (N=198)
Timepoint	Through Visit 6 (28 to 35 days after drug discontinuation)			Through study completion (average of 24 months)
AE definition	Treatment-emergent AEs by Preferred Term reported by at least 2 patients in either arm, MedDRA 13.0, N (%)			Treatment-emergent AEs by Preferred Term reported by ≥5% patients in LEROS, or by at least 2 patients in a RHODOS arm, MedDRA 24.0, N (%)
N (%) with at least 1 severe adverse event	2 (3.6)	0	2 (2.4)	13 (6.6)
Cardiac disorders				
Left ventricular hypertrophy	4 (7.3)	0	4 (4.7)	NR
Gastrointestinal disorders				
Upper abdominal pain	3 (5.5)	3 (10.0)	6 (7.1)	13 (6.6)
Constipation	2 (3.6)	3 (10.0)	5 (5.9)	2 (1.0)
Diarrhoea	5 (9.1)	3 (10.0)	8 (9.4)	19 (9.6)
Flatulence	0	2 (6.7)	2 (2.4)	NR
Vomiting	4 (7.3)	2 (6.7)	6 (7.1)	6 (3.0)
Nausea	NR	NR	NR	15 (7.6)
Infections and infestations				
Gastroenteritis	1 (1.8)	2 (6.7)	3 (3.5)	4 (2.0)
Influenza	6 (10.9)	3 (10.0)	9 (10.6)	8 (4.0)

Nasopharyngitis	14 (25.5)	5 (16.7)	19 (22.4)	33 (16.7)
Sinusitis	1 (1.8)	2 (6.7)	3 (3.5)	7 (3.5)
Investigations				
Alanine aminotransferase increased	1 (1.8)	3 (10.0)	4 (4.7)	17 (8.6)
Blood cholesterol increased	0	2 (6.7)	2 (2.4)	4 (2.0)
Blood creatine phosphokinase increased	1 (1.8)	2 (6.7)	3 (3.5)	15 (7.6)
Blood triglycerides increased	6 (10.9)	3 (10.0)	9 (10.6)	5 (2.5)
Gamma-glutamyl transferase increased	0	5 (16.7)	5 (5.9)	10 (5.1)
Aspartate aminotransferase increased	NR	NR	NR	14 (7.1)
Musculoskeletal and connective tissue disorders				
Arthralgia	0	2 (6.7)	2 (2.4)	3 (1.5)
Back pain	4 (7.3)	2 (6.7)	6 (7.1)	9 (4.5)
Nervous system disorders				
Dizziness	3 (5.5)	0	3 (3.5)	5 (2.5)
Headache	13 (23.6)	6 (20.0)	19 (22.4)	37 (18.7)
Respiratory, thoracic, and mediastinal disorders				
Cough	6 (10.9)	0	6 (7.1)	12 (6.1)
Oropharyngeal pain	5 (9.1)	3 (10.0)	8 (9.4)	14 (7.1)
Skin and subcutaneous tissue disorders				
Pruritus generalised	1 (1.8)	2 (6.7)	3 (3.5)	NR
Rash	2 (3.6)	2 (6.7)	4 (4.7)	6 (3.0)

Abbreviations: AE, adverse event; MedDRA, Medical Dictionary for Regulatory Activities.

Sources: CS Table 22; LEROS clincialtrials.gov record[39]

Limited safety data were presented from the RHODOS-OFU single visit: the CS quoted pages 69 and 70 of the EMA European Public Assessment Report of idebenone, which stated that: "*Of the 60 patients included in the Safety Population of RODOS-OFU, there was one SAE of hypertensive emergency experienced on the day of the RHODOS-OFU visit, which was over 3 years after completing. The investigator considered this event not related to study drug received in RHODOS. No other relevant safety findings were derived from RHODOS-OFU.*"[24]

In PAROS, the prospective non-interventional post-authorisation safety study of idebenone, the following primary endpoints were measured:

- Frequency of AEs of special interest ([redacted])
- Frequency and nature of AEs and serious AEs;
- Frequency and nature of adverse drug reactions and serious adverse drug reactions.

The frequency of AEs of special interest and the frequency of AEs and serious AEs observed in PAROS were reported in Table 11 and Table 12 of the company response to clarification, respectively. These results are in-line with the safety findings of RHODOS and LEROS.

Safety data were also available from the EAP safety population (N=111). [redacted] The EAG considers these data to be in line with RHODOS and LEROS.

AEs were not included in the economic model given that most AEs experienced were considered mild in all safety studies conducted.

3.3.6.1 EAG summary of safety data

The EAG notes the overall incidence of AEs across the idebenone clinical trial programme and real-world evidence studies was low, and few AEs were classed as severe. For the only data set in which a

placebo cohort were available, RHODOS, the proportion of patients experiencing AEs was similar in the idebenone arm compared to the placebo arm, with the potential exceptions of nasopharyngitis (idebenone: 25.5%; placebo: 16.7%); cough (idebenone: 10.9%; placebo: 0%); and dizziness (idebenone: 5.5%; placebo: 0%). These AEs were not classed as severe, and the EAG considers that even if the increases observed in RHODOS were observed in clinical practice, they would be unlikely to have a meaningful impact on the cost-effectiveness of idebenone.

The EAG notes that there is an absence of long-term placebo-controlled data on the safety of idebenone for treating LHON, but the EAG was reassured that the proportion of patients experiencing AEs did not meaningfully increase in the longer term LEROS clinical trial and the EAP and PAROS observational studies.

3.3.6.2 Left ventricular hypertrophy

In RHODOS, four (7.3%) idebenone patients compared to 0 (0%) placebo patients experienced left ventricular hypertrophy. The EAG notes this was explored by the EMA in the European Public Assessment Report, which considered that:[24]

- In RHODOS, only one case of left ventricular hypertrophy was considered related to idebenone treatment;
- All left ventricular hypertrophy events were non-serious and were reported by the same investigational site;
- The diagnoses was not supported by clinical or ultrasound evidence;
- The incidence of ECG findings suggestive of left ventricular hypertrophy developing after initiation of the study treatment was lower in the idebenone group (7.27%) than in the placebo group (13.33%);
- When considering data also from research in patients treated with idebenone for Friedreich's ataxia, there was no demonstrated signal of any ECG abnormality in heart rate for individuals treated with idebenone.

The EAG further notes that left ventricular hypertrophy was not reported as an observed AE in LEROS, PAROS, or the EAP. The EAG, therefore, considers it unlikely that idebenone is related with the development of left ventricular hypertrophy in people with LHON.

3.3.7 Discussion of clinical effectiveness evidence

As discussed further in Section 4.2.4, the EAG has concerns over the preference of the EAP over the LEROS trial data to inform the clinical and cost-effectiveness of idebenone beyond 6 months (6 to 36 months).

The company argues that the EAP should be preferred over LEROS due to heterogeneity in the proportion of males between the LEROS trial [redacted] and the RHODOS trial (85.9%) that was used to inform the clinical and cost-effectiveness of idebenone for the first 6 months of treatment. The EAG notes that although the proportion of males was more comparable between the RHODOS trial and the EAP (82%), males in the LEROS trial still constituted [redacted] patients. Taking this into consideration in addition to that it is unclear if sex is a prognostic factor impacting disease severity,[42] and that there was no substantial difference in outcome data between the LEROS trial, the RHODOS trial and EAP, the EAG has concerns over the rationale for the company's preference for the EAP over LEROS.

The EAG notes that the genetic mutation distribution of the RHODOS trial population consisting of people carrying three mutations (m.11778G>A [67.3%], m.14484T>C [20%], m.3460G>A [12.7%]) was more aligned with the EAP population compared to LEROS population, which consisted of patients from a wider range of LHON mutations (m.11778G>A [redacted] m.14484T>C [redacted] m.3460G>A [redacted] Negative [redacted] Other [redacted]) than the EAP study (m.11778G>A [62.1%], m.14484T>C [18.4%] m.3460G>A [19.5%]).

In addition, the EAG notes that the EAP provides more longer-term data for up to 36 months compared to the LEROS trial with data for up to 24 months, but that these data are limited with LEROS (N=199) providing data for a larger data set than the EAP (N=87), potentially making it a better choice to inform the long-term effectiveness of idebenone. The EAG also notes that time since onset in the RHODOS and LEROS trial is comparable (≤5 years) but differs in the EAP including patients with onset of vision loss in the second eye less than 12 months. EAG clinical experts agreed that time since onset is an important prognostic factor for that can impact treatment effectiveness. The EAG's concerns over the difference in time since onset between the data sets also include the greater chance of spontaneous recovery present during the first year of onset as highlighted by the clinical experts, which could introduce further bias in the interpretation of the results.

3.4 Critique of the indirect comparison treatment comparison

3.4.1 Trials informing the indirect treatment comparison

A direct head-to-head comparison of idebenone and SoC for 6 months of treatment is available from the RHODOS trial. After this, no RCT data comparing long-term treatment with idebenone and SoC are available. In the company base case, the long-term treatment effects of idebenone and SoC are modelled using two unmatched populations: the EAP population for idebenone and the CaRS-I and CaRS-II natural history (NH) populations for SoC. Such a comparison is at high risk of bias due to imbalances in prognostic factors between patients in the EAP and the CaRS studies, for example, differences in the prevalence of each major three mutation type (Section 3.2.5).

The EAG considers that matching the idebenone and SoC cohorts would provide a less biased method to model the long-term treatment effect of idebenone compared to SoC, but notes no matched control analyses were provided in the original CS. Following a request by the EAG at the clarification stage, the company provided a propensity-score matching (PSM) analysis of changes in patient's best visual acuity between LEROS and CaRS-I and CaRS-II at Month 24.

3.4.2 Statistical methods

The company's PSM analyses compared LEROS ITT patients at 24 months with SoC CaRS-I and CaRS-II SoC patients. The following prognostic factors were included in the calculation of the propensity score:

- Sex;
- Age at first symptom onset;
- Genotype;
- Months since the most recent symptom onset;
- Months since first symptom onset;
- Number of symptomatic eyes at baseline;
- Baseline logMAR for the left eye; and
- Baseline logMAR for the right eye.

Rather than matching patients at baseline and then including all subsequent follow-up visits in the analysis, the PSM analyses were conducted using a single baseline and 24-month visit window only. The following process was conducted to match patients:

- A subset of NH patients from CaRS-I and CaRS-II were selected who had a pair visits 24 months apart (within a window of 3 months) – matching the "baseline" and 24-month visit in LEROS. This created a total sample of 84 patients with 270 potential visit pairs 24 months apart for matching;
- Through PSM, 68 of the 84 available NH patients to be matched to 68 of 125 LEROS ITT patients.

PSM was conducted using a nearest neighbour approach with a calliper width of 0.2 time the standard deviation of the logit of the PS. The PSM analyses were implemented in SAS 9.4.

The baseline characteristics of the matched patients are displayed in Table 18.

Table 18. Baseline characteristics of idebenone treated patients (LEROS ITT) matched to SoC treated patients (CaRS-I and CaRS-II).

Baseline Characteristic	Matched SoC (CaRS) N = 68	Matched Idebenone (LEROS) N = 68
Gender		
Female	11 (16.2%)	15 (22.1%)
Male	57 (83.8%)	53 (77.9%)
Genotype		
m.11778G>A	40 (58.8%)	[redacted]
m.3460G>A	14 (20.6%)	[redacted]
m.14484T>C	9 (13.2%)	[redacted]
Other	5 (7.4%)	[redacted]
Age at 1st symptom onset		
Mean ± SD	26.2±15.3	29.7±13.6
Median (Q1-Q3)	21.0 (15.5 to 36.5)	26.7 (19.1 to 39.1)
Eyes affected at baseline		
1	0 (0.0%)	4 (5.9%)
2	68 (100.0%)	64 (94.1%)
Months since 1st symptoms onset		
Mean ± SD	18.2±22.3	18.1±16.6
Median (Q1-Q3)	13.2 (5.5 to 21.4)	11.8 (6.1 to 23.8)
Min - Max	0.0 to 134.1	0.3 to 58.3
Months since most recent symptoms onset		
Mean ± SD	17.1±21.9	16.3±16.5
Median (Q1-Q3)	11.3 (4.1 to 20.4)	9.4 (4.6 to 23.5)

Min - Max	0.0 to 134.1	0.0 to 57.6
Baseline best VA logMAR		
Mean ± SD	1.19±0.53	1.16±0.60
Median (Q1-Q3)	1.30 (0.90 to 1.80)	1.31 (0.69 to 1.65)
Min - Max	–0.20 to 1.80	–0.12 to 1.80
Baseline best VA		
Light Perception	0 (0.00%)	1 (1.47%)
Hand Motion	4 (5.88%)	6 (8.82%)
Counting Fingers	13 (19.12%)	9 (13.24%)
logMAR >= 1.3 and < 1.7	20 (29.41%)	20 (29.41%)
logMAR >= 1.0 and < 1.3	13 (19.12%)	9 (13.24%)
logMAR >= 0.6 and < 1.0	10 (14.71%)	9 (13.24%)
logMAR >= 0.3 and < 0.6	4 (5.88%)	4 (5.88%)
logMAR < 0.3	4 (5.88%)	10 (14.71%)

Abbreviations: ITT, intention to treat; logMAR, Logarithm of the minimum angle of resolution; NH, Natural history; Q, Quartile; SD, Standard deviation; SoC, standard of care; VA, Visual acuity

Source: Company response to clarification Table 5.

The EAG considers the baseline characteristics of matched cohorts to be reasonably balanced, and considerably less imbalanced than the original unmatched samples (Section 3.2.5). The EAG notes two small remaining imbalances in patient baseline characteristics:

- The age of first symptom onset is younger for the SoC cohort than the idebenone cohort, which is likely associated with a greater probability of spontaneous recovery in the SoC cohort;
- The prevalence of the milder T14484C genotypes is slightly higher in the idebenone cohort than the SoC cohort, which is likely associated with a greater probability of spontaneous recovery in the idebenone cohort.

The company implemented two analysis of covariance (ANCOVA) models to perform a comparison of change in best logMAR VA between matched idebenone and SoC patients. The first model included all patients and had treatment as a sole predictor. The second model included treatment, genotype and a treatment-by-genotype interaction as predictors, and limited the analysis population to patients with one of the three major genotypes.

3.4.3 LEROS ITT vs CaRS-I and CaRS-II ITC results at 24 months

Full results of the PSM comparison between LEROS and the CaRS studies are presented in the company response to clarification question A2. Here, the EAG focuses on the most relevant outcome for the economic model, change in best logMAR VA.

The results of the change in best logMAR VA model with treatment as a single predictor are presented in Table 19. There was no statistically significant difference in the change in best logMAR VA at 24 months between idebenone, –0.13 (95% CI: –0.27 to 0.02), and matched NH controls, –0.11 (95% CI: –0.26 to 0.04), with similar point estimates and confidence intervals between the groups.

Table 19. PSM analysis of change in best VA at 24 months between idebenone treated patients (LEROS ITT) matched to SoC treated patients (CaRS-I and CaRS-II). Adapted from company response to clarification Table 8.

Treatment	Change in best VA at 24 months, logMAR LS-Means (95% CI)	LS-Means p-value
Idebenone	–0.13 (–0.27 to 0.02)	0.097
SoC	–0.11 (–0.26 to 0.04)	0.152
Difference	–0.02 (–0.23 to 0.19)	0.871

ANCOVA with treatment as covariate

Abbreviations: CI, Confidence interval; ITT, intention to treat; LS, Least squares; logMAR, Logarithm of the minimum angle of resolution; SoC, Standard of care; VA, Visual acuity

Source: Company response to clarification Table 8

Table 20 presents the results of the alternative change in best logMAR VA model with treatment, genotype and a treatment-by-genotype interaction as predictors, and the limited analysis population to patients with one of the three major genotypes. Conditional on genotype and the interaction between genotype and treatment, there was no statistically significant difference in the change in best logMAR VA at 24 months between idebenone, –0.14 (95% CI: –0.29 to 0.02), and matched NH controls, –0.24 (95% CI: –0.41 to –0.07), although the point estimate numerically favoured SoC.

Table 20. PSM analysis of change in best VA at 24 months between idebenone treated patients (LEROS ITT, major 3 genotypes only) matched to SoC treated patients (CaRS-I and CaRS-II). Adapted from company response to clarification Table 9.

Treatment	Major 3 genotypes	Change in best VA at 24 months, logMAR LS-Means	LS-Means p-value
Idebenone	_	–0.14 (–0.29 to 0.02)	0.085
NH	_	–0.24 (–0.41 to –0.07)	0.007

Difference	_	0.10 (–0.13 to 0.34)	0.381
Idebenone	G11778A	–0.07 (–0.25 to 0.11)	0.436
Idebenone	G3460A	0.17 (–0.13 to 0.46)	0.263
Idebenone	T14484C	–0.50 (–0.82 to –0.19)	0.002
NH	G11778A	0.07 (–0.11 to 0.25)	0.471
NH	G3460A	–0.30 (–0.61 to 0.00)	0.049
NH	T14484C	–0.48 (–0.86 to –0.10)	0.013
Difference	G11778A	–0.14 (–0.39 to 0.12)	0.289
Difference	G3460A	0.47 (0.05 to 0.89)	0.029
Difference	T14484C	–0.02 (–0.52 to 0.47)	0.923

ANCOVA with treatment and mutation as covariates type 3 test of fixed effects p-values:

- Treatment p = 0.381
- Genotype p = 0.028
- Treatment*Genotype p = 0.534

Abbreviations: CI, Confidence interval; ITT, intent-to-treat; LS, Least squares; logMAR, Logarithm of the minimum angle of resolution; NH, Natural history; VA, Visual acuity

Source: Company response to clarification Table 9

An analysis of the subgroup of LEROS ITT patients matched to CaRS patients in the subacute phase of LHON, defined as only those visits where the most recent symptom onset occurred within the last year, was also presented in the company response to clarification question A2. The EAG notes the result of this smaller sample analysis was in-line with the full population: no statistically significant differences between idebenone and SoC in the change in best logMAR VA at Month 24.

3.4.4 EAG critique

The results of these PSM analyses do not provide strong evidence of a clinically meaningful long-term treatment benefit of idebenone over SoC for the treatment of LHON, and the company did not provide a detailed interpretation of the results of the PSM analyses. The EAG notes that:

- The PSM analysis is currently the only available matched cohort analysis of the effects of long-term idebenone treatment compared to SoC;
- These data do not provide strong evidence of a clinically meaningful long-term treatment benefit of idebenone over SoC for the treatment of LHON;
- There are several limitations to the PSM analysis – which was conducted over a short time frame during the clarification stage – but the EAG considers the analysis to be reasonably unbiased;

- The face validity of the point estimates of a negligible (Table 19) or negative (Table 20) long-term treatment benefit of idebenone over SoC may be low, and suggests this PSM analysis underestimates the long-term treatment effect of idebenone, despite being considered by the EAG as the most appropriate analysis of the comparative effectiveness of long-term idebenone treatment compared to SoC that is currently available.

The EAG completed a quality of effectiveness estimates from non-randomised studies (QuEENS) checklist for the PSM analyses, which is presented in Appendix 8.1.[43] The EAG notes that several areas highlighted as lower quality for the PSM analyses, such as comparing different analysis methods, were likely infeasible during the 4-week window the analyses were conducted. The EAG is most concerned that only a limited amount of CaRS follow-up data were included in the analyses by choosing to only analyse a single visit pair, rather than all available data, for SoC patients. The EAG also notes that this prevented the data from the matched-control analysis being used in the economic model, as the company explained in response to further clarification:

> *"The company [performed] the matching algorithm for individual patients as per the request for A2, however the per cycle transition counts cannot be derived as the same patient will not be followed over the trial duration and therefore, their movement across health states cannot be accurately captured (matching algorithm was performed de novo at each time point, implying that different patient subsets from the CaRS trial were included at each timepoint)."*

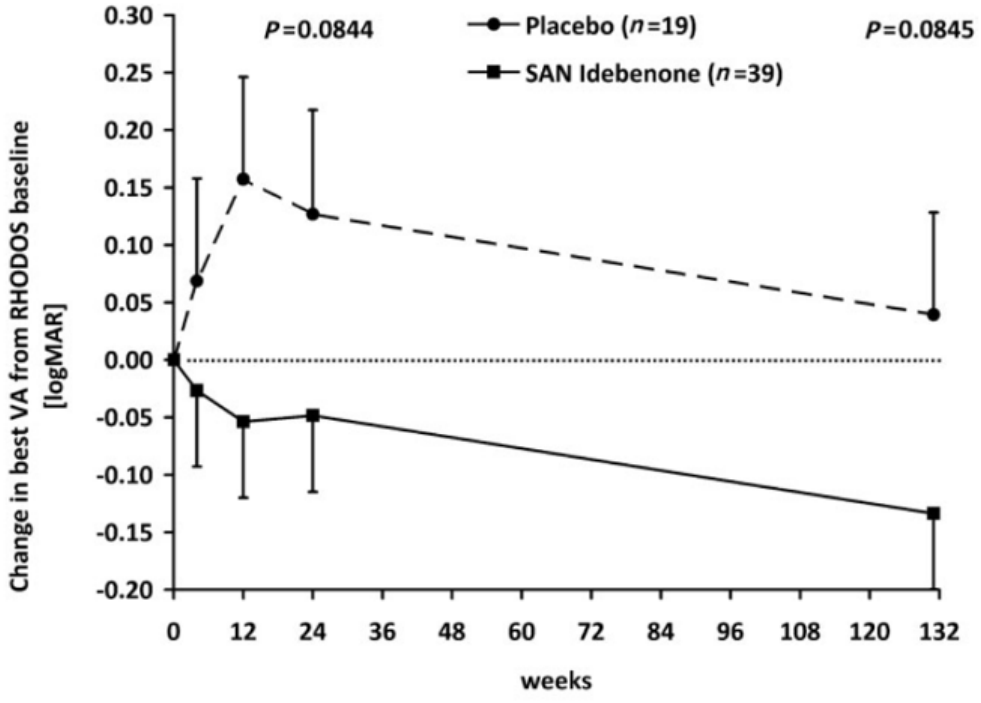
The EAG's preferred approach would have been to match patients between LEROS and CaRS at baseline and use all available follow-up data in analysis. The EAG considers that, since the median time between visits was 11.7 months in CaRS, restricting the analysis set to visit pairs 24 months apart likely does not make best use of the available data. Nevertheless, at the current time, the PSM analysis presented in the company's response to clarification is the only available long-term matched-control comparison of changes in patients' best logMAR VA over time between idebenone and SoC, which does not provide strong evidence of a clinically meaningful long-term treatment benefit of idebenone over SoC for the treatment of LHON.

The EAG notes that a single long-term follow-up datapoint is available from N=58 participants from the RHODOS trial in the OFU visit, presented in Section 3.3.1.2. This provides data on the long-term outcomes of people previously treated with idebenone (for 6 months) compared to SoC for a

duration of 6 months only, up to Week 132. These data are reproduced in Figure 6. Through Week 132, the estimated difference in best VA between patients treated with 6 months of idebenone compared to SoC was −0.173 (95% CI: −0.370 to 0.024), equivalent to 8 ETDRS letters. This difference was not statistically significant (p = 0.0845). The EAG notes that:

- The RHODOS RCT and OFU data are more consistent with a positive long-term treatment effect of idebenone compared to SoC compared to the PSM analyses;
- The RHODOS OFU visit data may slightly overestimate the long-term treatment effect of 6-months of idebenone treatment compared to SoC due to a selection bias at patient entry favouring idebenone:
    - In the full RHODOS RCT ITT population, the between group difference (idebenone – SoC) at Week 24 was −0.120 (95% CI: −0.255 to 0.014), equivalent to 6 ETDRS letters;
    - In the subgroup of patients comprising the RHODOS OFU population, the between group difference (idebenone – SoC) at Week 24 was −0.175 (95% CI: −0.375 to 0.024), equivalent to 8 ETDRS letters. That is, the patient population entering the RHODOS OFU visit had a larger treatment effect at Week 24 than the ITT population from which it was sampled.

Figure 5. Change in visual acuity over time for the best visual acuity (logMAR), RHODOS OFU cohort (Reproduced from CS Figure 15)



Data are estimated means ± SEM from MMRM, based on the change from baseline (in weeks) and plotted for the two treatment groups as defined in the RHODOS study. No treatment was given between Week 24 and Week 131. Worsening/improvement of visual acuity is indicated as positive/negative values in change of logMAR. A difference of logMAR 0.1 corresponds to five letters or one line on the Early Treatment Diabetic Retinopathy Study chart. The P-values are given for the difference between treatment groups.
Source: Reproduced from CS Figure 15. Klopstock T et al, 2013 (19)
Abbreviations: logMAR, Logarithm of the minimum angle of resolution; MMRM, Mixed-model of repeated measures; SEM, Standard error of the mean; VA, Visual acuity

Hence, at the time of the EAG report, the EAG notes three available approaches to modelling the long-term treatment effect of idebenone vs SoC:

- A PSM analysis of LEROS and the CaRS studies that does not provide strong evidence of a clinically meaningful long-term treatment benefit of idebenone over SoC for the treatment of LHON, but is the only source of matched-control data for patients treated with idebenone for > 6 months;
- The RHODOS OFU visit, providing data on the Week 132 outcomes of patients previously treated with idebenone or SoC for 6 months, followed by SoC for both treatment arms up to Week 132; and
- An unmatched comparison of the EAP or LEROS and the CaRS studies, employed in the current company base case, that is at high risk of bias due it being a naïve comparison with imbalances in prognostic factors between patient cohorts.

The EAG notes that a matched-control analysis of individual eyes between LEROS and CaRS-I and CaRS-II patients was presented in the LEROS CSR, and summarised in the CS Appendix M. As highlighted in Clarification Question A1, the EAG did not consider this analysis appropriate to inform a cost-effectiveness analysis of idebenone as the analysis focused on outcomes within individual eyes, rather than at the more relevant level of the individual patient. Moreover, the EAG noted in Clarification Question A1 (f) that the “matching” procedure only matched CaRS patients’ visit pairs times to the average time since onset of symptoms at baseline calculated for LEROS. The EAG does not consider that this procedure matches patients on key prognostic factors, such as LHON genotype baseline VA, and is therefore at high risk of bias. The EAG notes that the PSM analyses provided at the individual patient level follows a similar approach to the “by eye” analysis originally presented, but also matches patients based on these key prognostic factors.

3.5 Conclusions of the clinical effectiveness section

The EAG concludes that although the RHODOS trial provided randomised controlled evidence on the efficacy of idebenone compared to standard of care (SoC) for up to 6 months (24 weeks), evidence on the long-term effects of idebenone has been limited.

No RCT data comparing long-term treatment with idebenone and SoC are available as evidence has been limited to observational data with inherent limitations such as the open-label and uncontrolled nature of the data collection in the EAP, the retrospective analysis of patient records in the CaRS studies and the considerable loss of data during follow-up across sources of data. Long-term effectiveness beyond 6 months was modelled using two unmatched populations: the EAP population for idebenone and the CaRS-I and CaRS-II natural history populations for SoC. The analysis used to inform on the long-term efficacy and effectiveness of treatment with idebenone is considered by the EAG to be inadequate due to imbalances in prognostic factors in the study populations such as the prevalence of different mutation types and at high risk of bias. Therefore, the long-term efficacy of idebenone remains uncertain.

Further uncertainties arise in the interpretation of long-term evidence from the RHODOS-OFU trial, which was based on a single visit (approximately 30 months after completion of the RHODOS trial), where patients had not been receiving further treatment with idebenone between the completion of RHODOS and their follow-up visit. Although improvements in VA observed for idebenone and placebo after a mean time of 30 months (2.5 years) from week 24 of the RHODOS trial, suggested the benefit of 6 months treatment with idebenone is maintained after treatment is stopped, the lack of intermediate data collection between the end of RHODOS and OFU visit led to uncertainties in the interpretation of results.

Moreover, the EAG notes that overall, in the current evidence, efficacy of idebenone has been documented for up to 5 years after onset and this is highlighted in the CS, with the EAP only including patients with onset of vision loss in the second eye less than 12 months and RHODOS only including patients with onset of vision loss ≤5 years. However, it is noted that the majority of the prevalent population in UK clinical practice will have disease onset >5 years and the EAG has concerns about applicability of results from the trial populations to the prevalent population in UK clinical practice.

Furthermore, there were additional areas of potential bias influencing the interpretation of the evidence. The exclusion of one patient from the placebo group in the mITT analysis of the RHODOS trial, resulted in a considerable increase in the between-group difference, creating uncertainty over the robustness of the RHODOS data. Additionally, given the potential of spontaneous recovery in LHON, there was a risk of overestimating the effect of idebenone.

Finally, it has been noted in existing literature and by clinical experts that the magnitude of the benefit of treatment with idebenone may vary between different subgroups of patients, for example between people with different LHON-causative mDNA mutations. However, limited sample sizes resulting from the clinical trials currently available, did not provide sufficient power to allow for the detection of subgroup effects and support meaningful conclusions about potential differences in the magnitude of the effect of treatment with idebenone between different groups of patients.

4 Cost effectiveness

Table 21 below presents the incremental cost-effectiveness results of the company's updated (i.e., post clarification) base case results.

Table 21. Company's base case results

Interventions	Total Costs (£)	Total LY	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
Deterministic results							
SoC	[redacted]	[redacted]	[redacted]	-	-	-	-
Idebenone	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	18,758
Probabilistic results							
SoC	[redacted]	-	[redacted]	-	-	-	-
Idebenone	[redacted]	-	[redacted]	[redacted]	-	[redacted]	19,272

Abbreviations: ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality-adjusted life-year; SoC, standard of care.

*Probabilistic results do not account for idebenone and SoC treatment effectiveness uncertainty

4.1 EAG critique of the company's systematic literature review for cost effectiveness evidence

The company carried out two systematic literature reviews (SLRs) to identify published studies to inform the cost-effectiveness evaluation of idebenone. One SLR aimed to capture publications relevant to cost-effectiveness and costs and resource use, and the other health-related quality of life (HRQoL) associated with Leber's hereditary optic neuropathy (LHON), not limited by intervention. Searches were initially conducted in October and November 2020 with two updated searches being run in February and March 2023. A summary of the External Assessment Group's (EAG's) critique of the methods implemented by the company to identify relevant publications is presented in Table 22. Due to time constraints, the EAG was unable to replicate the company's searches and appraisal of identified abstracts.

Table 22. EAG critique of SLR methods

Systematic review step	Section of CS in which methods are reported			EAG assessment of robustness of methods
	Cost effectiveness evidence	HRQoL evidence	Resource use and costs evidence	
Search strategy	Appendix G1.1	Appendix G1.1	Appendix G1.1	Appropriate databases and

				HTA bodies searched.
Inclusion/ exclusion criteria	Appendix G1.1	Appendix G1.1	Appendix G1.1	Appropriate. Exclusion limited to diseases other than LHON.
Screening	Appendix G1.2	Appendix G1.2	Appendix G1.2	Appropriate.
Data extraction	Appendix G1.3	Appendix G1.3	Appendix G1.3	Appropriate.
Quality assessment of included studies	Appendix G1.6	Appendix H1.2	Appendix G1.6	Appropriate. Evaluated using Drummond and Efficace check lists.[44, 45]

Abbreviations: CS, company submission; EAG, External Assessment Group; HRQoL, health related quality of life.

The SLRs identified 10 relevant publications, five of which related to health care resource use and five to HRQoL. None, however, were economic evaluations or contained usable utility values or information that could be used to inform the model. Instead, the company used published NICE technology appraisals in related disease areas (specifically retinal dystrophies and macular degeneration) to inform the development of the *de novo* cost-effectiveness model from a separate targeted literature review.[46-49]

4.2 Summary and critique of company's submitted economic evaluation by the EAG

4.2.1 NICE reference case checklist

Table 23 summarises the EAG's appraisal of the company's economic evaluation against the NICE reference case checklist for the base-case analysis, with reference to the NICE final scope outlined in Section 2.3.

Table 23. NICE reference case checklist

Element of health technology assessment	Reference case	EAG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	The major health effects for patients with LHON have been included in the economic model.
Perspective on costs	NHS and PSS	All relevant costs have been included and are based on the NHS and PSS perspective.
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	Cost-utility analysis has been provided by the company with fully incremental analysis.

Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Lifetime horizon (100 years of age).
Synthesis of evidence on health effects	Based on systematic review	The company has performed an appropriate systematic literature review.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	Health outcomes have been expressed in terms of QALYs, within HRQoL values taken from Brown *et al*. 1999 which calculated HRQoL values using TTO and a VF-14 questionnaire.[23]
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	HRQoL values obtained from the RHODOS trial were not used in the model. Instead, utilities were informed using Brown *et al*. 1999 which were derived from general population patients.[23]
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	The utility data used can be considered relevant to the UK, however they are not LHON specific.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	The economic evaluation matches the reference case.
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Many of the costs included in the analysis have been sourced using NHS reference costs.[50] However, health state resource use costs have been sourced using Meads *et al.* 2003,[3] which have no clear relation to NHS and PSS costs.
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	A discount rate of 3.5% has been used for both costs and health effects.

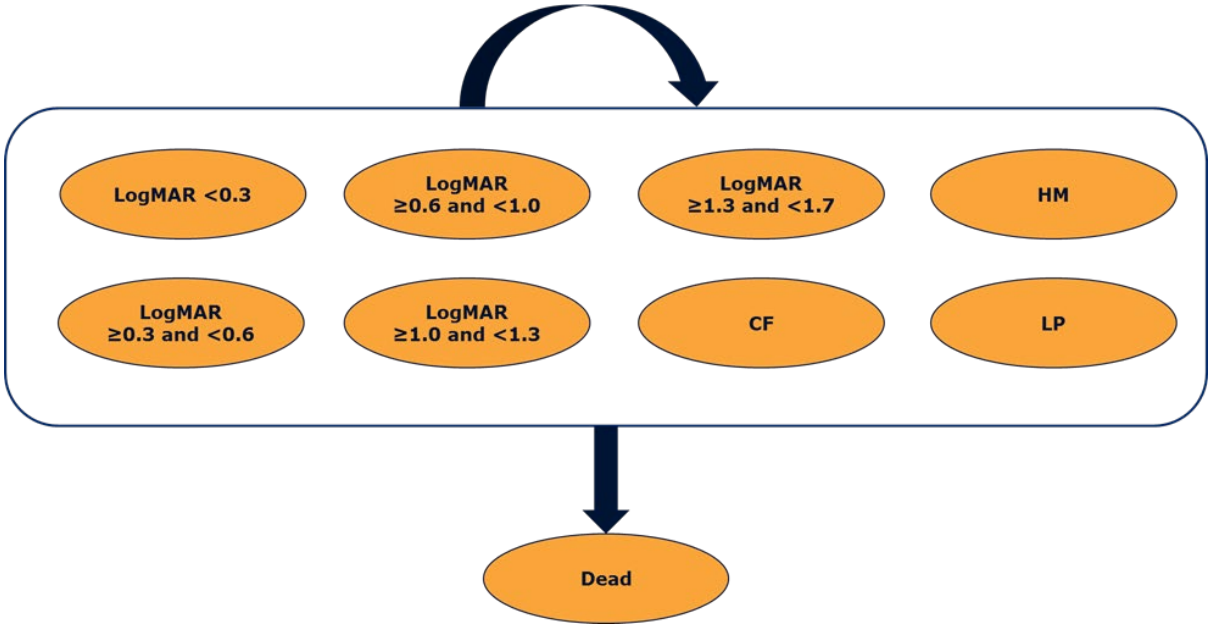
Abbreviations: EAG, External Assessment Group; NHS, national health service; PSS, personal social services; QALY, quality adjusted life year; TTO, time trade off.

4.2.2 Modelling approach and model structure

The company developed a *de novo* Markov model that included eight health states based on visual acuity (VA) expressed in terms of logMAR (logarithm of the Minimum Angle of Resolution), and

death as an absorbing health state (Figure 7). Patients were distributed across the model health states at baseline according to the baseline logMAR distribution of patients in the RHODOS trial.

Figure 6. Company model structure (reproduced from Figure 20 in the CS)



Abbreviations: CF, Counting Fingers; HM, Hand Motion; LP, Light Perception.
NB: CF, HM, and LP correspond to logMAR 2.0, 2.3 and 2.6 in the RHODOS and EAP studies.

The company justified their modelling approach through comparisons to cost-effectiveness models used in previous relevant NICE TAs, specifically HST11 which was conducted in 2019 for voretigene neparvovec in treating inherited retinal dystrophies[46] and NICE TA274[49], TA283[48] and TA298[47] which were conducted in 2013 for ranibizumab across multiple retinal related conditions. A Markov model was used in each approach; however, the models varied considerably between indications, with the HST11 model utilising five health states based on grouped logMAR values according to the American Medicines Association and the NICE TAs using eight or nine health states based on standardised readable EDTRS (Early Treatment of Diabetic Retinopathy Study) bands, which the company's model mirrors.

4.2.2.1 EAG critique

The EAG notes that the company model structure is comparable to NICE TA274[49], TA283[48] and TA298[47] conducted 10 years ago for ranibizumab; however, the EAG considers that a model that groups health states according to key changes in functional sight and HRQoL, similar to the HST11 model,[46] would be more appropriate.

The EAG's clinical experts outlined that patient HRQoL does not perfectly correlate to gain or loss of sight, but instead there are functions and capabilities of sight, which losing or gaining lead to significant changes in HRQoL. Examples of these being the ability to drive (<logMAR 0.3), being eligible for the Certificate of Vision Impairment in England (logMAR >1 [on-chart visually impaired]) and being unable to read any letters on a logMAR chart at six meters (logMAR >1.7 [off-chart visually impaired]). As discussed further in Section 4.2.4, the large number of health states in the economic model also significantly reduces the available patient data to inform each transition probability, leading to health state transitions being impossible and multiple data imputations being required. For example, under both probabilistic and deterministic conditions it is impossible for idebenone treated patients to remain in the Hand Movement health state past cycle 10 (2.5 years) in the company's model. The EAG therefore considers that the company model is flawed and potentially inappropriate for decision making as there is insufficient evidence to support the high number of health states in the economic model given the modest differences in HRQoL and functional capabilities between some of the health states according to the EAG's clinical experts.

While the HST11 model may therefore be preferred to the company model, the health states considered in HST11 conflicted with the opinions of the EAG's clinicians regarding key differences in patient HRQoL according to sight. For example, the HST11 model grouped together patients with logMAR scores of 0 to 1 (Figure 8); however, the EAG's clinical experts argued that a patient with no to limited visual acuities (LogMAR <0.3) will have a significantly higher HRQoL than a patient whose sight has deteriorated to the extent they are no longer able to drive but not considered sight impaired (LogMAR 0.6-1). The EAG's clinical experts also stated that HRQoL would be similar between patients considered off-chart visually impaired (CF, HM and LP) as any sight which may remain is unlikely to provide a level of autonomy.

Following the opinions of the EAG's clinicians, the EAG requested at the clarification stage that the company updated their base case model by grouping the logMAR based health states according to the EAG preferred health states as described in Figure 8. In contrast to the model used in HST11, treatment effectiveness is not capped in the EAG preferred model, allowing patients to have logMAR values and HRQoL utilities more similar to general population estimates. Patients able to drive (limited visual acuities) and unable to drive (moderate visual acuities) are also differentiated. Similarly, logMAR values eligible for the Certificate of Vision impairment in England have been grouped (on-chart visually impaired), and health states unable to read any letters on a logMAR chart (off-chart visually impaired) are also grouped. The EAG additionally considers that that the reduction

in health states also makes the best use of the limited available patient data as it avoids the implausible model transitions exhibited in the company's base-case model.

Figure 7. Company, EAG preferred and HST11 model health states

HST11 health states	Company health states	EAG preferred health states
Moderate visual impairment	LogMAR <0.3	Limited visual impairment
	LogMAR 0.3-0.6	Moderate visual impairment
	LogMAR 0.6-1.0	
Severe visual impairment	LogMAR 1.0-1.3	Visually impaired (on-chart)
Profound visual impairment	LogMAR 1.3-1.7	
CF	CF	Visually impaired (off-chart)
HM, LP	HM	
	LP	

Abbreviations: CF, counting fingers; EAG, external assessment group; HM, hand movement; LP, light perception.

The company did not comply with the EAG's request to update the base case model; however, the company did conduct a scenario using the EAG's proposed model structure. Following the adaptation of the model health states, the health state utility values (HSUVs) and health state resource use estimates were also recalculated to accommodate the changes in model structure, resulting in an increase in the ICER from £18,758 to £27,053. The EAG preferred model structure is assumed in the EAG's base case assumptions.

4.2.3 Perspective, time horizon and discounting

The model cycle length was three months (with a half cycle correction applied) and a lifetime horizon was adopted (up to age 100 years), allowing the model to run for 66 years given a patient starting age of 34 in the model, which was the mean age in RHODOS. The perspective of the analysis was based on the UK NHS and PSS (personalsocial service), with future costs and benefits discounted using an annual rate of 3.5%, as per the NICE reference case.[51]

4.2.3.1 EAG critique

The EAG notes that the half cycle correction applied by the company was calculated as the average of the current and subsequent cycle, applied from the first model cycle (cycle 0); however, the EAG considers the half cycle correction should have been applied to the current and previous cycle from

cycle one onwards. At clarification the company complied with the EAG's request to correct the half cycle correction methodology which led to a £241 increased in the ICER.

4.2.4 Treatment effectiveness

4.2.4.1 Measures of treatment effectiveness and use of the RHODOS trial

The clinical effectiveness of idebenone and SoC (standard of care) treatments were captured by the transitions between health states in the model. Health state transitions probabilities were derived using patient better seeing eye VA observations as this was considered by the company to be the endpoint most relevant to clinical practice and a patient's HRQoL.

The RHODOS study was used to inform the treatment effectiveness for both idebenone and SoC as it was a randomised, doubled-blind, trial comparing idebenone to the current SoC (placebo). However, as the trial length was limited to six months, a period deemed too short to fully demonstrate the full benefits of idebenone by the company, supplementary data were required to model long-term treatment effectiveness.

4.2.4.1.1 EAG critique

The EAG considers that using the treatment effects from RHODOS to inform the idebenone and SoC treatment effect for the first six months of the model is appropriate, as is the use of better seeing eye VA data to derive the transition probabilities.

4.2.4.2 Idebenone long term treatment effects

As described in Section 3.2.3, LEROS and the Expanded Access Programme (EAP) are single-arm open-label studies measuring long term idebenone treatment efficacy. While LEROS is a natural history-controlled intervention study (n=199) conducted over 24 months, the EAP study is a real-world evidence (RWE) non-controlled retrospective analysis (n=87) over 36 months. When deciding which study was most suitable to derive idebenone treated patient transition probabilities post six months (end of the RHODOS study), the EAP study was preferred by the company due to the lesser heterogeneity compared to the LEROS patient populations, with an additional advantage being the longer study time of the EAP.

Compared to the RHODOS study in which 85.9% of patients were male, LEROS contained [redacted]% males with the EAP study containing a more similar 82% male. Additionally, the genetic distribution of the EAP population was more aligned to that of RHODOS than LEROS, with the RHODOS trial only

consisting of idebenone-treated patients who carried three LHON mutations (m.11778G>A [67.3%], m.14484T>C [20%], m.3460G>A [12.7%]), compared to LEROS which consisted of patients from a wider range of LHON mutations (m.11778G>A [redacted], m.14484T>C [redacted], m.3460G>A [redacted], Negative [redacted], Other [redacted]) than the EAP study (m.11778G>A [62.1%], m.14484T>C [18.4%] m.3460G>A [19.5%]). Given that m.14484T>C patients are considered more likely to spontaneously recover by the company and their clinical experts and the difference in the proportion of male patients, the LEROS data were excluded from the economic model with the company preferring to derive transition probabilities from 6 to 36 months using the EAP for the idebenone treated patients.

As a means of validating the use of the EAP, the company compared the six-month outcomes of the RHODOS and EAP studies. The company concluded that the outcomes were broadly similar, with 30.2% (16/53) of idebenone treated patients in RHODOS and 46% (40/87) of patients in the EAP study achieving clinically relevant recovery (CRR) at six months.

After 36 months in the model, the company assumed that patient logMAR VA would stabilise and remain unchanged as this was the opinion provided to the company by their clinical experts. Therefore, patients are modelled to remain in their health state from cycle 12 (month 36) until death in both the idebenone and SoC treatment arms.

4.2.4.2.1 EAG critique

The company preferred to use the EAP over LEROS to derive a long-term treatment effect for idebenone due to the less heterogeneity in genetic distributions and sex proportions. Appendix N (clinical validation) however states that company's clinical experts concluded that the baseline characteristics of the RHODOS, RHODOS-OFU and LEROS trials were all representative of the patient population in clinical practice.

The company aimed to validate using the EAP by comparing CRR achieved at six months between trials; however, the LEROS study outcomes were more comparable to RHODOS than EAP, with [redacted] of LEROS achieving CRR compared to 46% of patients in EAP and 30% of patients in RHODOS. The EAG therefore considers that the RHODOS and EAP outcomes are highly varied with 50% more patients achieving CRR in EAP compared to RHODOS, and notes that LEROS provides more comparable clinical outcomes than the EAP.

In evaluation of CRR as a clinically relevant measure, the EAG notes that as CRR is defined as improvement of at least logMAR 0.2 (equal to two lines of readable letters on a logMAR chart) for

patients with “on-chart” VA at baseline, or an improvement from “off-chart” VA to at least logMAR 1.6 (equal to one line on-chart) for patients with off-chart VA at baseline, CRR may be achieved with no difference in functional sight or change in HRQoL (patients are still considered vision impaired or unable conduct key autonomous function such as driving). Therefore, CRR may not be a helpful indicator of improved HRQoL as it does not differentiate between sight recovery and functional sight recovery. For example, although 30% of idebenone patients achieved CRR in RHODOS after six months, mean recovery in terms of logMAR was 0.037, the equivalent of 1 letter on a logMAR chart.

The EAG is additionally concerned with the use of the EAP dataset over LEROS given the difference in eligible patients in each study. Only patients with symptom onset in the most recently affected for less than one year were included in the EAP study, while patients with symptom onset of less than five years were included in the LEROS study, which was the same inclusion criteria for RHODOS. Given that 44% of LEROS patients had experienced symptom onset for more than one year, if spontaneous recovery is more likely to occur earlier after nadir, then spontaneous recovery would be more likely to occur while being treated with idebenone in EAP than in LEROS, leading to a potential additional confounding of the estimated treatment effect in the EAP.

The LEROS study is also larger than the EAP (196 vs 87 patients) and therefore using LEROS may have lessened the key issue of missing data, which in combination with the high number of health states, leads to multiple transitions between health states in the model being impossible. For example, under both probabilistic and deterministic conditions, idebenone treated patients are unable to move to or remain in the Hand Movement health state past cycle 10 (two and a half years in the model). It’s similarly impossible for any idebenone treated patients to be logMAR 0.3 to 0.6 between cycles 8 and 9.

Even when data is not missing, due to the limited number of patient observations, the transition probabilities are highly uncertain and have far reaching consequences. For instance, the penultimate and final idebenone transition probabilities before VA is assumed to be fixed till death are calculated using only 9 patient observations across the eight health states with three additional imputed observations being required so that transition probabilities can be calculated for all health states.

For these reasons and those outlined in Section 3.2.3, the EAG considers that the LEROS study is more appropriate to inform the idebenone treatment effect after RHODOS, and as such, the company was requested at the clarification stage to conduct a scenario deriving idebenone

transition probabilities using the LEROS ITT population. The company complied with the EAGs request, with the ICER increasing from £18,578 to £21,129 in the scenario.

To validate the use of the LEROS treatment effect in the EAG's base case assumptions, the model mean change in logMAR was plotted against the RHODOS-OFU, LEROS and LEROS natural history matched analysis study findings.

Figure 9 shows that in the scenario using LEROS, logMAR change from baseline is equal to that of the EAP at 6 months as the RHODOS treatment effects are used up to this point. However, after 6 months, the LEROS change in logMAR is less volatile compared to the EAP, possibly due to more patient observations being available in LEROS. The EAP logMAR change from baseline is greater than LEROS and RHODOS-OFU at 36 months (from when logMAR is assumed to be fixed in the model), but the LEROS natural history matched analysis identified a comparable logMAR change after 2 years. The EAG notes that in the natural history matched analysis using LEROS and CaRS patients, no significant difference in logMAR was identified between idebenone and SoC treated patients. For these reasons the idebenone treatment effects from LEROS are assumed in the EAG's base case.

Figure 8. Idebenone mean logMAR change from baseline

4.2.4.3 SoC long term treatment effects

A similar approach was used by the company to select which data should be used to supplement the RHODOS SoC treatment effectiveness in the model. While the company submission (CS) does not explicitly draw on example comparisons between studies as was done between RHODOS, LEROS and EAP for the idebenone treated arm, the company outlined that baseline characteristics of the natural history patients in CaRS-I were similar in terms of age, sex and mutation type to RHODOS and so the study was suitable for deriving SoC transition probabilities in the model.

As data collected in the CaRS-I study had variable follow-up times the company used a "windowing approach" to classify CaRS-I patient observations into 3 months windows. For example, patients with a visit ≥ 1.5 months and < 4.5 months were assigned the 3-month window, while patients with a visit ≥ 4.5 months and < 7.5 months were assigned to the 6-month window. The company used a last observation carried forward (LOCF) approach to control for missing data and imputed the data to allow patients to remain in their health states when no data was available to inform the transition probabilities.

To validate the approach the company compared the proportion of patients who achieved clinically relevant recovery (CRR) at 6 months between RHODOS and CaRS-I, with 10.3% (3/29) of placebo-treated patients in RHODOS and 8.1% of natural history patients in CaRS-I (6/74) achieving CRR. Given the similarity of outcomes at six months, the company therefore concluded that despite some heterogeneity in terms of population and data analysis, CaRS-I was a suitable dataset to model SoC treatment effectiveness after RHODOS.

The company further justified the use of CaRS-I by outlining that the only alternative to using CaRS-I would be to assume no change in VA after six months, thereby using only the measured SoC treatment effects from RHODOS, due to the limited data available.

4.2.4.3.1 EAG critique

While the company outlined their preference for using CaRS-I to derive transition probabilities for the SoC arm, the EAG considers that using CaRS-II or combining CaRS-I and -II studies would have provided a more robust estimate of the SoC treatment effect. As described in Section 3.24, CaRS-II was a retrospective observational study conducted to establish the natural history of LHON patients, specifically aimed to gather data to serve as the natural history comparator group for the LEROS study.

Using the CaRS-II study may have reduced the uncertainty in the treatment effect introduced by the small number of patient observations in the CaRS-I study (n=87), as the CaRS-II population was much larger (n=219). As using the larger CaRS-II study (or a combination of CaRS-I and -II) could reduce the extent of missing data imputation required and the need for LOCF, the EAG requested a scenario that derived a SoC treatment effect using both CaRS -I and -II patients. The company conducted the scenario as requested, using the EAG's preferred model while removing SoC observations generated using LOCF which decreased the ICER to £6,463.

The EAG noted that in the company's scenario, all transition probabilities were informed using only 169 observations, compared to the 740 observations when using CaRS-I and LOCF. In contrast, the company reported in Table 1 of the supplementary clarification response that 944 appropriate and usable observations (from the 5,186 observations recorded in the studies) taken from the 385 appropriate patients from CaRS-I and -II are available. Additionally, individual transition probabilities appear to only be informed by observation from a maximum of 49 patients (the transition probability between months 6 and 9) and a minimum of nine patients (the transition probability from 18 to 21 months). In the scenario the EAG also notes that model mean logMAR is significantly higher than reported in the CaRS -II study. At five years, mean logMAR was approximately 1.64 in the model and 1.06 in the CaRS -II study.

The EAG therefore considers that the SoC treatment effect in the scenario is underestimated and highly uncertain given only a fraction of the patient observations are utilised and the model outcomes do not align to the clinical outcomes. The EAG additionally notes that a more robust treatment effect may be calculated and used in the model should the company have utilised all appropriate and available patient observations from the CaRS studies.

Given the large difference in patient observations when removing observation generated using LOCF, the company was requested to provide the number of SoC observations generated using LOCF used in their base case assumptions. In response the company provided the data in Table 24 .

Table 24. The number of patients whose observations were LOCF at each timepoint in the CaRS-I data (reproduced from Table 22 in the clarification response)

Timepoint	Number of patients whose observations were LOCF at each timepoint (%)
Baseline	0 (0%)
Month 3	21 (28.4%)
Month 6	35 (47.3%)

Month 9	48 (64.9%)
Month 12	59 (79.1%)
Month 15	61 (82.4%)
Month 18	63 (85.1%)
Month 21	66 (89.2%)
Month 24	69 (93.2%)
Month 27	71 (95.9%)
Month 30	70 (94.6%)
Month 33	71 (95.9%)
Month 36	71 (95.9%)

Abbreviations: CaRS, Case Record Survey; LOCF, Last observation carried forward

As shown, at six months where the SoC treatment effectiveness is informed by CaRS-I after RHODOS, almost half of the observations were generated using LOCF. By one year, almost 80% of observations were generated using LOCF, with this proportion increasing to approximately 90% by 21 months. Critically, the transition probabilities that dictate the health state a SoC patient will remain until death are calculated using observations only 4% of which were taken directly from patients at that time. When removing the observations not generated using LOCF, the number of observations used to derive transition probabilities throughout the model using the CaRS-I study falls from 740 to 88. The EAG therefore considers treatment effects associated with using the CaRS-I dataset are highly uncertain and inappropriate with and without LOCF.

Given the EAG consideration that a robust SoC treatment effect may be calculated from the available company data, the EAG requested that the following scenarios be conducted:

- A scenario using all appropriate CaRS-I and -II patient observations available;
- A scenario using the matched natural history patients from CaRS and idebenone LEROS patients, as used to conduct the matched control analysis;
- A scenario using only RHODOS-OFU to model long term treatment effects.

The company did not conduct the requested scenarios, stating that the previously conducted scenario used all available CaRS patient data. Additionally, the natural history matched controlled comparators could not be used due to the matching algorithm being performed *de novo* at each time point. The EAG, therefore, considers that the data are available to conduct a matched analysis but that the company's current matching algorithm is inappropriate for use. Lastly, the company did

not conduct a scenario using the RHODOS-OFU study as the company considered the scenario inappropriate.

As the requested scenarios were not supplied by the company, the EAG conducted an additional scenario. Given that the RHODOS-OFU study showed a maintained difference in change in logMAR from the end of RHODOS (6 months) to the end of RHODOS-OFU (30 months), between idebenone and SoC patients (Figure 15 in the CS), the scenario applied the idebenone transition probabilities from LEROS to SoC patients after RHODOS (see Section 6.3). The EAG consider this approach was preferred to anchoring the idebenone treatment effect to SoC given that the SoC treatment effects from CaRS are highly uncertain. As a means of validating the appropriateness of the scenario, the company base case and scenario mean change in logMAR from baseline results were plotted against the RHODOS-OFU and LEROS natural history matched (CaRS) analysis findings. The RHODOS mITT results were not plotted given the similarity in outcomes to the RHODOS-OFU study (+0.123 and +0.127, respectively).

As seen in Figure 10, SoC patient mean change in logMAR under the LEROS transition probabilities scenario aligns more closely to the RHODOS-OFU results compared to the company base case over time. The figure additionally highlights that although the RHODOS study treatment effects are stated to be applied to SoC patients in the model up to 6 months, change in logMAR from baseline between RHODOS mITT and the model are significantly different, with the mean change in logMAR of SoC patients in the model being 2.28 times worse than measured in the RHODOS trial at 6 months (+0.289 in model vs +0.123 in RHODOS mITT and +0.127 in RHODOS-OFU). If instead considering the RHODOS ITT population outcomes, which as outlined in Section 3.2.1 the EAG deems more appropriate, the model outcomes at 6 months are 3.44 times greater in the model than the trial (+0.289 in the model vs +0.084 in the RHODOS ITT population). The EAG, therefore, includes applying the LEROS treatment effects to SoC patients after the RHODOS treatment effects in its preferred assumptions but caveats that this assumption may be conservative. While at 36 months in the scenario, change from baseline logMAR is slightly better than measured in the RHODOS-OFU trial at 36 months (-0.0065 vs 0.039, respectfully). At the beginning of the model the SoC treatment effect in the LEROS scenario is greatly underestimated compared to the SoC treatment effect in the RHODOS study (6 months) the outcomes of which are less discounted in the model.

The EAG also notes in Figure 10 that the natural history matched analysis showed that SoC logMAR improved over time and was comparable to idebenone logMAR improvement (-0.24 and -0.28, respectively).

Figure 9. SoC mean logMAR change from baseline

The EAG considers it a key issue that company's base case assumptions informing the SoC treatment effect do not replicate the clinical trial findings in the model. Additionally, the EAG considers that a robust SoC treatment effect may be derived from the available CaRS patient data; however, limited patient observations are used from the studies compared to the potentially appropriate and available patient data. Similarly, as reported in Section 3.4.2, alternative matching methodologies could also have been employed to provide more robust treatment effects.

4.2.4.4 Sensitivity analyses

To assess the uncertainty of the idebenone treatment effects on the cost effectiveness results, the company varied the patient observations from the RHODOS and EAP studies, informing the transition probabilities between health states, using a Dirichlet distribution. The same method was not applied to SoC transition probabilities and so the cost effectiveness sensitivity to SoC treatment effectiveness uncertainty was not assessed in any probabilistic sensitivity analysis (PSA) conducted by the company.

4.2.4.4.1 EAG critique

As the model lacked the functionality to allow the SoC treatment effects to be made probabilistic, at the clarification stage the EAG requested that the company allowed the SoC treatment effects to be probabilistic using a similar approach to the idebenone treatment effects. The company however misinterpreted the request, instead removing the probabilistic functionality from the idebenone transition probabilities.

As a follow up clarification question, the EAG requested that both idebenone and SoC transition probabilities be made probabilistic; however, the company did not comply with the EAG's request, stating that, "*including the transition probabilities in the PSA creates substantial uncertainty in the probabilistic results of the CEA. Therefore, the transition probabilities have not been included in the PSA*". The company added that they strongly considered that including the transition probabilities in the PSA will create highly inaccurate probabilistic cost-effectiveness results that will be inappropriate for decision-making. As such, the PSA does not account for treatment effectiveness uncertainty for either idebenone or SoC treatment effects. This is a key issue, as the EAG considers the treatment effects to be highly uncertain given the limited patient data and that the NICE Guide to the Methods of Technology Appraisal states that PSA results are no longer simply recommended but are a mandatory requirement for all cost-effectiveness models submitted to NICE.

4.2.4.5 Treatment discontinuation

All idebenone patients in the model were assumed to experience treatment effects, with no patients experiencing no benefit of treatment. Patients who discontinued idebenone in the model were assumed to continue experiencing idebenone treatment effects and not SoC treatment effects.

4.2.4.5.1 EAG critique

While treatment discontinuation is accounted for in idebenone treatment cost calculations (approximately 40% discontinue treatment after two years), the company's model reflects that no idebenone patients who discontinue treatment go on to experience SoC treatment effects.

The EAP CSR states that of the 111 patients enrolled in the study, 12 patients discontinued treatment due to lack of efficacy, which the EAG considers should be incorporated into the model. When asked at clarification why discontinuation had been applied to treatment costs and not treatment effects, the company noted that although the EAP report v5.0, dated 11 October 2018,

stated that 12 patients out of the 111 patients enrolled did discontinue due to a lack of efficacy, the final EAP report dated 28 August 2019 stated that cumulatively, only nine out of 111 patients permanently discontinued idebenone treatment due to the lack of efficacy, or occurrence of AEs, or a fatal outcome which is captured in the EAP safety population.

The EAG considers that given the transition probabilities are derived from the EAP mITT (n=87) population in the company base case and not the EAP safety population (n=111), patients who experience no treatment benefit with idebenone may not be captured in the model. Similarly, patients who discontinue treatment may not have attended later appointments in the EAP or LEROS studies and so their lack of clinical benefit would not be included as observations in the model. At clarification the EAG requested a scenario exploring treatment discontinuation which the company provided. In the scenario, the company assumed that 4% of patients discontinued idebenone after two years. This proportion of patients was assumed to be in addition to the proportion of patients who discontinue treatment accounted for in the treatment costs and was calculated using a weighted average. The scenario led to an increase in the ICER, from £18,758 to £19,709.

The EAG notes that as the 4% patient treatment discontinuation was made in addition to the patients already discontinuing treatment within treatment costs, treatment costs in the scenario are likely underestimated as patient treatment discontinuation is potentially double counted. Similarly given that 12 of 111 EAP patients discontinued treatment due to a lack of efficacy, the EAG considers the proportion of patients who discontinue treatment in the scenario should be 10.8%. Lastly, patients in the company scenario who discontinue treatment after two years still experience idebenone treatment effects for two years before discontinuing treatment when no treatment effect should be accounted for due to the lack of treatment efficacy.

As such, using the same weighted average approach as the company, the EAG conducted a scenario in which 10.8% of idebenone treated patients incurred idebenone treatment costs and SoC treatment effects for two years before discontinuing treatment (see Section 6.2). The scenario led to an increase in the ICER to £21,022.

4.2.5 Mortality

Mortality assumed for LHON patients was that of all-cause mortality stratified by age and sex using England general population estimates from 2018 to 2020. The company noted that evidence exists demonstrating that the risk of mortality is higher in patients who are visual impaired and therefore

idebenone could be considered to reduce mortality risk compared to no treatment. However, given the lack of specific mortality data for idebenone, the conservative assumption was made to not include a survival benefit for idebenone treated patients.

4.2.5.1 EAG critique

The EAG considers that given the lack of evidence provided for an idebenone survival benefit the company's approach of assuming no survival benefit is reasonable.

4.2.6 Health-related quality of life

4.2.6.1 Health state utility values

The key clinical trials used for measuring the effectiveness of idebenone, discussed in Section 3.2.1, collected condition specific health-related quality of life (HRQoL) data only using the Visual Function Index (VF-14), Clinicians Global Impression of Change (CGIC) and energy levels. The NICE Reference Case[52] recommends the use of EQ-5D-3L directly measured from patients for the estimation of HRQoL. When not available from clinical trial data, EQ-5D data can be sourced from published literature or estimated by mapping from other measures of HRQoL collected in the clinical trials, using published mapping algorithms. No published mapping algorithm is available to map from VF-14, collected in the RHODOS clinical trial, to the EQ-5D. Therefore, the company undertook a systematic literature review (SLR) to identify appropriate health state utility values for use in the economic model.

The company's SLR identified no studies providing utility values for LHON patients that could be used in the economic model. Therefore, the company undertook a targeted literature review to identify utility values based on related diseases and those used in previous NICE TAs (HST 11,[46] TA298,[47] TA283[48] and TA294[53]). Based on the targeted search, the company's base case utility values were based on Brown *et al.* 1999.[23] Brown *et al.* derived utility values using time trade off (TTO) valuation from 325 patients with vision loss due to a range of vitreoretinal diseases, with the majority of patients having either age-related macular degeneration (ARMD) (33%) or diabetic retinopathy (33%). Utility values were provided separately for both the best seeing eye (BSE) and worst seeing eye (WSE), with the company using values for the BSE in their model. Visual acuity was reported across 12 states, represented as a fraction (out of 20 feet), which the company converted to the corresponding logMAR, to match the measurement used in the economic model. As the health states used in the economic model were based on logMAR range as opposed to the point estimate

as presented in Brown *et al.* 1999, the midpoints for each logMAR range used in the economic model were matched up with the closest utility value from Brown *et al.* The utility derived from Brown *et al.* and applied in the company's base-case economic model for each health state are shown in Table 25.

Table 25. LogMAR utility values derived from Brown *et al.* and corresponding model health state utility values (reproduced from Table 26 of the CS)

Brown *et al.* visual acuity	Brown *et al.* utility (95% CI)	Mid-point health state	Model utility value
LogMAR = 0	0.92 (0.87 to 0.97)	LogMAR <0.3	0.84
LogMAR = 0.1	0.87 (0.82 to 0.92)		
LogMAR = 0.2	0.84 (0.79 to 0.89)		
LogMAR = 0.3	0.80 (0.74 to 0.86)	LogMAR 0.3 to 0.6	0.77
LogMAR = 0.4	0.77 (0.70 to 0.84)		
LogMAR = 0.6	0.74 (0.67 to 0.81)	LogMAR 0.6 to 1.0	0.67
LogMAR = 0.7	0.67 (0.57 to 0.77)		
LogMAR = 1.0	0.66 (0.55 to 0.77)	LogMAR 1.0 to 1.3	0.63
LogMAR = 1.2	0.63 (0.54 to 0.72)		
LogMAR = 1.3	0.54 (0.43 to 0.65)	LogMAR 1.3 to 1.7	0.54
CF	0.52 (0.36 to 0.68)	CF	0.52
HM-NLP	0.35 (0.10 to 0.60)	HM/LP	0.35

Abbreviations: CI, Confidence interval; CF, Counting fingers; HM, Hand motion; LP, Light perception; NLP, No light perception

The company also provided alternative utility values identified via the targeted search from Lawrence *et al.* 2023b,[2] Czoski-Murray *et al.* 2009[54] and Rentz *et al.* 2014[55] and provided scenario analyses with these values applied. Following the clarification stage, the company also provided the same sources with the utility values adjusted to match the EAG's proposed model structure using a reduced number of health states, as discussed in Section 4.2.2.

Utility values were not originally adjusted to account for reductions in quality of life with age in the company's economic model. Following a clarification request, the model and company's base-case were updated to include age-adjusted utility values using the Health Survey for England (HSE) 2014 dataset, as recommended by the NICE Decision Support Unit (DSU).[56]

4.2.6.2 EAG critique

In light of a lack of EQ-5D values from the RHODOS trial and no mapping algorithm available from VF-14 collected during the clinical trial, the EAG considers the use of utility values from the literature

to be generally appropriate. However, the EAG notes that details of the targeted search undertaken by the company were not provided and therefore it is unknown if alternative values that may have also been appropriate or relevant are available.

The EAG notes that none of the utility values sourced by the company from the literature use directly reported EQ-5D-3L as preferred in the NICE Reference Case. The utility values used from Brown *et al.*[23] in the company's base case is the only study included by the company that derives utilities from patients, however this was not specific to LHON patients. The majority of patients (66%) in the study had either ARMD or diabetic retinopathy and an average age of 67 years. The EAG notes that the average age of patients from Brown *et al.*[23] is significantly higher than the average age of patients with LHON and the start age of the model (34 years), and patients with ARMD or diabetic retinopathy may have related co-morbidities that could result in utility values not being reflective of LHON patients. In addition, utility values from Brown *et al*. were based on patients from the United States of America (USA).

The company also considered utility values from three alternative studies; Lawrence *et al.* 2023b,[2] Csozki-Murray *et al.* 2009[54] and Rentz *et al.* 2014.[55] Both Csozki-Murray *et al.* 2009[54] and Rentz *et al.* 2014[55] were identified due to being considered in previous NICE TA's (HST 11,[46] TA298,[47] TA283[48] and TA294[53]). The EAG notes Rentz *et al.* provides utilities for eight descriptive health states which were developed based on the Visual Function Questionnaire-Utility Index. For these utility values to be used it would require making assumptions regarding how the health states match up to the equivalent logMAR score. Csozki-Murray *et al.*[54] provided utility values for four health states based on logMAR scores of ≤0.30, 0.31– 0.60, 0.61 – 1.30 and ≥1.31. Based on the grouping of the health states used in the current economic model, the logMAR categories from Csozki-Murray *et al.*[54] may not align well with the health states used, requiring assumptions to be made. The EAG notes that one of the UK clinical experts in the company's validation survey (Appendix N) commented that ([redacted]).

The EAG notes that Lawrence *et al*. 2023b[2] is the only study to provide utility values specifically based on LHON. Although currently only published in poster/abstract form, the EAG considers there is sufficient detail available to review. Lawrence *et al*. 2023a[57] describes the development of eight health state vignettes which varied by level of visual acuity, defined by logMAR score. Draft vignettes

were developed based on literature reviews and trial data and then subsequently revised following feedback from nine LHON patients and five clinicians from the United Kingdom (UK) and Republic of Ireland (ROI). The eight health state vignettes were then valued by 362 members of the public from both UK and ROI using the Health Utilities Index-3 (HUI-3) and EQ-5D-5L via and online survey and a sub-sample of 120 participants also completed TTO interviews. Although referred to in the abstract as EQ-5D-5L health state utility values, it is noted that the EQ-5D-5L data were scored using the Hernandez *et al.*[58] mapping function, which maps to EQ-5D-3L. The EAG confirmed with authors of the study that the EQ-5D-5L data had been mapped to EQ-5D-3L utility values. From the available data, the vignette descriptions seem well defined and used a variety of evidence sources and information to develop them, in line with the NICE DSU recommendations.[59] The NICE DSU report suggests that vignettes should not include value-laden or irrelevant phrases or content, such as "devasting". The EAG does note, however, that the worst health state vignette (logMAR≥4) describes emotional impact of the disease as "*vision loss is devastating, and you find it very difficult to come to terms with*". Although this deviates from the recommendations from the NICE DSU report,[59] the EAG notes that the term 'devasting' is only used in one aspect of the vignette (emotional impact).

The average age of patients completing the valuation survey was 46.5 years old, which is substantially lower than that of the patients used in Brown *et al.* and closer to the average age of patients experiencing LHON. Due to this, and the fact that these values are estimated specifically for LHON and valued by the UK population (and ROI), the EAG considers that this is the most appropriate source of utility values for the economic model. The EAG notes that the choice of valuation method used resulted in wide variation in the utility values estimated for each logMAR health state, as shown below in Table 26, with HUI-3 valuation consistently giving lower utility values. Table 26 also shows the utility values from Lawrence *et al*. 2023b grouped into the EAG's preferred health states, as discussed in Section 4.2.2. from taking the average of the values when grouped. The EAG notes that in Lawrence *et al.* 2023b it is noted that during vignette development interviews, clinical experts discussed the potential overlap between health states and that they were unable to differentiate HRQoL impacts with similar health states. The EAG considers this to be a further validation of using a reduced number of health states.

Table 26. Estimated utility values by logMAR visual acuity, produced based on Figure 2, Lawrence *et al.* 2023b

	HUI-3 (n= 362)	EQ-5D-5L (mapped to EQ-5D-3L)	TTO (n=120)	HUI-3 (n= 362)	EQ-5D-5L (mapped	TTO (n=120)

		(n=358)			to EQ-5D-3L) (n=358)	
LogMAR <0.3	0.837	0.790	0.882	0.837	0.790	0.882
LogMAR ≥0.3 to <0.6	0.511	0.632	0.756	0.473	0.603	0.729
LogMAR ≥0.6 to <1.0	0.435	0.574	0.702			
LogMAR ≥1.0 to <1.3	0.347	0.495	0.565	0.336	0.496	0.545
LogMAR ≥1.3 to <1.7	0.325	0.497	0.525			
LogMAR ~2	0.211	0.368	0.406	0.194	0.352	0.398
LogMAR ~2.3	0.190	0.347	0.426			
LogMAR ~4	0.180	0.341	0.363			

Abbreviations: EQ-5D-5L, EuroQol-5 dimensions-5 level; EQ-5D-3L, EuroQol-5 dimensions-53; level HUI-3, Health Utilities Index-3; TTO, time trade off

Following confirmation from the study authors for Lawrence et al. 2023b[2] that utility values are mapped to EQ-5D-3L, the EAG considers these to be the most appropriate for decision making. Therefore, the EAG's preferred values for their base case analysis is using the EQ-5D values from Lawrence et al. 2023b,[2] with the values using alternative valuation methods applied in scenario analyses (see Section 6.2). The EAG notes that the values applied in the company scenario analysis for Lawrence et al. 2023b, provided as part of the clarification response, differ slightly to those used in the EAG preferred analysis, shown in Table 26, as the EAG has used the values reported for both the UK and ROI due to the larger sample size, whereas the company only reported the values for the UK.

4.2.6.3 Utility decrements

The company applied a lifetime utility decrement of 0.04 for all patients with a logMAR >1.0 to represent the disutility associated with LHON caregivers HRQoL. This disutility is applied for a patient's lifetime. As no quantitative caregiver HRQoL had been collected, the company utilised data from a published systematic review exploring the disutility of caring for an ill or disabled family member, previously used in HST11.[46] The value used by the company is based on a study identified in the systematic review stating that parents of children with activity limitations have a 0.08 lower EQ-5D score than parents of children without activity limitations. The company applied the same approach employed in HST11 and assumed that the disutility of carers of adults with activity limitations would be half of that applied to children.

No adverse events (AEs) disutility was applied in the economic model as, based on the RHODOS trial, most AEs experienced were considered mild.

4.2.6.4 EAG critique

During the clarification stage, the EAG requested that the Company remove the disutility of a caregiver for the proportion of patients who would be in residential care, as these patients would already be receiving separate care service. As part of the clarification response, the Company updated the base-case to remove the carer disutility for the proportion of patients receiving residential care. This had a small impact in the ICER.

The EAG notes that although the same disutility values were applied in HST11 to reflect the impact on caregivers for caring for a family member experiencing blindness, the committee concluded that these values should only be applied to carers of children and not adults and therefore the exclusion of a carer disutility for adult patients was used for decision making. Although the EAG recognises that patients experiencing blindness will require additional assistance from a caregiver, based on the available evidence, the disutility impact on caregivers is uncertain. The EAG preferred analysis applies no caregiver disutility in the base-case, with a scenario analysis provided to explore the impact of its inclusion (see Section 6.2).

4.2.7 Resource use and costs

4.2.7.1 Drug acquisition costs

The list price for idebenone 150mg is £6,364 per pack of 180 tablets. A confidential patient access scheme (PAS) is in place for idebenone (a simple discount of [redacted]) and all results presented in this report include the corresponding PAS. Dosing and subsequent drug acquisition costs in the model follows the SmPC recommended dose of 300mg three times per day. The three month cycle drug acquisition costs are shown in Table 27. The company assumed no administration costs associated with idebenone due to being an oral treatment.

Table 27. Idebenone drug acquisition costs

Dose per day (mg)	Dose per 3 month cycle (mg)	Packs required per cycle	Cost per 3 month cycle (list price)	Cost per 3 month cycle (PAS price)
900	82,181	3.04	£19,370	[redacted]

Abbreviations: mg, milligrams; PAS, patient access scheme

Based on compliance rates observed in the RHODOS trial, the company applies a reduction in drug acquisition costs based on a 96% compliance rate, assumed to apply for the duration of time patients remain on treatment. Patients are assumed to remain on treatment up to three years only, with a proportion of patients discontinuing each cycle based on pooled Kaplan-Meier (KM) treatment duration data observed in RHODOS and EAP, shown in Figure 11. As noted in Section 4.2.4 patients discontinuing treatment affects costs only.

Figure 10. Kaplan Meier curve for time of treatment with idebenone based on RHODOS/EAP data, reproduced from Figure 21 of the CS

As no therapeutic treatments are currently available for LHON, SoC is assumed to consist of established clinical management which includes visual aids, low vision rehabilitation and lifestyle management. No separate treatment costs are assumed to apply for SoC as this is instead captured in the health state resource use costs and also applied to patients on idebenone.

4.2.7.1.1 EAG critique

Clinical experts outlined to the EAG that they may continue to treat patients up to three years and beyond if patients were responding to treatment or had only recently stabilised. In addition, within the EAP study, treatment duration ranged from 2.4 – 70.4 months. During the clarification stage, the

EAG requested that the company extrapolated the time on treatment KM data using parametric curves to inform treatment costs for idebenone patients beyond three years. The company stated that due to the low patient numbers available beyond three years, extrapolating these data would be highly uncertain and inappropriate. In addition, it was noted that as clinical data is only measured up to three years, extrapolating time on treatment data beyond this time point would be biased against idebenone as costs are accrued without clinical benefit.

Although the EAG agrees that there are limited patient numbers available to produce meaningful extrapolations, it is noted that in years 2–3 of the economic model, idebenone patients are still moving between health states and improving, suggesting that patients may not yet have stabilised. As clinical experts suggested to the EAG that they may continue to treat patients who recently stabilise, and two of the company's UK clinical experts stated they would treat until stabilisation or plateau, the EAG considers that in clinical practice treatment costs will continue to be accrued for recently stabilised patients beyond three years. Therefore, the EAG considers it plausible for treatment costs to continue beyond the three year time period despite patients no longer moving between health states in the economic model. Not applying any treatment costs beyond this period may underestimate the true costs and is an area of uncertainty in the economic model. However, the EAG does not consider there to be robust data available to estimate the duration of treatment costs beyond three years.

As noted in Section 4.2.4, the EAG considers that LEROS is the more appropriate study to inform the idebenone treatment effectiveness after RHODOS, compared to EAP used by the company. However, as LEROS was conducted over a shorter time period than EAP (up to two years only), treatment discontinuation is curtailed by the shorter study duration. Therefore, the EAG considers the use of the LEROS data to inform treatment effectiveness, with EAP data used to inform the time on treatment to be more appropriate. Despite the EAG considering that patients may continue to receive treatment past three years, the use of the longer-term EAP time on treatment data (up to three years) is assumed to provide an illustration of using a longer treatment duration when combined with the LEROS data used for treatment effectiveness.

4.2.7.2 Routine monitoring

In line with their clinical expert opinion, the company applied the cost of an ophthalmologist visit three times per year for patients treated with idebenone for the first year of treatment, followed by one visit per year for subsequent years. This was based on their clinical expert opinion stating that

they would expect to see patients on idebenone every 4–6 months in the first year. For patients on SoC, this was assumed to be once per year for the entire duration of the model. Annual resource use was converted to every three months to match the cycle length of the model.

4.2.7.2.1 EAG critique

The EAG's clinical experts stated that patients with LHON would have optical coherence tomography (OCT) undertaken each time they had an outpatient visit with ophthalmology. Similarly, the EAG noted the cost difference between first and follow-up attendance for ophthalmology visits in the NHS Reference costs. Therefore, at the EAG's request during the clarification stage, the company updated the costs of ophthalmology to use a separate cost associated with first visit and follow-up visits, while including the cost of an OCT alongside every ophthalmology visit and updated their base-case results to reflect this, resulting in a small increase in the ICER of £96. The final resource use and costs associated with routine monitoring in the model are shown in Table 28.

Table 28. Routine monitoring costs and resource use

Resource	*Unit cost*	*Per cycle resource use: idebenone (cycle 1-4)*	Per cycle resource use: idebenone (cycle 5+) and SoC	*Source*
Ophthalmology visit (first visit)	£166.64	0.75	0.25	NHS reference costs 2021/2022. outpatient care - Ophthalmology service, Non-Admitted Face-to-Face Attendance, First
Ophthalmology visit (subsequent visit)	£143.93			NHS reference costs 2021/2022. outpatient care - Ophthalmology service, Non-Admitted Face-to-Face Attendance, Follow-up (WF01A)
OCT	£158.23			NHS reference costs 2021/2022. Retinal Tomography, 19 years and over' (BZ88A)

Abbreviations: NHS, National Health Service; OCT, optical coherence tomography; SoC, standard of care

4.2.7.3 *Health state resource use*

The company included costs for each health state, assumed to represent the costs associated with blindness, varying by logMAR score. The included resources associated with blindness were informed from a published study by Meads et al. 2003,[3] used in previous NICE appraisals for eye conditions (HST11,[46] TA155,[60] TA294[53] and TA274[49]). The company's included resource use consisted of hospitalisations (assumed to be due to injurious falls), outpatient visits (obtaining low vision aids and rehabilitation), blind registration, supportive living, residential care (aged 65+ only) and depression. Both blind registration and depression were assumed to be one-off costs applied in the first year, whereas all other costs are assumed to occur per cycle.

The company stated that as Meads et al.[3] was based not specifically on patients with LHON and rather in an older population of patients strictly classed as blind, as such the reported resource use was not applicable to the LHON population. They therefore obtained estimates of each resource use across the included model health states, classified by logMAR value, from a survey [redacted]. The average of the estimated resource use from the [redacted] was then calculated and these estimates were validated by five UK clinical experts.

The unit costs for hospitalisations due to injurious falls was assumed to be the cost of an A&E visit, sourced from NHS Reference Costs 2021/22. All other included costs were taken from Meads et al. and inflated from 2001 prices to 2022 using the Personal Social Services Research Unit (PSSRU) hospital and community health service (HCHS) pay and price indices. Following a clarification request from the EAG, the company updated the cost used for residential care to be sourced from PSSRU 2022 rather than inflated from Meads et al.[3] Resource use and unit costs applied in the company's model are shown in Table 29 and Table 30, respectively.

Table 29. Resource use for each health state defined by logMAR used in the company's model

Resource	Health state resource use							
	LogMAR <0.3	LogMAR 0.3 to 0.6	LogMAR 0.6 to 1.0	LogMAR 1.0 to 1.3	LogMAR 1.3 to 1.7	CF	HM	LP
Hospitalisations	2%	3%	10%	18%	20%	22%	27%	30%
Outpatient visits	13%	38%	80%	83%	83%	83%	83%	83%

Blind registration*	0%	25%	78%	100%	100%	100%	100%	100%
Supportive living	0%	0%	20%	40%	48%	57%	63%	70%
Residential care (age 65+)	0%	2%	7%	7%	8%	20%	22%	35%
Depression due to LHON onset	7%	20%	30%	33%	42%	45%	58%	65%

Abbreviations: CF, counting fingers; HM, hand motions; LHON, Leber hereditary optic neuropathy; LP, light perception

* Applied in the first year only (cycles 1 to 4)

Table 30. Unit costs for health state resource use applied in the company's model

Resource	Unit cost from Meads (2000 prices)	Annual (inflated to 2022)	Per cycle	Source
Hospitalisations	-	£1,728.82	£432.20	NHS Reference Costs 2021/2022: Emergency Medicine, Category 3 Investigation with Category 1-3 Treatment (VB02Z) Cost representing hospitalisation due to falls
Outpatient visits for low vision	£341.63	£577.26	£144.31	Cost from Meads et al. inflated using PSSRU Cost representing low vision aids (£136.33) and low vision rehabilitation (£205.30)
Blind registration	£97.41	£164.74	£41.18*	Cost from Meads et al. inflated using PSSRU. Cost representing doctor's sessional fee for completing Certificate of Vision Impairment (59.70) and mean cost of a community occupational therapist for the initial assessment (£37.71)
Supportive living	£2,848.63	£4,818.23	£1,204.56	Cost from Meads et al. inflated using PSSRU Cost representing community home care worker
Residential care (age 65+)	-	£75,241.50	£18,810.38	PSSRU 2022. Local authority own-provision residential care for older people (age 65+)
Depression due to LHON onset	£391.97	£662.90	£165.72*	Cost from Meads et al. inflated using PSSRU. Cost representing costs depression due to the onset of LHON

Abbreviations: LHON, Leber hereditary optic neuropathy; NHS, national health service; PSSRU, Personal Social Services Research Unit

* Applied in the first year only (cycles 1 to 4)

In response to a clarification question (question B1), the company also provided the proportion of patients requiring each resource when using the EAG's preferred modelled health states, as discussed in Section 4.2.2. The estimates for the updated health states were calculated by taking the average of the proportions from the combined health states, reported below in Table 31.

Table 31. Company resource use estimates applied to the EAG preferred model structure

Resource	Limited visual acuities (logMAR <0.3)	Moderate visual acuities (logMAR 0.3 to 1.0)	On chart visual acutities logMAR 1.0 to 1.7	Off-chart visual acuities (logMAR >1.7)
Hospitalisations	2%	7%	19%	26%
Outpatient visits for low vision	13%	59%	83%	83%
Blind registration	0%	52%	100%	100%
Supportive living	0%	10%	44%	63%
Residential care (age 65+)	0%	4%	8%	26%
Depression due to LHON onset	7%	25%	38%	56%

Abbreviations: LHON, Leber hereditary optic neuropathy

4.2.7.3.1 EAG critique

The EAG notes that it is not aware of any published literature available on resource use for patients with LHON specifically, hence why the company used estimates derived from clinical experts. The resource categories included for health state costs were informed by Meads *et al*. [3] which, as previously stated, has been used in numerous NICE TAs. The EAG notes that in previous NICE TAs (HST11,[46] TA155[60] and TA294[53]), the proportion of patients expected to require each resource is taken directly from Meads *et al.* and applied to patients who are classified as blind, dependent on the visual acuity measure used in the economic models.

During the clarification stage, the EAG requested that the company included a scenario analysis using the proportion of patients requiring each resource taken directly from Meads *et al.* and applied only to patients with a logMAR >1. In response to clarification, the company included a scenario analysis using data from Mead *et al.* However, within this scenario, the company deemed it inappropriate to assume patients with logMAR<1 do not require any resource use and therefore

applied the same proportions obtained from their clinical experts for these health states. In addition, the company stated that as the proportion of patients requiring hospitalisation in Mead *et al.* reflects patients requiring a hip replacement rather than due to injurious falls in their own model, the proportions for hospitalisations estimated from their clinical experts is still applied in this scenario. The company noted that the estimated proportions reported in Meads *et al.* represents an elderly population which does not align with the LHON population in this current appraisal. The EAG notes that many of the estimated proportions used by the company are actually higher than those estimated by Meads *et al.* Clinical experts advising the EAG suggested that as LHON typically occurs in a younger population, patients often adapt to their eyesight, more so than if it had developed later in life. Therefore, it might be more likely that the resource use proportions in Meads *et al.* representing an older population are higher than that for the LHON population.

While the EAG agrees that the proportions estimated in Meads *et al.* may not be fully reflective of the younger LHON population, the EAG considers the proportions used in the company's base-case to be highly uncertain. The EAG notes that in the initial survey of [redacted], from which the resource use estimates were obtained, there was often a wide range between the highest and lowest estimates provided for many resource categories. The company then presented the averages of the three experts to five UK clinical experts for validation. The company reported that one clinical expert stated that they would expect outpatient care to be higher and therefore ran a scenario analysis in which this was 2 times higher. However,

[redacted]

The EAG clinical experts also provided comments on the resource use proportions included in the model. One expert stated that they would not expect young people with vision equal to driving vision to be experiencing regular falls, as estimated by the company's resource use. In response to a clarification question (B16b), the company referenced a study undertaken by the Royal Institute of Blind People (RNIB), reporting that 8,021 falls related to partial sightedness and blindness occurred in patients aged 18-59. The company also stated that the report estimated that half of fallers have reoccurring falls, thus supporting the application of hospitalisations as a regular per cycle cost in the economic model. While the EAG considers it plausible that a proportion of patients may have regular falls requiring hospitalisation, it is uncertain if this would apply to patients with good visual acuity,

every 3 months for the entire model duration, despite only being applied to a small proportion. The EAG, therefore, considers it more plausible to apply the proportion of patients requiring hospitalisation from Meads *et al*. to all patients with a logMAR >1 and assume that this proportion is representative of patients requiring hospitalisation due to injurious falls.

The cost of supportive living from Meads *et al.* 2003 used in the company's model is assumed to reflect the cost of a community home care worker. A clinical expert advising the EAG noted that they expect this would entail assessing the home environment and installing features that may help. It was noted that this would generally be a one-off visit rather than a regular on-going cost.

Clinical experts also stated to the EAG that supplying low vision aids, in the form of magnification tools and rehabilitation would not be an ongoing regular cost throughout a patient's lifetime but more of a one-off cost required on sight deterioration (likely when considered sight impaired [logMAR>1]). Following a request during the clarification stage, the company provided a scenario in which outpatient care costs were applied as a one-off costs rather than per cycle. The EAG deem this to be more reflective of clinical practice and applies this in the EAG preferred analysis (see Section 6).

Due to the uncertainty in the company's estimates derived from their clinical experts and a lack of available evidence for resource use in the LHON population, the EAG considers it more appropriate to use estimated resource use from Meads *et al.*, applied to patients with a logMAR>1. However, as the cost of depression is applied as a one-off cost, the EAG considers it more appropriate to apply the proportion of patients experiencing this cost to all health states as clinical experts advised that this is likely to affect all patients with a diagnosis of LHON as they adjust to their prognosis. Meads *et al.* reported separate resource use for low vision aids (11%) and low vision rehabilitation (33%); however, the company's model structure combined these into one resource use (Outpatient visits for low vision). Therefore, in order to implement the Meads *et al*. proportions in the company's model it was necessary to take the average of low vision aids and low vision rehabilitation (22%). As both the low vision aids and rehabilitation had similar costs (see Table 30), the EAG does not consider this will have a considerable impact. The EAG's preferred assumptions are summarised below and Table 32 shows the EAG's preferred resource use estimates in line with the EAG's preferred model structure:

- Proportion of patients requiring each resource sourced from Meads *et al.* 2003, applied only to patients with logMAR>1, except depression costs which are assumed to apply to all health states.
- Costs for outpatient visits for low vision (vision aids and rehabilitation), blind registration, supportive living and depression all applied as a one-off cost in the first year.
- Proportion experiencing hospitalisation assumed to be applicable to those having injurious falls.

Table 32. EAG preferred resource use assumptions applied to EAG preferred model health states

Resource	Limited visual acuities (logMAR <0.3)	Moderate visual acuities (logMAR 0.3 to 1.0)	On chart visual acuities logMAR 1.0 to 1.7	Off-chart visual acuities (logMAR >1.7)
Hospitalisations	0%	0%	5%	5%
Outpatient visits for low vision	0%	0%	22%	22%
Blind registration	0%	0%	95%	95%
Supportive living	0%	0%	6%	6%
Residential care (age 65+)	0%	0%	30%	30%
Depression due to LHON onset	39%	39%	39%	39%

Abbreviations: LHON, Leber hereditary optic neuropathy

5 Cost effectiveness results

5.1 Company's cost effectiveness results

Table 33 presents the cost-effectiveness results of the company's updated (i.e., post clarification) base case deterministic and probabilistic analyses. The probabilistic sensitivity analysis (PSA) conducted to assess the joint parameter uncertainty around base case results used a Monte Carlo simulation and derived probabilistic results from 1,000 generated simulations. The EAG notes that as described in Section 4.2.4, transition probabilities were not made probabilistic in the PSA and so treatment effectiveness uncertainty has not been accounted for. Therefore, while probabilistic results have been provided, the EAG considers these results may be inappropriate for decision making given the extent of the treatment effectiveness uncertainty.

Table 33. Company's base case results

Interventions	Total Costs (£)	Total LY	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
Deterministic results							
SoC	[redacted]	[redacted]	[redacted]	-	-	-	-
Idebenone	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	18,758
Probabilistic results*							
SoC	[redacted]	-	[redacted]	-	-	-	-
Idebenone	[redacted]	-	[redacted]	[redacted]	-	[redacted]	19,272

Abbreviations: ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality-adjusted life-year; SoC, standard of care.

*Probabilistic results do not account for idebenone and SoC treatment effectiveness uncertainty

A PSA scatterplot is presented in Figure 12 and a cost-effectiveness acceptability curve (CEAC) is presented in Figure 13. Based on these analyses, the probability that idebenone is cost effective versus SoC is 50% at a willingness to pay (WTP) threshold of £20,000 and 85% at £30,000, using the company base case assumptions.

The EAG notes that as the idebenone and SoC treatment effects have not been varied according to uncertainty in the estimated treatment effectiveness, the deterministic results may be considerably different from the probabilistic results.

Figure 11. Company base case PSA case scatter plot

Figure 12. Company base case cost-effectiveness acceptability curve

5.2 Company's sensitivity analyses

5.2.1 One-way sensitivity analysis

The company conducted a one-way sensitivity analyses (OWSA) to assess the sensitivity of the ICER to varying specific parameters in isolation and to identify the main model drivers. The results are illustrated in Figure 14. The EAG notes that while the ICER was most sensitive to the utility of patients with a logMAR of less than 0.3, the analysis did not vary treatment effectiveness which the EAG considers the ICER may be most sensitive to given the results of the EAG scenario conducted.

Figure 13. Company base case one-way sensitivity analysis

5.2.2 Scenario analysis

The company undertook a series of scenario analyses to assess the impact of applying alternative assumptions to key model parameters, in addition to several scenario analyses requested by the EAG. Results of all scenario analyses conducted by the company are presented in Table 34.

Table 34. Company conducted scenario analyses

Parameter	Scenario number	Base-case	Scenario	Incremental costs (£)	Incremental QALYs	ICER (£)
Discount rate for costs and outcomes	1	3.5%	0%	[redacted]	[redacted]	2,929
	2		1.5%	[redacted]	[redacted]	9,964
	3		6%	[redacted]	[redacted]	34,074
Time horizon	4	66 years	50 years	[redacted]	[redacted]	21,375
	5		30 years	[redacted]	[redacted]	29,754
Utility source	6	Brown *et al.* (1999)	Rentz *et al.* (2014)	[redacted]	[redacted]	18,787
	7		Lawrence *et al.* – EQ-5D-5L	[redacted]	[redacted]	22,070
	8		Lawrence *et al.* – HUI3	[redacted]	[redacted]	15,680
	9		Lawrence *et al.* – TTO	[redacted]	[redacted]	18,714
	10		Czoski-Murray *et al.*	[redacted]	[redacted]	20,094
Baseline characteristics source	11	RHODOS	EAP	[redacted]	[redacted]	20,333
	12		CaRS	[redacted]	[redacted]	19,224
	13		Pooled RHODOS, EAP and CaRS	[redacted]	[redacted]	19,484
Caregiver disutility	14	Included	Excluded	[redacted]	[redacted]	22,181
Compliance	15	RCT compliance	Full compliance – 100%	[redacted]	[redacted]	21,453
	16		RWE compliance – 87%	[redacted]	[redacted]	17,454
Resource use inputs	17	Informed by KOL survey (2022) with the exception of ophthalmologist visits	Base-case + outpatient care use adjusted according the UK clinical input	[redacted]	[redacted]	21,615

Using the EAG's proposed health states	B1	Company preferred model	EAG preferred model	[redacted]	[redacted]	27,053
Applying the LEROS data from month 6 to month 24 in the idebenone arm	B2	EAP	LEROS	[redacted]	[redacted]	21,129
Removing the LOCF assumption from the CaRS data	B3	Using LOCF	No LOCF	[redacted]	[redacted]	1,963
Applying a 4% idebenone treatment discontinuation rate for idebenone	B8	No discontinuation rate applied to treatment effect	Discontinuation rate applied to treatment effect	[redacted]	[redacted]	19,709
Applying various HRQoL sources using the EAG's proposed health states	B10	Company preferred model	EAG preferred model	[redacted]	[redacted]	19,107 – 29,407
Applying a one-off outpatient care cost for idebenone	B16	Recurring outpatient care cost	One-off outpatient care cost	[redacted]	[redacted]	19,595
Applying the adjusted resource use inputs based on Meads et al.	B20	Additional health care resource use for those not visually impaired	No additional health care resource use for those not visually impaired	[redacted]	[redacted]	22,277
Using the EAG's proposed health states with LEROS data for idebenone	B21	Company preferred model, EAP	EAG preferred model, LEROS	[redacted]	[redacted]	26,798

Abbreviations: CI, confidence intervals; CaRS, Case record survey; EAG, external assessment group; EAP, Extended access programme; LOCF, last observation carried forward; NHS, National Health Service; PSA, probability sensitivity analyses; PSS, Personal social services; RWE, Real-world evidence; SE, standard error

5.3 Model validation and face validity check

The company states that the cost-effectiveness model was quality assured by a senior health economist not involved in the model build who reviewed the model for coding errors, inconsistencies, and plausibility of inputs and outputs. The company also states that the model was subject to stress testing of extreme scenarios to test for technical modelling errors and plausibility of results.

6 Additional economic analysis undertaken by the EAG

6.1 Model corrections

The External Assessment Group (EAG) did not identify any model errors requiring correction.

6.2 Exploratory and sensitivity analyses undertaken by the EAG

In Section 4 of this report, the EAG has described several scenarios that warranted further exploration in addition to the company's own sensitivity and scenario analyses. The EAG cost effectiveness scenarios comparing idebenone to SoC (standard of care) are listed below, with results presented in Section 6.3.

- Applying the idebenone LEROS transition probabilities to SoC patients after RHODOS – Section 4.2.4;
- Adjusting the idebenone treatment discontinuation weighted average calculation and increasing the proportion who discontinue treatment to 10.8% – Section 4.2.4;
- Adjusting the EQ-5D utilities calculated from Lawrence *et al.*[2] to include patients from the Republic of Ireland – Section 4.2.6;
- Applying healthcare resource costs associated with visual impairment according to Meads *et al.*[3] – Section 4.2.7;
- Applying supportive living cost as a one-off cost – Section 4.2.7;
- Applying outpatient care cost as a one-off cost – Section 4.2.7.

6.3 EAG scenario analysis

All additional scenarios conducted by the EAG were done so using the EAG's preferred model structure. The results of EAG scenarios are outlined in Table 35.

Table 35. Results of the EAG's scenario analyses

	Results per patient	Comparator	Intervention	Incremental value
-	Company base case			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	18,758
0	EAG preferred model structure			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]

	ICER (£/QALY)	-	-	27,053
1	Applying the LEROS transition probabilities to SoC patients after RHODOS			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	59,061
2	Adjusting the idebenone treatment discontinuation weighted average calculation and increasing the proportion who discontinue treatment to 10.8%			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	30,316
3	Adjusting the EQ-5D utilities calculated from Lawrence *et al.*[2] to include patients from the Republic of Ireland			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	27,780
4	Applying additional healthcare resource costs according to Meads *et al.*[3]			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	31,631
5	Applying supportive living cost as a one-off cost			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	35,456
6	Applying outpatient care cost as a one-off cost			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER (£/QALY)	-	-	28,128

Abbreviations: EAG, External Assessment Group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

6.4 EAG preferred assumptions

Table 36 outlines the EAG's preferred assumptions and the independent and cumulative impact on the ICER of each assumption. The EAG's base case deterministic and probabilistic cost-effectiveness results are provided in Table 37.

- EAG preferred model structure;
- Using the LEROS study to derive the idebenone long term treatment effects;

- Applying the idebenone transition probabilities to SoC patients after RHODOS;
- Adjusting the idebenone treatment discontinuation weighted average calculation and increasing the proportion who discontinue treatment to 10.8%;
- Using the EQ-5D utilities calculated from Lawrence *et al.*[2] that include patients from the Republic of Ireland;
- No carer disutility applied;
- Applying additional healthcare resource costs according to Meads *et al.*[3] ;
- Applying supportive living cost as a one-off cost;
- Applying outpatient care cost as a one-off cost.

Table 36. EAG preferred model assumptions

Preferred assumption	Section in EAG report	Independent ICER (£/QALY)	Cumulative ICER (£/QALY)
Company base case	-	18,758	18,758
EAG preferred model structure	4.2.2	27,053	27,053
Using the LEROS study to derive the idebenone long term treatment effect*	4.2.4	28,459	35,736
Applying the LEROS transition probabilities to SoC patients after RHODOS	4.2.4	59,061	99,366
Adjusting the idebenone treatment discontinuation weighted average calculation and increasing the proportion who discontinue treatment to 10.8%	4.2.4	21,022	111,280
Using the EQ-5D utilities calculated from Lawrence et al[2] that include patients from the republic of Ireland**	4.2.6	27,780	109,432
No carer disutility applied	4.2.6	21,019	127,207
Applying additional healthcare resource costs according to Meads et al.[3]**	4.2.7	31,631	128,419
Applying supportive living cost as a one-off cost	4.2.7	25,899	129,704

Applying outpatient care cost as a one-off cost	4.2.7	19,595	130,269

Abbreviations: EAG, External Assessment Group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

*EAP discontinuation rate applied given the limited LEROS study duration

**EAG preferred model structure assumption also required

Table 37. EAG base case results

Interventions	Total Costs (£)	Total LY	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
Deterministic results							
SoC	[redacted]	[redacted]	[redacted]	-	-	-	-
Idebenone	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	130,269
Probabilistic results*							
SoC	[redacted]	-	[redacted]	-	-	-	-
Idebenone	[redacted]	-	[redacted]	[redacted]	-	[redacted]	126,422

Abbreviations: ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality-adjusted life-year; SoC, standard of care.

*Probabilistic results do not account for idebenone and SoC treatment effectiveness uncertainty

6.5 Conclusions of the cost effectiveness sections

Overall, the EAG considers that many of the company's assumptions informing the economic model are bias in favour of idebenone, leading to its cost-effectiveness being overestimated.

Structurally the company's model does not align with key HRQoL sight related thresholds outlined by the EAG's clinical experts and the high number of health states used in the company's base case lead to health state transitions being impossible when coupled with the equally critical issue of the small number of patient observations informing the model.

Of all the issues identified by the EAG in the company's model, the EAG considers the modelling of the SoC treatment effect is the most impactful. The EAG has identified that SoC logMAR recovery is considerably worse in the model than in the RHODOS, RHODOS-OFU and the LEROS natural history matched analysis, with SoC logMAR worsening 2.28 times as much in the model compared to the end of the RHODOS study (6 months) if considering the mITT population and 3.84 worse with respect to the ITT population. Additionally, while the RHODOS OFU and CaRS -II studies recorded that mean SoC logMAR recovered to baseline values by three and four years respectively, these results are also not reflected in the model.

While limited available patient data is a key issue in many health technology assessments with rare genetic conditions influencing the ability of companies to provide a robust SoC treatment effect, the EAG considers that there appears to be sufficient trial data to inform a robust SoC treatment effect, which the company has not utilised. For example, the EAG requested a scenario which utilised all available and appropriate CaRS patient data (approximately 944 observations); however, only 169 observations were used in the scenario. The EAG, therefore, considers that the modelled SoC treatment effect is underestimated and highly uncertain.

The company has additionally failed to account for the uncertainty in the estimates of treatment effectiveness within their deterministic and probabilistic sensitivity analyses (PSA). While the company has attempted to justify the exclusion of treatment effectiveness uncertainty from the sensitivity analysis by suggesting it's inclusion will create substantial uncertainty in the results, the EAG considers this a critical flaw in the development of the model. Investigating the impact of parameter uncertainty on the incremental cost effectiveness ratio (ICER) is a critical step in the evaluation of new health technologies. The NICE Guide to the Methods of Technology Appraisal states that PSA results are no longer simply recommended but are a mandatory requirement for all cost-effectiveness models submitted to NICE.

7 References

1. Chen BS, Yu-Wai-Man P, Horton M. Psychometric Validity of the Visual Function Index in Leber Hereditary Optic Neuropathy. *Transl Vis Sci Technol* 2023; **12**: 23.
2. Lawrence C, Bojakowski S, Williams E, Lloyd A. Utility Estimates for Vision Loss in Leber Hereditary Optic Neuropathy POSTER CODE: PCR227. *Value in Health* 2023; **26**: SUPPLEMENT-S355.
3. Meads CaH, C. What is the cost of blindness? *British Journal of Ophthalmology* 2003; **87**: 1201-4.
4. Poincenot L, Pearson AL, Karanjia R. Demographics of a Large International Population of Patients Affected by Leber's Hereditary Optic Neuropathy. *Ophthalmology* 2020; **127**: 679-88.
5. Yu-Wai-Man P, Chinnery PF. *Leber Hereditary Optic Neuropathy.* In: Adam MP, Feldman J, Mirzaa GM, Pagon RA, Wallace SE, Bean LJH, et al., editors. GeneReviews((R)). Seattle (WA)2021.
6. Ghelli A, Zanna C, Porcelli AM, Schapira AH, Martinuzzi A, Carelli V, et al. Leber's hereditary optic neuropathy (LHON) pathogenic mutations induce mitochondrial-dependent apoptotic death in transmitochondrial cells incubated with galactose medium. *J Biol Chem* 2003; **278**: 4145-50.
7. Lin CS, Sharpley MS, Fan W, Waymire KG, Sadun AA, Carelli V, et al. Mouse mtDNA mutant model of Leber hereditary optic neuropathy. *Proc Natl Acad Sci U S A* 2012; **109**: 20065-70.
8. Mackey DA, Ong JS, MacGregor S, Whiteman DC, Craig JE, Lopez Sanchez MIG, et al. Is the disease risk and penetrance in Leber hereditary optic neuropathy actually low? *Am J Hum Genet* 2023; **110**: 170-6.
9. Topic Selection Oversight Panel. Highly Specialised Technologies (HST) criteria checklist: Idebenone for treating visual impairment in Leber's hereditary optic neuropathy [ID547]. 2023.
10. Yu-Wai-Man P, Griffiths PG, Brown DT, Howell N, Turnbull DM, Chinnery PF. The epidemiology of Leber hereditary optic neuropathy in the North East of England. *Am J Hum Genet* 2003; **72**: 333-9.
11. Newman NJ, Carelli V, Taiel M, Yu-Wai-Man P. Visual Outcomes in Leber Hereditary Optic Neuropathy Patients With the m.11778G>A (MTND4) Mitochondrial DNA Mutation. *J Neuroophthalmol* 2020; **40**: 547-57.
12. Yu-Wai-Man P, Turnbull DM, Chinnery PF. Leber hereditary optic neuropathy. *J Med Genet* 2002; **39**: 162-9.
13. Carelli V, Carbonelli M, De Coo IF, Kawasaki A, Klopstock T, Lagrèze WA, et al. International Consensus Statement on the Clinical and Therapeutic Management of Leber Hereditary Optic Neuropathy. *Journal of Neuro-Ophthalmology* 2017; **37**: 371-81.
14. Petrovic Pajic S, Jarc-Vidmar M, Fakin A, Sustar Habjan M, Brecelj J, Volk M, et al. Case report: Long-term follow-up of two patients with LHON caused by DNAJC30:c.152G>A pathogenic variant-case series. *Front Neurol* 2022; **13**: 1003046.
15. Royal National Institute of Blind People. The criteria for certification. 2023. Available from: https://www.rnib.org.uk/your-eyes/navigating-sight-loss/registering-as-sight-impaired/the-criteria-for-certification/. Date accessed: November 2023.
16. Combal J-P, Chipot A, Holzinger E, Hillgenberg M, Roche M, Sargent A, et al. Leber Hereditary Optic Neuropathy (LHON): a qualitative study on the experiences and hopes of patients and caregivers. 2017.
17. Harper L. Living with leber hereditary optic neuropathy: Exploring experiences and perceptions of a disruptive mitochondrial condition, Cardiff University. 2018.
18. Steinberg EP, Tielsch JM, Schein OD, Javitt JC, Sharkey P, Cassard SD, et al. The VF-14. An index of functional impairment in patients with cataract. *Archives of ophthalmology (Chicago, Ill : 1960)* 1994; **112**: 630-8.
19. Kirkman MA, Korsten A, Leonhardt M, Dimitriadis K, De Coo IF, Klopstock T, et al. Quality of Life in Patients with Leber Hereditary Optic Neuropathy. *Invest Ophthalmol Vis Sci* 2009; **50**: 3112.

20. Mashima Y, Kigasawa K, Shinoda K, Wakakura M, Oguchi Y. Visual prognosis better in eyes with less severe reduction of visual acuity one year after onset of Leber hereditary optic neuropathy caused by the 11,778 mutation. *BMC Ophthalmol* 2017; **17**: 192.
21. Kirkman MA, Yu-Wai-Man P, Korsten A, Leonhardt M, Dimitriadis K, De Coo IF, et al. Gene-environment interactions in Leber hereditary optic neuropathy. *Brain* 2009; **132**: 2317-26.
22. The M, Healthcare products Regulatory A. Great Britain Marketing Authorisations (MAs) for Centrally Authorised Products (CAPs). *GOVUK* 2022.
23. Brown GC. Vision and quality-of-life. *Trans Am Ophthalmol Soc* 1999.
24. EMA. Raxone (idebenone) European Public Assessment Report. 2015.
25. Klopstock T, Yu-Wai-Man P, Dimitriadis K, Rouleau J, Heck S, Bailie M, et al. A randomized placebo-controlled trial of idebenone in Leber's hereditary optic neuropathy. *Brain* 2011; **134**: 2677-86.
26. Awmsg. Idebenone (Raxone®). *All Wales Therapeutics and Toxicology Centre* 2021.
27. SMC. Idebenone (Raxone) 150 mg film-coated tablets (SMC No. 1226/17). Scottish Medicines Consortium, 2017 07 April. Report No.
28. Electronic Medicines C. Raxone 150 mg film-coated tablets - Summary of Product Characteristics (SmPC) - (emc). 2022.
29. Ema. Raxone (idebenone). European Medicines Agency. *European Medicines Agency* 2018.
30. Klopstock T, Tomasso L, Lloria X. Long-term efficacy and safety of idebenone in patients with LHON in the subacute/dynamic phase: Results from the LEROS study. *Investig Ophthalmol Vis Sci* 2022; **63**.
31. Ishikawa H, Masuda Y, Ishikawa H, Shikisima K, Goseki T, Kezuka T, et al. Characteristics of Japanese patients with Leber's hereditary optic neuropathy and idebenone trial: a prospective, interventional, non-comparative study. *Jpn J Ophthalmol* 2021; **65**: 133-42.
32. van Everdingen JAM, Pott JWR, Bauer NJC, Krijnen AM, Lushchyk T, Wubbels RJ. Clinical outcomes of treatment with idebenone in Leber's hereditary optic neuropathy in the Netherlands: A national cohort study. *Acta ophthalmologica* 2022; **100**: 700-6.
33. Catarino CB, Von Livonius B, Priglinger C, Banik R, Matloob S, Tamhankar MA, et al. Real-World Clinical Experience With Idebenone in the Treatment of Leber Hereditary Optic Neuropathy. *Journal of Neuro-Ophthalmology* 2020; **40**: 558-65.
34. Chiesi Farmaceutici SpA. Data on file: Case record survey. The natural history of Leber's Hereditary Optic Neuropathy (LHON). SNT-IR-006. Version 1.0. 2014.
35. Chiesi Farmaceutici SpA. FINAL STUDY REPORT SNT-CSR-002. HISTORICAL CASE RECORD SURVEY OF VISUAL ACUITY DATA FROM PATIENTS WITH LEBER'S HEREDITARY OPTIC NEUROPATHY (LHON). 2019.
36. Chiesi Farmaceutici SpA. A Non-interventional Study of Clinical Experience in Patients Prescribed Raxone® for the Treatment of Leber's Hereditary Optic Neuropathy (LHON). SNT-IV-003 (PAROS). Amended Clinical Study Rerport. Version 2.0. Data on File2022.
37. ClinicalTrials.gov. Study to Assess Efficacy,Safety and Tolerability of Idebenone in the Treatment of Leber's Hereditary Optic Neuropathy (RHODOS). NCT00747487. Last update 2013.
38. Chiesi Farmaceutici SpA. Data on file: Update Report - SNT-EAP-001. Expanded Access Program (EAP). Raxone® in the Treatment of Patients. 2015.
39. ClinicalTrials.gov. Study to Assess the Efficacy and Safety of Raxone in LHON Patients (LEROS). NCT02774005. Last update 2021. 2016.
40. ClinicalTrials.gov. RHODOS Follow-up Single-visit Study. RHODOS-OFU. NCT01421381. Last update 2013.
41. Royal National Institute of Blind P. The criteria for certification. *RNIB* 2016.
42. Poincenot L, Pearson A, Rustum K. Demographics of a Large International Population of Patients Affected by Leber's Hereditary Optic Neuropathy, 2019. Available from:

https://www.sciencedirect.com/science/article/pii/S016164201932295X. Date accessed: 19 Jan 2024.
43. Faria R, Hernandez Alava A, M,, Manca A, Wailoo A, J. NICE DSU Technical Support Document 17: The use of observational data to inform estimates of treatment effectiveness for Technology Appraisal: Methods for comparative individual patient data. *Available from http://wwwnicedsuorguk* 2015.
44. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996; **313**: 275-83.
45. Efficace F, Bottomley A, Osoba D, Gotay C, Flechtner H, D'Haese S, et al. Beyond the development of health-related quality-of-life (HRQOL) measures: a checklist for evaluating HRQOL outcomes in cancer clinical trials--does HRQOL evaluation in prostate cancer research inform clinical decision making? *J Clin Oncol* 2003; **21**: 3502-11.
46. Nice. Voretigene neparvovec for treating inherited retinal dystrophies caused by RPE65 gene mutations. HST 11. NICE.
47. Nice. Ranibizumab for treating choroidal neovascularisation associated with pathological myopia. TA298. 2013.
48. Nice. Ranibizumab for treating visual impairment caused by macular oedema secondary to retinal vein occlusion | Guidance | TA283. 2013.
49. Nice. Ranibizumab for treating diabetic macular oedema | Guidance | TA274. 2013.
50. England N. NHS Reference Costs 2021/22, 2023. Available from: https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/. Date accessed: July 2023.
51. National Institute for Health and Care Excellence (NICE). NICE health technology evaluations: the manual. Process and methods [PMG36]. 2022. Available from: https://www.nice.org.uk/process/pmg36/resources/nice-health-technology-evaluations-the-manual-pdf-72286779244741. Date accessed: Sept 2022.
52. NICE. NICE health technology evaluations: the manual (Process and methods [PMG36]), 2022. Available from: https://www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation. Date accessed: 03 Oct 2023.
53. Nice. Aflibercept solution for injection for treating wet age-related macular degeneration. TA294. 2013.
54. Czoski-Murray C, Carlton J, Brazier J, Young T, Papo NL, Kang HK. Valuing Condition-Specific Health States Using Simulation Contact Lenses. *Value in Health* 2009; **12**: 793-9.
55. Rentz AM, Kowalski JW, Walt JG, Hays RD, Brazier JE, Yu R, et al. Development of a Preference-Based Index from the National Eye Institute Visual Function Questionnaire-25. *JAMA ophthalmology* 2014; **132**: 310-8.
56. Hernández Alava M, Pudney S, Wailoo A. Estimating EQ-5D by Age and Sex for the UK. NICE DSU Report. 2022.
57. Lawrence C, Williams E, Bojakowski S, Lloyd A. PCR177 Characterizing Patient Burden in Leber's Hereditary Optic Neuropathy Using Mixed Methods Research. *Value in Health* 2023; **26**: S345.
58. Hernández Alava M, Pudney S, Wailoo A. Estimating the Relationship Between EQ-5D-5L and EQ-5D-3L: Results from a UK Population Study. *PharmacoEconomics* 2023; **41**: 199-207.
59. Rowen D BJ, Wong R, Wailoo A. Measuring and valuing health-related quality of life when sufficient EQ-5D data is not available. NICE DSU Report. 2020.
60. Nice. Ranibizumab and pegaptanib for the treatment of age-related macular degeneration. TA155. 2008.
61. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharm Stat* 2011; **10**: 150-61.

8 Appendices

8.1 Quality assessment

Table 38. Quality of effectiveness estimates from non-randomised studies (QuEENS) checklist for the propensity score matching (PSM) analysis of LEROS ITT idebenone-treated patients compared to CaRS SoC treated patients.[43]

Question	Answer
Q1: Have different methods been compared within the study?	No - Only a single method, propensity score matching, was used
Q2: Have the results of the study been compared to others in the literature?	Not compared – no other estimates were available
Q3: Is there a discussion of what treatment effect is identified and of the assumptions needed?	No discussion of either
Q4: Is the model chosen consistent with the outcome variable if using a parametric method?	Yes/unclear. ANCOVA is used but no justification was provided. LogMAR values from LEROS included those measured in a continuous fashion through ETDRS charts, and ordinal measurements of "finger counting", "hand motion" and "light perception". The distribution of logMAR values was presented as binned categories, but skewness was not assessed. LogMAR values from the CaRS studies were converted from Snellen measurements.
Q5: Were any checks conducted on the model specification?	No checks reported
Q6: On selection: Is the assumption of selection on observables assessed?	Partially, the EAG suggested some key baseline characteristics for matching, to which the Company added others, albeit without explicit justification or discussion of whether any prognostic factors were unmatched and/or not available.
Q7: What checks were conducted to assess overlap?	No checks reported, although there was a reasonable degree of overlap for each baseline characteristic between unmatched populations.
Q8: Has balancing of the covariates been checked after matching and propensity score methods?	Yes, through standardised mean differences presented in Figure 1 of Company response to clarification.
Q9: Is the propensity score function sufficiently flexible?	Unclear/unlikely to be sufficiently flexible, it was not reported that polynomial or interaction terms were allowed for in the calculating of the propensity score.
Q10: Are potential IVs excluded from the set of conditioning variables?	Yes, other than sex, each variable included in the matching set is likely a meaningful prognostic factor
Q11: Data quality: Are there data quality issues?	(a) Data and definitions comparable for treated and control groups: Partially, logMAR values directly measured in LEROS but are converted Snellen measures in CaRS;

	(b) Treated and controls come from the same area or environment: No, CaRS is a historical natural history study whereas LEROS a clinical trial; (c) Rich set of variables used for matching: Yes, a reasonable set of variables were used for matching; (d) Reasonable sample sizes: Partially, given the rarity of the disease the sample sizes appear reasonable, despite no formal consideration of statistical power in the matched-control analysis. However, the EAG is concerned that only a single time point (24 months) was used from the follow-up data.
Q12: For Nearest Neighbour: Has bias adjustment been conducted if more than one variable was included when matching on covariates?	No bias adjustment was reported.
Q13: Is the choice of replacement (with/without) reasonable?	Partially, the decision was not justified but most NH patients were successfully matched without replacement.
Q14: Is the choice of the number of calliper matching reasonable?	Yes, a standard calliper width was used that did not result in an excessive loss of sample size.[61]
Abbreviations: ANCOVA, analysis of covariance; EAG, external assessment group; ETDRS, Early Treatment of Diabetic Retinopathy Study; logMAR, logarithm of the minimum angle of resolution; NH, natural history	