Fluocinolone acetonide intravitreal implant for treating chronic diabetic macular oedema

NICE ID1421

Produced by: Warwick Evidence

Authors (Warwick Medical School unless stated otherwise):
Ewen Cummins, health economist, McMDC, Harrogate
Lena Al Khudairy
Noemi Lois, professor of ophthalmology, Queen's University, Belfast
Martin Connock
Rhona Johnston, computer programmer, McMDc, Harrogate
Daniel Gallacher
Pamela Royle
Christopher Stinton
Norman Waugh

Correspondence to:
Professor Norman Waugh
Division of Health Sciences
Warwick Medical School
University of Warwick, Coventry CV4 7AL norman.waugh@warwick.ac.uk

Date completed: 6th March 2019

Source of funding: This report was commissioned by the NIHR Programme as project number 127335
Competing interests: None
Rider: The views expressed in this report are those of the authors and not necessarily those of the HTA Programme.
This report should be referenced as: Fluocinolone acetonide intravitreal implant for treating chronic diabetic macular oedema. A single technology appraisal. Warwick Evidence, 2019.

Conflicts. No financial competing interests. Norman Waugh is an active patient member of the Macular Society but has a form of macular disease unrelated to diabetes and DMO.

Contents

List of Tables

List of Figures

List of Abbreviations

AAO	American Academy of Ophthalmology
AE	Adverse event
AIC	Akaike' Information Criterion
AMD	Age-related macular degeneration
AREDS	Age-Related Eye Disease Study
ARVO	Association for Research in Vision and Ophthalmology
AUC	Area under the curve
BCVA	Best corrected visual acuity
BIC	Bayesian Information Criterion
BSE	Better-seeing eye
CCG	Clinical Commissioning Group
CEAC	Cost-effectiveness acceptability curves
CRD	Centre for Reviews and Dissemination
CSMO	Clinically significant macular oedema
DCCT/EDIC	Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications
DMO	Diabetic macular oedema
DR	Diabetic retinopathy
DRCRN	Diabetic Retinopathy Clinical Research Network
DVLA	Driver and Vehicle Licensing Agency
EMA	European Medicines Association
EoP	End of Period
ERG	Evidence review group
EURETINA	European Society of Retina Specialists Congress
FA	Fluorescein angiography
FAc	Fluocinolone acetonide
FAME	Fluocinolone Acetonide for Macular Edema
FDA	Food and Drug Administration
FTH	Central foveal thickness
HbA1c	Glycated haemoglobin
HRG	Healthcare Resource Group
HS	Health state
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio
ICE-UK	Iluvien Clinical Evidence study in the United Kingdom
IOP	Intraocular pressure
IRISS	Iluvien Registry Safety Study
ITT	Intention-to-treat

LYG	Life years gained
LYs	Life years
META-EYE	Meta-analysis for Eye disease
MHRA	Medicines and Healthcare products Regulatory Agency
MMRM	Mixed Model for Repeated Measures
NEI VFQ-25	National Eye Institute Visual Functioning Questionnaire- 25 item
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMA	Network meta-analysis
NPDR	Non-proliferative diabetic retinopathy
OCT	Optical coherence tomography
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSUR	Periodic Safety Update Report
QALYs	Quality-adjusted life years
QIC	Quasi-likelihood Information Criterion
Qol	Quality of life
RCT	Randomized controlled trial
RR	Relative risk
SD	Standard deviation
SE	Standard error
SGLT2	Sodium-dependent glucose transporter 2
SLR	Systematic literature review
SmPC	Summary of product characteristics
SoP	Start of Period
STA	Single technology assessment
TPMs	Transition Probability Matrices
TTO	Time trade off
VA	Visual acuity
VEGF	Vascular endothelial growth factor
VFQ	Visual Function Questionnaire
VFQ-UI	Visual Functioning Questionnaire-Utility Index
WSE	Worst-seeing eye

1 SUMMARY

1.1 Critique of the decision problem in the Alimera submission

The patient group was identified by NICE as:

"Eyes with phakic lenses and with visual impairment associated with chronic diabetic macular oedema considered insufficiently responsive to available therapies".

Phakic eyes still have their natural lens, i.e. have not had cararacts removed and replaced by artificial lenses. Those with artificial lenses are called "pseudophakic". Available therapies approved by NICE include laser photocoagulation (for thinner retinas, with central retinal thickness less than 400 microns) and the anti-VEGF drugs, ranibizumab and aflibercept (for thicker retinas). The NICE scope identifies the patient group as "People with chronic diabetic macular oedema that is insufficiently responsive to available therapies who have phakic lenses" but does not define "insufficiently responsive" or "chronic".

The Alimera Summary says (A7, fourth bullet) that

"In clinical practice, 0.2 µg/day fluocinolone acetonide implant would only be considered for use in phakic patients in the following instances: a) in patients with pre-existing cataract and who would require cataract surgery in the next 1-3 years; b) in a small group of patients who are contraindicated for first line therapies or are needle phobic where the benefit of protecting the retina outweighs the risk of cataract formation."

The submission does not say how cataract was diagnosed or defined. It is difficult to predict which cataracts will need extraction "in the next 1-3 years" unless there is already a clear cataract present at baseline (e.g. $\geq$2 nuclear sclerosis or $\geq$2 posterior subcapsular cataract as per AREDS classification). But one reason for exclusion from FAME was;

"14. *Any lens opacity which impairs visualization of the posterior pole or significantly impairs vision, in the opinion of the investigator"*

The patients in clause (a) will be a subgroup of all phakic patients, who comprise three subgroups;

- Those who are phakic at baseline and at 3-year follow-up
- Those who are phakic at baseline with no cataract, but who develop cataract and have it removed, becoming pseudophakic

- Those who are phakic at baseline, but who do have detectable cataract, have it removed later, and become pseudophakic – the clause (a) patients.

As regards the first subgroup, Alimera provided no data on the phakic-phakic subgroup in their submission, because it was a very small subgroup with only 17 patients in the FAME trial, and no data were reported in the analysis by Yang et al which underpins this Alimera submission. However almost all have cataract at baseline, so they would receive fluocinolone.

The second group, no baseline cataracts, is excluded by Alimera in their Summary, but those patients appear to be included in the main submission and economic modelling. In discussion with Alimera, it was clarified that all phakic patients are treated whether they have cataract or not.

The Alimera approach to comparators is more detailed than the NICE scope, which mentioned only "established clinical management without fluocinolone". Alimera say;

"The following technologies alone or in combination with laser photocoagulation: ranibizumab, aflibercept, dexamethasone intravitreal implant not for phakic, bevacizumab ([off-label use] for people in whom other DMO treatments are unsuitable)."

However, the patients for whom fluocinolone is being considered have failed to respond to such treatments, so the expensive drugs, ranibizumab and aflibercept, are unlikely to be cost-effective. We lack evidence on this patient group. Alimera's view, supported by data from an observational study, is that if fluocinolone is not available, treatment with anti-VEGF drugs is likely to continue despite lack of evidence of efficacy, because of a lack of effective alternatives.

The FAME patients could not match the whole patient group in the NICE scope, because the scope says "*People with chronic diabetic macular oedema that is insufficiently responsive to available therapies*" which includes the anti-VEGF drugs (ranibizumab, aflibercept and bevacizumab) as well as laser photocoagulation.

Alimera therefore provided evidence from recent observational studies of fluocinolone in routine care, including for patients who had had insufficient responses to laser and/or anti-VEGF drugs. These provided evidence that fluocinolone was also effective after anti-VEGF failure.

1.2 Summary of clinical effectiveness evidence submitted by the company

Alimera submitted data from the initially phakic subgroup from the FAME (Fluocinolone Acetonide for Macular Edema) trial, and from some observational studies, some reporting experiences with

fluocinolone in routine care, sometimes called “real life” studies. They did not submit any NMA, and the ERG supports that decision.

The outcomes from the FAME patients who were phakic at baseline but pseudophakic at 3 years, show that treatment with fluocinolone improves mean BCVA by about 8 letters, followed by a decline of about 7 letters as cataract develops, then an improvement after cataract extraction back to 8 letters above baseline by about 34 months (from Figure B2.6, page 75). The sham group gained only 2 letters by 34 months. So the mean marginal gain in BCVA by 3 years is about 6 letters. The improvement in the sham group may be spontaneous resolution, or a response to improved glycaemic control.

1.3 Summary of the ERG’s critique of clinical effectiveness evidence submitted

FAME was a good quality trial but done at a time before anti-VEGF drugs had come into common use, so most patients had had only laser treatment. They were recruited into FAME if foveal thickness was 250 microns or more despite at least one prior macular laser treatment. So some may have responded to laser treatment but not sufficiently so. For example, someone with pre-treatment retinal thickness of 350 microns, reduced to 260 after laser treatment, would have been eligible. The FAME papers do not say how many had only one laser treatment. So inclusion in FAME does not necessarily mean patients were unresponsive to laser, but only that laser treatment did not reduce retinal thickness to 250 microns or less.

The most useful of the observational studies are in patients with chronic DMO that has not responded to previous treatment, including with anti-VEGF drugs. The improvements in BCVA are not dramatic – a mean 5.3 letters at 24 months in the Medisoft study, and a mean of about 3 letters at 12 months in the ICE-UK study. However these results are mainly in pseudophakic patients. Some studies reported a decline in VA in the year prior to fluocinolone treatment. So stability without improvement may be a useful outcome.

The Alimera submission assumed that if fluocinolone was not available, patients would continue on anti-VEGF treatment. There are some important unknowns including;

1. In patients not responding to anti-VEGFs, what is the marginal gain with fluocinolone compared to continuing anti-VEGFs?
2. What is the cost per QALY of continuing anti-VEGFs in people who don’t respond to these drugs? We suspect this would be high unless bevacizumab was used. And perhaps high even then.
3. In patients not responding to anti-VEGFs, what is the marginal gain with fluocinolone compared to no treatment?

1.4 Summary of cost effectiveness submitted evidence by the company

The company developed a de novo markov model with a quarterly cycle and a 30 year time horizon, effectively a lifetime. Based upon ICE-UK data fluocinolone is compared to a usual care composite comparator of 28% no drug treatment / laser, 63% ranibizumab and 9% bevacizumab, so 72% anti-VEGF with clinical equivalence between ranibizumab and bevacizumab being assumed.

The model independently simulates a cohort of study eyes and a cohort of fellow eyes. These are then combined into bilateral health states. The eyes are distributed across 8 health states defined by 10 letter BCVA bands, though the best health state stretches from 86 letters and up, while the worst stretches from 25 letters and down. The baseline distributions across the health states for both the study eyes and the fellow eyes are taken from ICE-UK.

For the study eye the balance between those with and without cataract at baseline is taken from the Retro-IDEAL study: 50:50, which is considerably higher than FAME phakic patients. Due to the limited number of patients in ICE-UK and conflicting evidence on the impact of cataract removal, the company assumes identical distributions at baseline for those with and without cataract.

For the fellow eye the balance between those with and without DMO at baseline is assumed to be the percentage of fellow eyes in ICE-UK with a history of treatment for DMO: 77%. It is assumed that among fellow eyes with DMO at baseline the balance between those with and without cataract is as for the study eye: 50:50. It is also assumed that all fellow eyes are phakic at baseline.

Within the model patients divide into the four patient subsets of

- phakic without cataract,
- phakic with cataract,
- undergoing cataract surgery, or
- pseudophakic.

Probabilities of developing cataract and of having cataracts removed once having developed them are estimated from FAME for fluocinolone and for sham. Those receiving fluocinolone are modelled using the fluocinolone probabilities, while those on laser or anti-VEGF are modelled using the sham probabilities. These determine patient movements between the four patient subsets, with cataract surgery being a tunnel health state lasting a single cycle as patients transition between being phakic with cataract and being pseudophakic.

Sets of four quarterly transition probabilities matrices (TPMs), one for each of the four patient subsets above, are estimated from FAME data for both fluocinolone and for sham. In a pooled analysis, a set is estimated for both fluocinolone and for sham from the quarterly transitions during the 1st 3 months of FAME. A second set is estimated for both fluocinolone and for sham from the quarterly transitions during months 4-36 of FAME. Extrapolation during years 4, 5 and 6 of the model reapplies the second set.

For years 1-6 the company applies the FAME fluocinolone TPMs to model the study eyes in the fluocinolone arm that are on treatment. The company base case assumes that there will be no change in the BCVA of the study eyes in the comparator arm. A scenario analysis applies the FAME sham TPMs to the study eyes in the comparator arm with the intention of modelling the net effect observed during FAME.

The company also estimates an odds ratio of gaining letters of 1.54 for anti-VEGF relative to laser. This permits it to construct TPMs for the anti-VEGFs. A weighted average set of TPMs for a usual care comparator based upon 28% experiencing the sham TPMs and 72% experiencing the anti-VEGF TPMs is then constructed. The company does not apply this composite comparator set of TPMs to the study eyes on usual care. It applies it to the fellow eyes on usual care, usual care being assumed for the fellow eyes with bilateral DMO in both arms.

Those who cease treatment are assumed to have a constant BCVA during years 1-6. The FAME 64% proportion of fluocinolone patients not achieving a 36 month 15 letter gain is modelled as discontinuing fluocinolone at the end of year 3, while the remaining 36% receive a 2nd fluocinolone implant. It is assumed that no patients in the usual care arm discontinue treatment.

From the end of year 6 it is assumed that all cease treatment. Over the next 24 years there are no net treatment costs. An equal probability of worsening BCVA is applied in both arms, which maintains the modelled year 6 net BCVA gain for the next 24 years.

Treatment specific adverse events are included. Mortality multipliers for diabetes and for blindness are also included.

The resulting distributions of study eyes and fellow eyes during each quarterly model cycle are combined into bilateral health states. The company estimates bilateral quality of life values by applying a published time trade off mapping function for the NEI-VFQ-25 to the FAME NEI-VFQ-25 data specific to each bilateral health state, with some subsequent smoothing of results between health states. These yield the main QALY estimates by arm.

Quality of life effects of adverse events and injection anxiety are also included but these have little impact upon results.

The company assumes all in the fluocinolone arm receive 1 implant at baseline, and those retreated at the end of year 3 receive another 1 implant.

ICE-UK data suggests that in the year prior to the fluocinolone implant those receiving anti-VEGF had a mean of around 4 injections, and in the year after the fluocinolone implant around a third of patients continued to receive anti-VEGF injections with roughly the same frequency. Rather than model this in the fluocinolone arm, the company nets these out to estimate that patients on anti-VEGF in the comparator arm receive roughly a net additional 3 anti-VEGF injections compared to the fluocinolone arm. As noted above, all patients in the comparator are assumed to remain on treatment until the end of year 6, but the number of anti-VEGF injections is assumed to linearly decline to zero by the end of year 6. Those receiving laser are assumed to have an average of 1.2 administrations each year.

Administration of each treatment is costed at £108.

Treatment specific monitoring costs are applied, fluocinolone monitoring frequency during year 1, 2 and 3 being taken from FAME. The monitoring frequency for anti-VEGFs is taken from the NICE STA of aflibercept, with 12 in year 1, 6.3 in year 2 and 4.0 in year 3. For fluocinolone the year 3 monitoring frequency is applied in years 4, 5 and 6 while for the comparator arm the average monitoring frequency over years 1, 2 and 3 is applied in years 4, 5 and 6.

The company estimates an annual cost of blindness of £19,795.

The company base case estimates that fluocinolone results in a net drug cost of £2,306 and additional adverse event costs, but that these are more than offset by administrative savings of £487, monitoring savings of £1,713 and costs of blindness savings of £987 resulting in an overall net saving of £330. A small overall survival gain of 0.043 discounted life years is estimated but the net gain of 0.236 QALYs main arises from the improved BCVA distribution through time. As a consequence, the company estimates that fluocinolone dominates the composite comparator.

In the opinion of the ERG the main company scenario analyses that are of interest are:

- Modelling the net rather than the absolute FAME treatment effect result in a net cost of £523 and a net gain of only 0.036 QALYs and so an ICER of £14,753 per QALY.
- Retaining the absolute FAME treatment effect and assuming that those in the comparator arm have a constant BCVA but only incur the costs of laser slightly revised the net gain to 0.238 QALYs

due to a slightly different side effects profile. But the main effect is to increase net costs to £3,763 which results in an ICER of £15,842 per QALY.

- A scenario analysis that has not been parsed by the ERG estimates 1st line fluocinolone use against fluocinolone use in the comparator arm only after cataracts have been removed. The costs in the fluocinolone arm and the composite comparator arm are identical to one another, and while the net gain falls to 0.145 QALYs it results in the company concluding that fluocinolone use should not be restricted to only once cataracts have been removed.
- A scenario analysis that has not been parsed by the ERG in which fluocinolone is reportedly modelled as being used in the fellow eye still results in dominance.
- Among those with cataract at baseline a net saving of £579 and a net gain of 0.282 QALYs hence a better costs effectiveness than the base case.
- Among those who are cataract free at baseline a net saving of £81 and a net gain of 0.190 QALYs hence a worse costs effectiveness than the base case, but still dominance.

1.5 Summary of ERG critique of cost effectiveness submitted evidence by the company

An immediate question is whether a composite comparator is appropriate. Those receiving laser may be a distinct subgroup from those receiving anti-VEGF in which case it would be better to model them as separate subgroups. If patients can move from anti-VEGF to either no further drug treatment or to fluocinolone this would seem to argue for a comparison between three alternative treatments.

In the opinion of the ERG the FAME sham arm reflects a degree of natural recovery, and the FAME fluocinolone arm will also include this natural recovery element. Given the patient group this natural recovery element may have been previously exhausted and may not apply. To the ERG this argues for the model removing the natural recovery element from the usual care arm and also from the fluocinolone arm, in order to model the FAME net treatment effect. But the model structure does not permit this. The closest is to retain the natural recovery element in both arms, and so model the FAME net treatment effect; i.e. apply the FAME fluocinolone TPMs in the fluocinolone arm and the FAME sham TPMs in the comparator arm.

The model validation section suggests that the model inputs and structure overestimate the treatment effectiveness of the FAME sham arm. As a consequence, the resulting clinical effectiveness estimates are probably biased against fluocinolone. But in itself this is not an argument against trying to model the FAME net treatment effect, just an observation that the model does not do it very well.

In the opinion of the ERG it is not suitable to model additional net benefits from ongoing treatment beyond the end of year 3, and it is more reasonable to assume that the net benefits that are realised at the end of year 3 are maintained but do not increase further. The ERG revised modelling only applies

this as a scenario analysis due to the validation problem when sham effectiveness is assumed for the comparator arm, as outlined above. But if the modelling of the comparator arm can be better aligned with the FAME sham arm effectiveness at 36 months, the ERG thinks that additional net gains in BCVA beyond this should not be modelled, other than those arising from cataract removal.

An issue is whether those with cataracts will have them removed more quickly if receiving fluocinolone than if receiving laser or anti-VEGF. The company treatment effect for this has a p-value that at best is [redacted]. The company account of its development of the cataract removal model would seem to argue for its exclusion.

Another issue is whether those with cataract at baseline should be considered separately from those without cataract at baseline. Similarly, should those with thinner retinas be considered separately from those with thicker retinas, does the company regression analysis sufficiently explore this given the 28% assumed to receive laser and what should be assumed for the anti-VEGFs in this regard.

A key consideration is whether patients will only ever receive 1 fluocinolone implant during each 3 year period or will receive more, 1.3 being received during the 3 years of FAME. In the opinion of the ERG the number observed during FAME should be applied, unless the company can clearly demonstrate that all additional implant during FAME occurred at or very close to 36 months. There are also questions around the proportion of those who receive a fluocinolone implant who continue to receive anti-VEGFs after implant, and in both arms what number of anti-VEGF administrations should be extrapolated among those receiving them.

The company base case cost for bevacizumab relates to cancer care. The 2018 NICE NG82 guideline for AMD estimated an ophthalmic cost of £49.

The company implementation of monitoring costs is in the opinion of the ERG biased and overestimates the savings that will result from this element. It is unreasonable to average anti-VEGF monitoring costs over years 1, 2 and 3 and carry forward this average, rather than carry forward the year 3 monitoring cost as in the fluocinolone arm modelling. The number of monitoring visits for fluocinolone may be too low, the SmPC appearing to suggest quarterly monitoring for the lifetime of the implant. The number of monitoring visits for those on anti-VEGF also looks more like that for a new anti-VEGF patient than that for a patient established on anti-VEGF treatment with only quarterly injections. It also seems more reasonable to the ERG for bilateral monitoring costs to be whichever of the monitoring cost of the study eye and the monitoring cost of the fellow eye is the more intensive.

The company estimate of the cost of blindness is biased and is roughly double the estimate that the company has used in previous assessments, despite citing the same paper. But the company model

incorrectly quarters this estimate meaning that the costs of blindness are roughly half what they should be.

The blindness mortality multiplier is oddly implemented. Individual eyes are modelled as dying and patient deaths are assumed to be the average of these. This overestimates deaths due to blindness. The model also does not include treatment and monitoring costs for either eye beyond the end of year 6. Any modelled survival benefit is consequently not properly costed. To the ERG any bias from not applying the blindness mortality multiplier is likely to be rather less than the bias if it is applied.

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

1.6.1 Strengths

There is considerable sophistication in the regression models that derive the transition probabilities from FAME. The model validation work suggests that while the modelled change in BCVA for fluocinolone between baseline and month 36 does not track the evolution of the FAME fluocinolone BCVA changes, by month 36 the modelled change in BCVA reasonably approximates that of FAME. Unfortunately, the same cannot be said of the modelling of the sham arm. The model validation work suggests that the modelled change in BCVA for sham between baseline and month 36 lies above the FAME sham BCVA changes and at month 36 remains above it by a reasonable amount.

The derivation of bilateral health state quality of life values based upon the NEI-VFQ-25 is a welcome innovation, though it should be borne in mind that the mapping algorithm study was sponsored by the company.

1.6.2 Weaknesses and areas of uncertainty

The main issues relate to how well the available clinical effectiveness estimates match the decision problem.

- Is the patient group homogeneous? If so, is it reasonable to consider a single composite comparator or should the composite comparator be split into its constituent parts? If not, are those with insufficient response to laser a distinct subgroup from those with insufficient response to anti-VEGF and should they be considered separately?
- For those on non-drug treatment receiving laser what is the clinical effect of remaining on current treatment? What would be the net effect of switching to fluocinolone? Does the FAME data used in the model match this? To what extent should this comparison be expected to differ from the

assessment of the chronic phakic during TA301, given that th current model is largely built upon FAME data?

- For those on anti-VEGF what is the clinical effect of remaining on current treatment? What would be the net effect of having anti-VEGF withdrawn? What would be the net effect of switching to fluocinolone? How should the odds ratios for ranibizumab estimated by the company and the ERG be viewed within this? How well does the FAME data used in the model inform this?

After year 6 the net gain in BCVA between the arms is maintained at no additional treatment cost for 24 years. If this extrapolation is not reasonable the model is biased in favour of fluocinolone, and a waning of the net gain in BCVA during the 24 years should be applied.

The company submission concentrates upon treatment in the study eyes. It does not particularly consider differentiating treatment in the fellow eyes by arm, and the company may not have faith in the model structure for this. The ERG has not rebuilt the elements of the model that would be used for this, but has some concerns about it. Treatment with fluocinolone in the fellow eyes also requires assumptions about their chronic status and phakic status, and where within the treatment pathway they fall. But this might result in a higher proportion of chronic, phakic eyes that are treated with fluocinolone being the better seeing eye. This might improve the cost effectiveness estimate for fluocinolone.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG makes a number of revisions to the company model as detailed in section 4.4 of this report. The most important of these revisions are:

- The net gain from fluocinolone over the comparator arm is modelled as the net gain observed during FAME. This is implemented by applying the FAME fluocinolone TPMs for the fluocinolone arm and the FAME sham TPMs for the comparator arm.
- The model for cataract removal without a treatment effect is applied.
- Over three years 1.3 fluocinolone implants will on average be used, as observed in FAME, rather than the 1 assumed by the company.
- The cost of bevacizumab for ophthalmic use is £49 as in the 2018 NG82 for AMD.
- The costs of monitoring during years 4, 5 and 6 are the costs of monitoring during year 3.
- Fluocinolone monitoring is quarterly after the 1st year.
- Anti-VEGF monitoring for those on established treatment is bi-monthly in the 1st year and then declines as per the company assumptions.
- The annual cost of blindness is roughly double that implemented in the company model.

The ERG provides pairwise comparisons with laser, with anti-VEGF and with the pooled composite comparator of the company base case, with the following revised base cases and scenario analyses.

Table 1. ERG revised base case ICERs and scenario analyses

Comparator	Laser	Anti-VEGF	Composite
Base case ICER	£334k	£176k	£223k
SA01a: 1.54 OR for anti-VEGF vs laser	..	Dominated	Dominated
SA01b: 1.23 OR for anti-VEGF vs laser	..	Dominated	Dominated
SA02: Fluocinolone net gain = FAME change from baseline	£25,550	£8,302	£13,300
SA03: No additional BCVA in years 4, 5 and 6	Dominated	Dominated	Dominated
SA04: Cataract removal regression with treatment effect	£226k	£113k	£146k
SA05a: 100% cataract at baseline	£187k	£90,959	£119k
SA05b: 0% cataract at baseline	£1.3mn	£1.0mn	£1.2mn
SA06a: 100% retinas < 400µm at baseline	£463k	£259k	£322k
SA06b: 100% retinas ≥ 400µm at baseline	£289k	£148k	£190k
SA07a: anti-VEGF 100% ranibizumab	..	£153k	..
SA07b: anti-VEGF 100% bevacizumab	..	£331k	..
SA08a: Natural history 2% worsening per quarter	£304k	£161k	£204k
SA08b: Natural history 5% worsening per quarter	£366k	£193k	£245k
SA09a: Brazier et al NEI-VFQ-25 QoL algorithm	£353k	£188k	£237k
SA09b: Brazier et al EQ-5D QoL algorithm	£913k	£614k	£718k
SA09c: Czoski-Murray QoL with WSE 15% of BSE	£198k	£99,467	£128k
SA09d: Czoski-Murray QoL with WSE 30% of BSE	£224k	£113k	£145k
SA10: Combine study eye and fellow eye joint distribution	£321k	£159k	£207k
SA11: Fluocinolone administration costs +33%	£337k	£179k	£226k
SA12: anti-VEGF administration costs -25%	..	£185k	£232k
SA13a: anti-VEGF 4 monitoring visits in year 1	..	£161k	£214k
SA13b: anti-VEGF 12 monitoring visits in year 1	..	£187k	£231k
SA14: 50% fluocinolone retreatment at end of year 3	£195k	£92,233	£122k
SA15: Only 1 fluocinolone implants every 3 years	£265k	£98,241	£148k
SA16: Blindness residential care 30% self-funded	£332k	£174k	£221k
SA17: 18 year time horizon	£425k	£235k	£293k
SA18: SA02 + SA03	£34,947	£13,176	£19,612
SA19: SA02 + SA03 + SA17	£40,224	£16,985	£23,998

The worsening cost effectiveness of fluocinolone compared to the company analyses arises from three main sources:

- The model apparently overestimating the effect of the FAME sham arm which reduces the modelled net gain at 36 months to below that observed during FAME, as outlined in greater detail in the validation section 4.2.11.
- The increase in the number of fluocinolone injections.
- The revisions to the treatment of monitoring costs which reduce the cost offsets arising from this source.

The cost effectiveness estimates where fluocinolone is dominated by the comparator are likely to be due to the model not estimating the FAME sham arm particularly well. But there remains a question about whether an odds ratio of gaining letters for the anti-VEGFs relative to laser should be applied. Would withdrawal of anti-VEGFs in the patient group of interest result in a loss of vision compared to them remaining on anti-VEGFs? And is this what is being modelled, or are those receiving laser and those receiving anti-VEGFs different patient subgroups who should be modelled separately?

In the light of the model apparently not estimating the FAME sham arm particularly well, it may seem tempting to fall back on a comparison that assumes the net effect from fluocinolone at 36 months is the FAME fluocinolone absolute change from baseline rather than the FAME net effect at 36 months. The ERG disagrees with this if the changes observed during FAME include an element of natural recovery in both arms, as seems likely. But the sensitivity analysis SA02 makes this assumption and there is a corresponding improvement in the cost effectiveness estimates.

Again due to the model apparently not estimating the FAME sham arm particularly well, the ERG revised base cases do not assume that for those remaining on treatment beyond year 3 they will retain their BCVA gains from baseline but there will not be further additional BCVA gains from treatment, other than from cataract surgery. This is SA03, and if the model is revised to better model the FAME sham arm in the opinion of the ERG it should form part of the base case.

The model is reasonably sensitive to whether the probability of cataract removal among those who have developed a cataract varies by treatment. The company statistical analysis finds the treatment effect to have a p-value of at best [redacted] and by the company argument it should not be included. If there is a treatment effect, it is difficult to apply the central estimate with much confidence.

The model outputs appear particularly sensitive when split into those with and those without cataract at baseline, though this is in part due to the small net QALY gains causing large swings in the ICERs. There may also be concerns about the clinical effects among those with thin retinas compared to those with thick retinas. The data is from FAME, so there are concerns about the retinal subgroup analyses when anti-VEGFs are the comparator, or are a large element of the composite comparator.

The company uses a mapping algorithm based upon a company sponsored time trade-off study of the NEI-VFQ-25 relationship with quality of life. This permits quality of life values to be associated with bilateral health states directly, rather than by assumption as has occurred in previous NICE assessments when using the values of the Czoski-Murray experimental lenses study.

The model estimates a net BCVA gain at the end of year 6. From this point all treatment is assumed to cease and there are no net treatment costs. But the end of year 6 net BCVA gain is extrapolated largely unchanged for the next 24 years which may not be reasonable. In the model there is no ready means of waning the net BCVA effect during this period, and the closest approximation is to reduce the time horizon to 18 years. The net QALY gains fall and the ICERs worsen.

1.8 Conclusions

The evidence base for clinical effectiveness is not ideal, given the absence of a trial of fluocinolone versus continuing anti-VEGF drugs in eyes that have not responded sufficiently to those. The ERG considers that several observational studies show that fluocinolone is beneficial in eyes that have responded poorly to anti-VEGF treatment. However uncertainties remain around the size of the effect.

Note on process

An unusually large number of clarification questions had to be submitted to Alimera, and there had to be several rounds of clarification questions. The last response from Alimera was received on 1st March. Not all issues could be resolved before the ERG report had to be submitted.

2 BACKGROUND

2.1 The macula

The macula is part of the retina at the back of the eye. It is only about 5mm across but is responsible for all of our central vision, most of our colour vision and for the fine detail of what we see, such as reading and recognising faces.

The macula has a very high concentration of photoreceptor cells (rods and cones) that detect light and send signals to the brain, which interprets them as images. The rest of the retina processes our peripheral (side) vision. Macular disease causes loss of central vision.

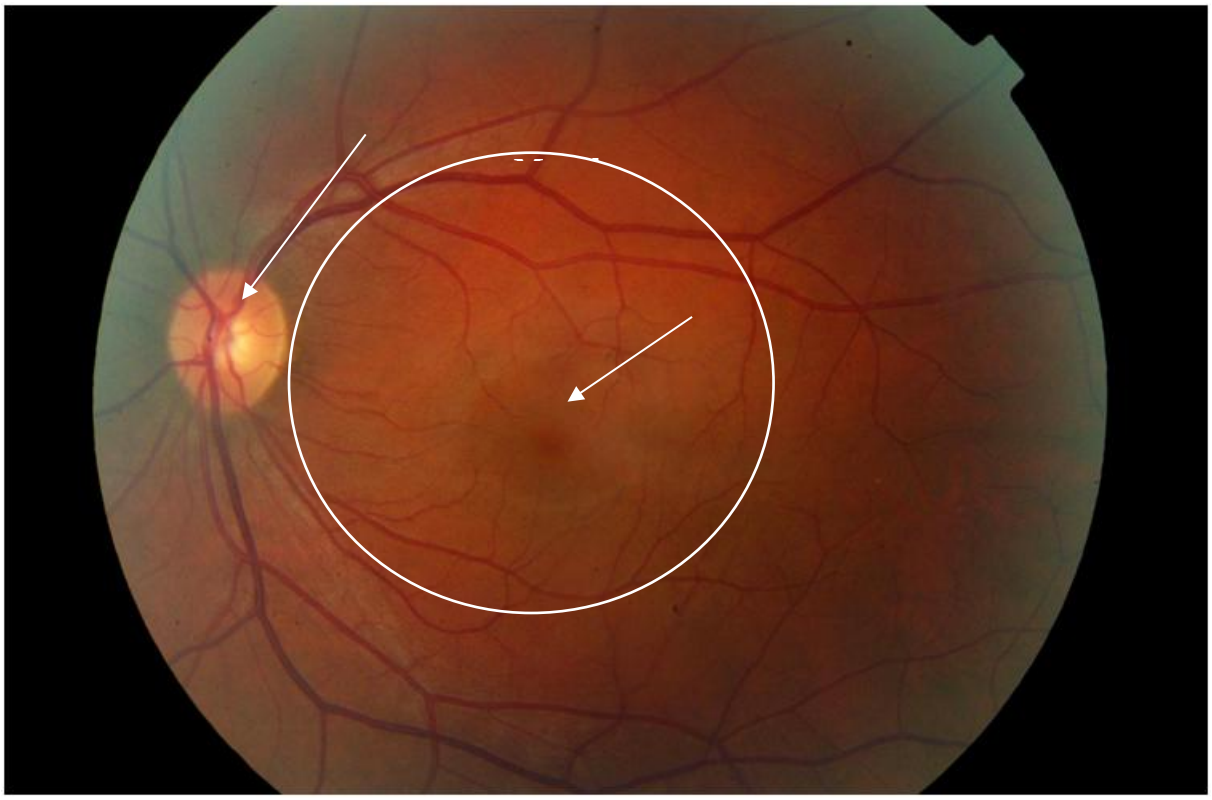


Figure 1. Retina, macula & fovea

2.2 Visual loss in diabetes

People with diabetes are at risk of visual loss from a number of conditions, some unrelated to diabetes, some not specific to diabetes but increased in diabetes, notably cataract, and some specific to diabetes, including proliferative retinopathy and macular oedema.

Two eye conditions are important in this appraisal – cataract and glaucoma, both being adverse effects of intravitreal steroids.

2.2.1 Cataract

Cataract occurs when the lens of the eye becomes opaque, preventing light from reaching the retina. The risk of cataract is increased in people with diabetes. Becker and colleagues [1] reported that the incidence amongst people with diabetes was about 50% higher than in the general population (12.4 (95% CI 12-12.7) compared to 7.9 (95% CI 7.6-8.2) per 1000 person years. However in diabetics with macular oedema, the incidence of cataract is about five times as high as in diabetics without DMO, and about 7.4 times the general population risk. A study from the Gloucestershire Eye Unit [2] reported that in people with diabetes, cataract was the commonest cause of visual impairment (49%), but cataract is easily treated by removal of the natural lens and replacement with an artificial lens. The causes of visual loss vary with age, with DMO accounting for 28% of visual impairment in people with diabetes in the 5th and 6th decades.

In the FAME trial, 86% of phakic eyes in the fluocinolone arm developed cataract, compared to 52% in the sham arm, and 41% in the fellow eyes not in the study. So the extra cataracts caused by fluocinolone were seen in 34% (86 – 52). Those in the fluocinolone arm had cataracts diagnosed and extracted on average 100 days earlier than those in the sham arm, with extraction at a mean of 18 months, and almost all extractions were performed by 24 months.

So most patients who were phakic at baseline developed cataract, but under half could be blamed on the fluocinolone. When considering the use of fluocinolone for chronic DMO in phakic eyes after all other treatments have failed, we need to consider possible outcomes;

- If we do not use fluocinolone, there is a high likelihood of central visual loss due to DMO.
- If we do use fluocinolone, an extra 34% will develop cataract, will suffer visual impairment as the cataract develops, but will have it removed, restoring vision.

So the fluocinolone trade-off is that in order to preserve central vision, many patients will have to have a period of deteriorating vision due to cataract, followed by its extraction. In cost-effectiveness analysis, there will be some temporary disutility and the cost of extraction.

It should be noted that some patients with DMO may also have peripheral visual loss due to proliferative retinopathy, but in most patients this will be treated with pan-retinal laser photocoagulation to preserve vision.

One issue is that if cataract surgery is done in someone with DMO, the DMO may get worse.[3] So some clinicians will treat the DMO first, before extracting the cataract. Others would treat the DMO with anti-VEGF drugs such as bevacizumab at the time of the cataract extraction. A meta-analysis by Feng et al[4] reported that trials of administering bevacizumab at the time of cataract surgery showed reductions in central macular thickness and improved VA compared to untreated control groups. Fluocinolone can be used in a similar way.

Another issue is what is meant by cataract. Some may define the threshold for cataract as 1+ nuclear sclerosis on the AREDS cataract grading system[5], others may prefer nuclear sclerosis 2+.

There are also different types of cataract: nuclear sclerotic, posterior-subcapsular and cortical. Nuclear sclerotic is the most common form, strongly age-related. The form most typically caused by steroids is the posterior-subcapsular, which may develop more quickly than other forms. The nuclear sclerotic form causes myopia which can be helped by spectacles, so may be less likely to require extraction.

2.2.2 Raised intra-ocular pressure and glaucoma

One adverse effect of steroids in the eye is an increase in pressure in the eye – intraocular pressure, IOP.

Glaucoma is characterised by increased pressure inside the eye, usually defined as IOP of 21 mm Hg or more with subsequent visual field defects and optic nerve damage. However some people may have signs of optic nerve damage and visual field loss at lower pressure. This is known as normal tension glaucoma. The intraocular pressure (IOP) rises because the normal drainage of aqueous fluid is impaired.

The increased pressure can cause progressive damage to the optic nerve, leading to impaired vision and blindness if not treated. Because of the way in which the nerve fibres are damaged, peripheral

vision is lost first, with central vision being affected later. There may be no symptoms in the early stages.

Glaucoma is treated by lowering the IOP. Treatment is initially by eye drops, sometimes using several different drugs, but these are not always sufficient and some people will require surgery. Patients in whom IOP rises and in whom surgery is required, will need frequent follow-up visits, perhaps 10 visits in the first year, reducing to 3 visits in year 2 and then 6-monthly. The NICE Clinical guideline on glaucoma[6] recommends that those at risk of glaucoma due to raised IOP are monitored at six monthly intervals, adjusted for their risk of developing glaucoma. However patients with DMO would be followed up regularly, so not all these visits would be additional.

The main types of glaucoma surgery are trabeculoplasty and trabeculectomy. (NB Not everyone regards trabeculoplasty as surgery.)

Trabeculectomy is the creation of a small hole in the eye to allow fluid to escape in order to reduce raised IOP, in those in whom it cannot be controlled by eye drops, even with two or three types of drops being used. 70% of patients can stop using drops after trabeculectomy. The operation takes about 60 minutes. Frequent visits are required afterwards, weekly at first, to monitor IOP, which can stay too high or fall too low.

Trabeculectomy increases the risk of cataract – about 10% of patients develop one by 3 years. Serious side-effects are rare, and include severe visual loss (about 1 in 1000). Other complications include endophthalmitis (infection in the eye) which can occur years later because the hole remains open, suprachoroidal haemorrhage, and cystoid macular oedema, usually transient. The RCO audit of trabeculectomy[7] reported early complications in 47% and late complications in 42%. However the most frequent late complication was cataract, which would apply less in the DMO population because so many have had cataracts removed or are going to.

If trabeculectomy fails, other forms of surgery such as tube drainage can be used. Trabecular stenting aims to reduce IOP by creating a bypass channel between the anterior chamber and Schlemm's canal to improve drainage of aqueous fluids. This procedure is often combined with cataract extraction and lens implant.[8]

Trabeculoplasty is a much simpler procedure where a laser is applied to the natural drainage system within the eye, leading to an increase in the outflow of aqueous humour. It is an outpatient procedure taking only minutes, though 2-3 hours of post-procedural monitoring is required. It may need to be repeated, perhaps a few years later.[9]

Deep sclerectomy is a less invasive surgical procedure used to reduce IOP. It has the advantage of fewer postoperative complications. [10]

For further details see Cochrane review by Burr et al. [11].

2.3 Diabetic eye disease

Diabetic retinopathy results from retinal changes arising from damage to small blood vessels and neural tissue due to high blood glucose levels over a period of years. About 90% of people with type 1 (insulin-dependent) diabetes will have some degree of retinopathy after having had the disease for 10 years. The prevalence is less in type 2 diabetes, but well over half will have some retinopathy by 10 years.

Diabetic macular oedema (DMO) is the most common cause of sight loss due to diabetes. [12] Oedema means fluid retention. The accumulation of fluid in the retina is due to increased leakage from blood vessels as well as incompetency of the retina to clear this fluid, which then builds up in the macula.[13]

Leakage can be from vascular abnormalities, mainly microaneurysms, or from macular ischaemia (areas of capillary non-perfusion) affecting the perifoveal capillaries or other capillaries at the macula. Ischaemic areas release vascular endothelial growth factor (VEGF), and this increases the permeability of the blood vessels leading to oedema.

DMO leads to a deterioration in detailed vision. If oedema persists, photoreceptor cell damage occurs with subsequent loss of vision.

Clinically significant macular oedema (CSMO) was defined in the ETDRS[14] as

- '*Thickening of the retina at or within 500um of the centre of the macula;*
- *Hard exudates at or within 500um of the centre of the macula if associated with thickening of adjacent retina;*
- *Zone(s) of retinal thickening one disc diameter or larger (1500um), any part of which is within one disc diameter of the centre of the macula*.'

Diabetic maculopathy can also be classified as focal or diffuse and ischaemic or non-ischaemic or mixed, depending on the location and cause of the leakage based on fluorescein angiography. Focal maculopathy is localised leakage at the macula. Diffuse maculopathy refers to generalised thickening of the caused by widespread leakage from dilated capillaries/microaneurysms or inefficiency of the

outer blood-retinal barrier (retinal pigment epithelium - RPE) or Muller cells to pump out fluid from the retina

Ischaemic maculopathy occurs when there is loss of blood vessels in the macula, which starves it of oxygen and nutrition, and is associated with a significant risk to vision.

Macular oedema can have also a tractional component from contraction of the innermost layers of the vitreous (posterior hyaloid) or contraction of epiretinal membranes (i.e. vitreomacular interface abnormalities). In the era of anti-VEGF treatment, diabetic macular oedema has been defined also as central involving, when there is fluid present within the central 1 mm as determined by optical coherence tomography (OCT) or non-central involving.

Macular oedema was classified by an EU Regulatory Framework Workshop based on the following features[15];

- Location – central, peri-central
- The amount of oedema as measured by macular thickness
- The presence or absence of vitreo-retinal interface abnormalities
- The presence or absence of hard exudates in the central subfield.

2.3.1 Prevalence of DMO

The prevalence of DMO increases with increasing duration of diabetes. In type 1 diabetes, the prevalence of CSMO is very low in the first few years after diagnosis, but rises to over 20% after 25 years, though that figure is based on data from the 1980s and 1990s, [16] and there is some evidence that better management of type 1 diabetes is reducing or postponing retinopathy. A global meta-analysis by the Meta-analysis for Eye disease (META-EYE) study group [17] concluded that prevalence of DMO under 10 years duration of diabetes was 3%; at 10-19 years, 13%; and after 20 years, 20%.

The risk of DMO is increased by smoking, by poor glycaemic control and by hypertension. It may be precipitated by pioglitazone which can cause oedema. [18]

In the META-EYE global review the prevalence amongst people with normal blood pressure was 5.5% compared to 10.6% in those with hypertension (BP >140/90 or already on anti-hypertensive medications).

There was a strong link between poor glycaemic control and prevalence (see Table 2).

Table 2. Diabetic control and DMO

HbA1c	Prevalence of DMO
7.0% or less	3.6%
7.1 to 8.0%	6.3%
8.1 to 9.0%	7.7%
Over 9.0%	12.5%

In the DCCT/EDIC study the group that received intensive treatment during the trial phase, still had less retinopathy 10 years after the trial ended.[19]

Hence good control of blood glucose and blood pressure should reduce the number of people developing DMO, and improving control may lead to regression of DMO.

A study by Minassian and colleagues[20] estimated the prevalence of DMO in England. Based on the estimate in the NICE scope of 3.3 million people in England with diabetes, and an expected prevalence of DMO of 7.12% in one or both eyes, we can estimate that there would be almost 86,000 people with clinically significant DMO with varying degrees of visual impairment. The Minassian estimates were based on the excellent data from the Diabetic Retinopathy Screening Service for Wales, based on many years of screening and a dataset from 27,178 screened people with diabetes. (A caveat is required. The Minassian data may not reflect the advent of highly sensitive OCT which may detect DMO that is not clinically significant, so the recorded prevalence now may be greater. Studies may report DMO differently, either as prevalence of clinically significant DMO, or as any DMO detected on OCT.)

The richness of the data meant that Minassian et al were able to subdivide people with DMO into groups of varying severity, as shown in Table 3. Percentages rounded to one decimal place.

Table 3. Estimated number of people with varying severity of DMO in England

	Description	**Prevalence % of all people with diabetes**	**Expected number in England**
All DMO	DMO of any severity in one or both eyes	7.1%	234,300
	In one eye only	4.7%	
	In both eyes	2.3%	
Slight sight loss	Sight loss (BCVA <6/6) attributable to DMO in at least one eye	2.8%	92,400
Visual impairment	BCVA < 6/6 to >6/60 attributable to DMO in at least one eye	2.6%	85,800
Blindness	BCVA ≤ 6/60 in both eyes, one or both due to DMO	0.1%	3300
Partial sight	BCVA <6/19 to >6/60 in the better seeing eye, due to DMO	0.2%	6600

	in one or both eyes		

It has been suggested that as many as a third of cases of DMO may resolve after six months [21] even if untreated, but that most becomes chronic. We take the definition of "chronic" to be lasting more than six months. However the comment about a third resolving spontaneously does not fit with the observation from the RISE and RIDE studies that 74% of pts in the sham arms, crossed over to ranibizumab after 24 months, though it should also be noted that at 24 months, 28% of the RISE and RIDE sham group had not received rescue laser.[22].

2.4 Treatment of DMO

The mainstays of treatment have been laser photocoagulation and anti-VEGF drugs. The NICE guidance on the anti-VEGF drugs ranibizumab [23] and aflibercept (TA 346)[24] recommended their use in patients with DMO and a central retinal thickness of 400 microns or more. NICE has not issued guidance on another anti-VEGF drug, bevacizumab, which is as effective and much less expensive, but not licensed for use in the eye. However a High Court ruling has said it is not unlawful for clinical commissioning groups to use bevacizumab instead of ranibizumab or aflibercept in wet AMD.

In TA349 in July 2015, NICE recommended the corticosteroid dexamethasone (the implant Ozurdex) for treatment of DMO only in pseudophakic patients who had had no response to non-steroid treatments, or in whom such treatments were unsuitable.[25] The definition of "no response" is not given, but there is a comment in the text that it would be inappropriate to define response as a gain of 5 or more letters, because DMO was a progressive condition, so avoiding loss of letters was regarded as a benefit. So intravitreal steroids could be regarded currently as being the treatment of last resort when all else has failed.

The updated Cochrane review regards a gain of fewer than 5 letters or less than 0.1 logMAR units as lack of response.[26] Most trials have used the proportion of patients gaining 10 or more, or 15 or more letters as the primary outcome, including trials of anti-VEGF drugs such as RISE and RIDE[22] and the FAME trial.[27]

2.5 The previous appraisal of fluocinolone

The NICE guidance from the last appraisal, para 4.21 states;

"*On balance, the Committee concluded that fluocinolone acetonide intravitreal implant could be a cost-effective use of NHS resources and recommended it as an option for people with chronic diabetic macular oedema that is insufficiently responsive to available therapies and if the implant is to be used in an eye with an intraocular (pseudophakic) lens."*

Note that unlike in the dexamethasone guidance, there is no mention of patients "in whom such treatments were unsuitable". Anti-VEGF drug treatment requires multiple injections in the first year with monthly visits for injections or monitoring, with perhaps four injections in years 2 and 3, and fewer in subsequent years, with some patients still requiring some in years 4 and 5. That will be a considerable burden for some people. One consequence of DMO is that patients may be unable to drive. The DVLA minimum requirement for central vision to drive in the UK is 6/12 vision. (Peripheral vision is also important so even if people meet 6/12 central vision, they may not meet driving criteria.) However visits will also be necessary after steroid injections because of the need to monitor intra-ocular pressure.

We are aware that in some centres, fluocinolone may be given simultaneously with cataract extraction – the cataract is removed, an artificial lens inserted, and the fluocinolone implant inserted into the now pseudophakic eye.

2.6 Issues

The scope says:

"People with chronic diabetic macular oedema that is insufficiently responsive to available therapies who have phakic lenses".

The scope does not define "chronic", "insufficiently responsive" or "available therapies". Response can be functional (vision) or anatomic (reduction in retinal thickness on OCT), but the latter may or may not be accompanied by change in vision. The anti-VEGF drugs act by removing fluid from the retina, but vision may or may not improve. Vision may take time to deteriorate even if oedema is present, so a lack of deterioration may not necessarily indicate a good response if the fluid has not cleared.

The Alimera submission (page 20) gives definitions of insufficiently responsive as worsening or static foveal thickness on OCT, or "partially dry DMO" in patients with FTH (central foveal thickness)

initially ≥400 micros, and in patients with FTH <400 microns “any thickening or decrease in visual acuity that in the opinion of the clinician, is unlikely to respond to laser”.

The DRCN.net group[28] defined non-response to ranibizumab treatment as <20% reduction in central retinal thickness after repeated injections over 12 months. About 20% of patients fell into this category.

Responsiveness can be defined in terms of visual acuity, expressed as number of letters that can be read on a standard chart, or as central retinal thickness in microns. One problem with using visual acuity, is that it may be good (6/9 or even 6/6, which can occur in people with DMO) at time of diagnosis, which means that letter gain may not be feasible. Some studies recruit only people with visual impairment. In patients with good baseline VA, response could be reported as reduction in retinal thickness. This was not a problem in FAME where recruits had BCVA between 19 and 68 letters.

There are four issues with reporting changes in retinal thickness;

- What part of the retina? Some studies report foveal thickness, others central retinal thickness (referring to the central 1mm thickness)
- Which device is used? Different OCT devices can give different readings in the same patient. This does not matter if the same device is used throughout the patient’s diagnosis and follow-up
- Test/re-test variability.
- Variability of central retinal thickness, which is not constant, but varies by time of day. One study reported differences before and after a meal. So we need to define a minimum change that allows for variations over time
- We also need to define what a clinically important difference in CRT is.

There are also issues in measuring VA. We know that spontaneous improvement can occur. But there can also be test/re-test issues with measurement of visual acuity. If patients are feeling downcast, they may read less. If the clinic is busy and the optometrists or nurses are pushed for time, patients may not be given enough time to read the visual acuity chart, and vision may appear to be worse than it is. Conversely, encouragement by the optometrist, may persuade them to attempt one more line, as might a feeling that treatment should be helping. So we suggest using a 5-6 letter testing variability, and using a 10-letter change as indicating a clinically significant improvement in vision.

We need to decide how to define anatomic response. How many microns change in thickness? Or by absence of fluid, since there may be diabetic neurodegeneration leading to a thinner retina, which may

mean that the central retinal thickness may not be too thick despite the presence of oedema. And at how many months, or after how many anti-VEGF injections?

We will assume that;

- Chronic is defined as present for more than six months since first detected without clearance during that time. However, in this STA we are referring more to the time since starting treatment with anti-VEGFs or laser. We also need to distinguish between persistent oedema and recurrent.
- Insufficiently responsive means a gain of fewer than 10 letters, or any loss of letters
- Available therapies means laser and anti-VEGF drugs. Dexamethasone is not included because it is not currently approved for phakic patients.

The anti-VEGF drugs have been a major advance in DMO, and are regarded as first line treatment in the EURETINA (European Society of Retina Specialists) Guidelines. [29] (In contrast to NICE guidance, they say that laser is no longer recommended, but they do not consider cost-effectiveness, despite commenting on the high costs of ranibizumab and aflibercept.)

However, after anti-VEGF treatment, only about half of the patients get a gain of 10 or more letters, as shown in the RESTORE trial[30] and a small proportion lose more than 10 letters. Some patients respond very well, some show little response, and some respond partially. The updated Cochrane review[26] notes that 30% of patients gained 15 or more letters on anti-VEGF drugs, compared to only 10% after laser treatment, but expressed concern that results in routine care would not be as good as in the trials. This concern is justified. The UK audit report by Egan and colleagues[31] on results with ranibizumab for DMO reported that 17% of eyes gained 15 or more letters, 60% were “stable”, meaning 0-15 letters gained, but that 23% lost letters. The mean letter gain was only 5 letters. The reduced effectiveness in routine care may simply reflect that the resources available in the NHS may not match those in the trials, for example for monthly injections/reviews. Patients may be being seen only every 6-8 weeks because of pressure in the NHS. In addition, patients in the trials seem to have had better diabetic control than seen in routine clinics. In FAME, HbA1c at baseline was 7.9%. In the current DIAMONDS trial in macular oedema, the average is about 9%. (Unpublished data.) In an ideal world, patients might even be treated at 3-weekly intervals initially though this is contrary to the licence.

In RISE and RIDE[22] 30-40% of patients did not gain 10 or more letters after 3 years treatment with ranibizumab. At the 3-year point in the DRCRN study by Elman et al[32], about half the patients on ranibizumab had failed to gain 10 or more letters and about 5% lost 15 or more letters.

A key question in cost analysis is how phakic people who are insufficiently responsive are treated if fluocinolone is not available. Will anti-VEGF injections be continued? Or laser, if thickness is <400 microns? Anti-VEGF treatment may not provide a satisfactory response, but if they do not fully dry the oedema but reduce the thickness to < 400 microns, then laser could be tried.

Another issue is how we define failure on anti-VEGF treatment. There are various studies suggesting that if ranibizumab is insufficiently effective, it is worth trying aflibercept. The effectiveness of switching has been reviewed by Banaee and colleagues.[33] The rationale for trying aflibercept if ranibizumab or bevacizumab are ineffective, is that ranibizumab binds VEGF-A, whereas aflibercept binds VEGFs A and B, and PlGF (placental growth factor, which acts in combination with VEGF-A), and so neutralises a larger number of the cytokines that may be involved in the development of retinopathy. Aflibercept also has a longer intra-ocular half-life. Banaee and colleagues report that 8 studies of switching from ranibizumab to aflibercept all showed improvements in central macular thickness, and five showed improvements in vision. They found no studies of switching from aflibercept to ranibizumab.

However, Ferris et al[34] have suggested that the improvement in these before and after studies could be due to regression to the mean. In a study using data from AMD studies, they studied patients that met the criteria for switching drugs, but had not switched, and found that they improved VA by 3-5 letters in the three months after the switching rule was met.

There is one very small study[35] of SGLT2 inhibitors in DMO showing a before and after improvement. With no controls, in a condition where spontaneous improvement can occur, little weight can be given to this study. However, if we extrapolate from the reduction in heart failure with the SGLT2 inhibitors, due to their diuretic effect, then an effect on DMO may be plausible. A proper RCT is required.

3 CLINICAL EFFECTIVENESS

Alimera submitted evidence from the FAME (Fluocinolone Acetonide in Diabetic Macular Edema, NCT00344968) trial, and from some observational studies.

3.1 The FAME trial

FAME[27] was carried out as two identical trials, but analysed as one, and was done in 101 sites in North America, Europe and India, with FAME A being done in northern sites and FAME B in southern sites, in much the same countries. Several UK centres were involved - Bristol, Southampton and Wolverhampton. The FAME trial compared fluocinolone with sham injections. The Iluvien device releases fluocinolone into the eye very slowly, over 3 years, after an initial burst.[36]

FAME was a good quality trial. The quality assessment is reported in Appendix 1.

Details of the baseline characteristics, results and adverse effects of the phakic only group are presented in Appendix 2. In brief, there were 97 patients in the phakic at baseline, pseudophakic at 3 years group. Mean age was 61, 88% had type 2 diabetes with mean duration 16 years, and 9% had type 1 diabetes with mean duration 31 years. Over two-thirds were from North America, 14% were from Europe, and 14% from India. 74% had some degree of cataract at baseline, but none severe. I conversation with Alimera, it appears that diagnosis of baseline cataract appears to have been highly sensitive, based on photographic detection of any degree of opacity. Cataract serious enough to impair visualisation of the retina led to exclusion from the FAME trial.

In the group that started phakic and stayed so, 94% had cataract at baseline.

Mean HbA1c was 7.9% (the NICE target is 6.5%) and only 31% had HbA1c under 7.0%. The mean centre point thickness in chronic DMO was 461.8 (SD 153.5, implying a skew towards higher levels because the minimum for inclusion was > 250 microns) making laser less likely to be effective – the NICE guidance on ranibizumab and aflibercept uses a central retinal thickness threshold of 400 microns for preferring drugs to laser photocoagulation.

The results of the FAME trial varied by duration of DMO. The overall results failed to show a significant difference between the fluocinolone and sham arms, but there was a significant difference in those with longer duration of DMO. This appraisal is concerned only with chronic DMO. The FDA

noted that the analysis by duration was not pre-specified in the protocol or the statistical analysis plan, [37] but Alimera did inform the FDA that though not mentioned in these documents, the duration of DMO analysis had been pre-planned. The division was initially described as being at 3 years but was later revised to 1.7 years. Details of the method are given in Appendix 3, for reference. The median duration in the >3 years group was 5.2 years.

The main problem with data from the FAME trial is that the study was done mainly in the era before widespread use of the anti-VEGF drugs. So patients had been treated with laser only, and therefore do not match the whole population in the NICE scope. Ideally, we would have a trial treating DMO with laser or anti-VEGFs, then if response was insufficient, randomising people to fluocinolone or to continuing with laser or anti-VEGF, though that design could be criticised for continuing ineffective treatments.

In the FAME trial, two doses were used, aiming at 0.2 µg and 0.5 µg daily, being referred to as low dose and high dose inserts. There was little or no difference in VA gains, but the high dose was associated with more cataracts and more glaucoma. Surgery to relieve IOP was required in 8.1% of the high dose group, 3.7% of the low dose group and 0.5% of the sham group.

So the 0.5 µg/day dose has been consigned to history, and no details of results with that are included in this ERG report.

Cunha-Vaz 2014 [38]

This paper[38] (Table B2.1 and pages 68 to 71 in the Alimera submission) reports the FAME analysis by duration of DMO, divided into chronic and non-chronic. The results for the phakic at baseline group are not presented separately apart from cataract frequency. In the non-chronic group, the proportions gaining 15 or more letters were 22% in the sham group and 28% in the fluocinolone group. The non-chronic sham group received rescue laser more often than the fluocinolone group (63% versus 43%) and more had anti-VEGFs (14% vs 3%) and triamcinolone (15% vs 7%).

In the chronic group, only 13% of the sham group gained 15 or more letters, compared to 34% of the fluocinolone group. However the 13% shows that some recovery occurs. More of the sham chronic group received rescue laser (62% versus 41%), anti-VEGFs (15% vs 3%) and triamcinolone (24% versus 8%). Some recovery may be due to improved glycaemic control – the diagnosis of DMO may be an incentive to improve control.

The mean improvement in BCVA was 7.6 letters on fluocinolone and 1.8 letters in the sham group, a difference of 5.8 letters, which is less than seen in the anti-VEGF trials. In the RISE and RIDE trials of ranibizumab, the placebo groups gained 2.5 letters by 24 months, and the ranibizumab groups gained 12.2 letters, a difference of 9.7. However all patients in the FAME trial had had laser treatment (and failed to reach 250 microns or less), whereas only 66% in the RISE and RIDE trials had had previous laser treatment, and about 30% had had no previous treatment of any kind. Duration since diagnosis of DMO was only 2.3 years in RISE and RIDE. So it could be argued that the FAME patients with chronic DMO were more advanced and less likely to respond to treatment. There have been no trials of anti-VEGFs versus fluocinolone, so the relative potencies are unknown.

Central retinal thickness appears to have been measured by OCT every three months in the 3-year FAME duration. The difference in thickness at 36 months was only 38 microns, which was not statistically significant ($p = 0.036$).

As noted earlier, 86% of those phakic at baseline developed cataracts, but so did 52% in the sham eyes (and 41% in the untreated fellow eyes). So it could be argued that only 34% of the 86% cataracts were due to fluocinolone.

The mean time to cataract sufficient to be reported as an adverse event in phakic subjects was 15 months in the fluocinolone group, compared to 22 months in the sham group. Similarly, the mean time to cataract extraction was shorter in the fluocinolone group, at 18 months versus 27 months in the sham group. (Figures rounded to whole numbers.) The reasons are not clear, but one possibility is that the type of cataract induced by fluocinolone is posterior-subcapsular which causes symptoms at an early stage. Almost all those who developed cataract in the fluocinolone group had extractions, whereas only 70% of sham group cataracts were extracted. So 85.1% of the phakic fluocinolone group had a cataract operation in the study eye, versus 36.4% in the sham group. The FAME study did not collect data on type of cataract.

Table 3 of the Cunha-Vaz paper shows that 41% of the fluocinolone group received rescue laser treatment, compared to 62% of the sham group. Criteria for rescue are given in the Campochiaro 2011 FAME paper[27] as persistent or recurrent oedema.

This raises an important point. If treatment successfully dries out the macula (i.e. removes all the oedema) the retina starts functioning again and that will allow recovery of function. It may not last – recurrence happens after both successful anti-VEGF and steroid treatment. However the outlook is

better than after no response (no reduction in oedema) or partial response (reduced but persistent oedema). If the retina is never dry, deterioration in function is to be expected.

Yang 2015 [39]

This paper from the FAME trial is the most relevant to this appraisal. Yit Yang (from Wolverhampton) and colleagues compared outcomes for the initially pseudophakic subgroup in the FAME study, and the initially phakic but who developed cataract and became pseudophakic after treatment with fluocinolone.

Figure 2 and Figure 3 provided by Alimera at clarification stage (figure B2.3 in the original submission includes non-chronic patients) shows the rapid improvement by about 6 letters in the group that had cataracts removed after fluocinolone implant (CAI group – line with squares – the line with triangles is not relevant), followed by a decline in vision as cataract develops, then an improvement after cataract removal, ending up with about 11 letter gain. However the sham group (data from Cunha-Vaz not Yang) improved by two letters (data not provided in Yang or by Alimera). So the mean additional benefit of fluocinolone was 9 letters. 42% of patients phakic at baseline, with chronic DMO, gained 15 letters BCVA by month 36.

Figure 2. Mean change in BCVA letter score after fluocinolone implant

Figure 3 shows the proportion with an improvement of 15 or more letters from baseline (line with black squares is the phakic at baseline group – please ignore the line with triangles).

Figure 3. Proportion of patients experiencing a ≥15-letter improvement in BCVA

The FAME trial recruited patients who had foveal thickness of 250 microns or more despite at least one macular laser treatment and BCVA in the range 19 to 68 ETDRS letters (Snellen 20/50 to 20/400). It was carried out before the anti-VEGF drugs were commonly used. For the purposes of this STA, we ideally need a trial in which eyes that have not responded to laser or anti-VEGFs (or a combination of initial anti-VEGFs then laser once the retinal thickness has been reduced) are randomised to fluocinolone or to continuing with anti-VEGFs or laser.

Adverse effects.

Cataract was the commonest adverse effect. The other main hazard was raised intra-ocular pressure. This was reported in 34% of the phakic-pseudophakic fluocinolone group and in 14% of the sham group, though 39% of the fluocinolone were reported to be using IOP-lowering medications. Amongst the 97 patients in this group, one required a trabeculoplasty and four had trabeculectomies, compared to none in the sham group.

The table of adverse effects in the Alimera submission (Table B2.22) is based on the Cunha-Vaz paper[38] which has all chronic DMO patients, not just the phakic ones. The columns for non-chronic DMO are not relevant. Alimera provided data for only those phakic at baseline – see appendix 2.

In the fluocinolone arm in the chronic group, 5.3% of patients needed surgical interventions for raised IOP (compared to none of the control arm), and 85% had cataracts extracted (compared to 51.1% of the control eyes). IOP-lowering drugs were needed by 39% of fluocinolone patients and 15% of control patients.

Very rarely, the fluocinolone insert has to be removed or re-positioned. The few reports of this cite two reasons. One is difficult to manage glaucoma. The other is migration of the implant into the front of the eye with a risk of corneal damage, which does not happen in phakic patients so is not relevant to this appraisal.

Re-treatment with fluocinolone within 3 years.

The FAME protocol allowed for more than one insert of fluocinolone to be used, at the discretion of local investigators, but no earlier than 12 months. Re-treatment was to be based on criteria as follow:

"Subjects were eligible for retreatment with the masked study medication to which they were randomised any time after the Month 12 assessments if they experienced vision loss or retinal thickening per optical coherence tomography (OCT). As a result, subjects received various numbers of treatments (1−4) during the studies. The protocol was designed to allow flexible timing of retreatment because the duration of therapeutic effect was not known prior to the start of the study."

The licence follows the trial protocol in allowing a second fluocinolone insert after 12 months. In the FAME trial, chronic group, 24% of patients received more than one fluocinolone implant. [38] However in the modelling, Alimera do not expect treatment at less than 3-year intervals.

Many patients with DMO will have diabetic retinopathy, and there is evidence from FAME[40] that a by-product of treating DMO, is a slowing of progression of retinopathy. This has not been included in the modelling – it would increase the already considerable complexity, but can be noted.

One uncertainty is the proportion of patients who will not need further treatment with fluocinolone after the first implant runs out after 3 years. The oedema may be cleared and vision improved, and the improvement may remain. However in a proportion (since we do not have data for, say, 10 year follow up, we do not know how many) the fluid is expected to recur at some time. Most recurrences may happen around the 36 month point when the fluocinolone runs out, but vision would not always go down right away.

3.2 Other studies submitted by Alimera

Given the lack of trial data for the group of eyes poorly responsive to anti-VEGF treatment, Alimera provided data from a number of observational studies.

The Medisoft Group[41]

Bailey and colleagues from the Medisoft Audit Group[41], from 14 UK centres, report results of fluocinolone implants in 305 patients (345 eyes), of whom 85% of eyes (79% of patients) had had previous anti-VEGF drug treatment, and 28% had had prior laser. Recruitment was up to August 2016. The mean central subfield thickness was 451 microns before fluocinolone, showing that previous treatment had failed. Most were either pseudophakic at baseline (90%) or received fluocinolone at the same time as cataract surgery (7%). So there were only about 10 phakic eyes. The value of this study for our purposes is that it provides data on the effectiveness of fluocinolone in patients previously treated with anti-VEGFs, which the FAME trial cannot. It also provides data on a wider range of patients than was in FAME, by including patients with baseline VA ranging from 5 to 85 ETDRS letters. (FAME 19 to 68). Statistical and writing support was provided by Alimera.

The duration of DMO prior to fluocinolone treatment is not reported but patients had had a mean of 7.4 intra-vitreal injections before fluocinolone so chronicity can be assumed. Though previous treatment with anti-VEGFs was in routine NHS care, and whether injections were given at optimal intervals (to ensure that patients were indeed "anti-VEGF non-responders") is not reported.

Baseline retinal thickness was reduced but only to a mean of 356 microns at 24 months. (For comparison, FAME baseline was 461 microns, reducing to about 300 at 24 months, from graph[27]) The mean number of letters read rose by 5, from 52 at baseline to 57 at 24 months. (FAME 7 letters increase at 24 months in the pseudophakic group [27]). The proportions improved or stable (defined as any gain, or less than four letters loss from baseline) were 79% at 12 months, 82% at 18 months and 87% at 24 months. The proportion with 6/12 or better vision rose from 18% at base line to 39% at 18 months and 40% at 24 months. 21% achieved a gain of 15 or more letters (29% in FAME). The proportion able to drive (minimum VA 6/12) rose from 18% before fluocinolone to 40% after.

In summary, overall 88% of patients had stable or improved vision at 24 months. The Medisoft patients do not match the phakic patients who are the subject of this appraisal, but the results do show that fluocinolone works in patients in whom anti-VEGFs have failed.

After fluocinolone insertion, 64% of eyes received no additional treatments. The other 36% had additional treatments, mainly ranibizumab and aflibercept.

When should anti-VEGF treatment be regarded as ineffective?

One issue is how many anti-VEGF injections should be given before deciding on ineffectiveness. Another issue is whether interval between injections were optimised – NHS pressures may sometimes make that difficult. The Alimera submission includes a treatment pathway from West Essex CCG (page 19)[42] recommending that anti-VEGF treatment be discontinued after three injections if there

was no response (but response is not defined), but unpublished data from the RESTORE and RESOLVE trials showed that some people respond more slowly, as noted in the ERG report for NICE for TA274 (Aberdeen HTA Group 2011) and by Bottoni and colleagues. [43] So an arbitrary cut-off after three injections might prevent benefit in slow responders. The West Essex pathway recommends discontinuation if there is "no response". As noted earlier, response could be anatomical/structural or visual, and so could be expressed as reduction thickness (number of microns), or resolution of fluid as seen on OCT, or number of letters gained.

In the initial DRCR Network trials, the loading dose was considered to be four injections. It is not clear if anyone has done cost-effectiveness analysis of the different stopping rules.

The mean increase of 5 letters in the Medisoft study would not be regarded as a significant response by bodies such as the EMA which regards the minimal clinically relevant improvement to be around 10 letters. [15] This figure has also been used in assessing new treatments for retinal diseases in NICE appraisals in recent years.

The issue of whether patients not responding sufficiently to ranibizumab or bevacizumab should be tried on aflibercept has been discussed earlier. Only 1.7% of the Medisoft eyes had been previously treated with aflibercept. Most (68%) received ranibizumab, with 21% receiving bevacizumab.

Raman 2018

In a hypothetical modelling exercise, Raman[44] from Plymouth has compared the costs of continuing aflibercept monotherapy versus a switch from aflibercept to fluocinolone after five injections of the former, in phakic patients with a sub-optimal response (not specified). The current NICE guidance restricts steroids to pseudophakic patients. The study was sponsored, and the model development funded, by Alimera. Costs included cataract extraction for almost half of the patients on fluocinolone. Over a 3-year timescale, costs were lower when fluocinolone was used. However the costs included continuing aflibercept despite lack of response – no stopping rule was applied. The study is only of costs, and does not take into account the clinical effectiveness gains from using fluocinolone in people with a poor response to aflibercept. If visual outcomes at 3 years were better, and costs lower, the fluocinolone combination strategy would dominate. However a key issue is continuing the costs of ineffective aflibercept.

Quhill and Beiderbeck 2017

Another cost analysis[45] (supported and co-authored by Alimera), using data from a Sheffield hospital, compares the costs of ranibizumab (14 injections over 3 years) and fluocinolone (one implant) when

used from the start of treatment (which would be contrary to the NICE scope and the licence, which assume fluocinolone is only used when other treatments fail). Quhill and Beiderbeck provide a useful split by phakic and non-phakic, including the costs of cataract extraction and glaucoma management. Using fluocinolone is much less expensive over a 3-year timescale than ranibizumab, because of the high cost of 14 ranibizumab injections, even allowing for the cost of cataract extraction in patients phakic at baseline, and of treatment for raised IOP. The authors appear to assume equal clinical effectiveness of the ranibizumab and fluocinolone treatments, but there is no trial comparing the two. The study provides very useful cost data, but its usefulness for this appraisal is limited by the fact of fluocinolone being used from the start of treatment. The cost analysis is more relevant to the use of fluocinolone as first-line drug.

Carneiro et al – ICEPT - a study of the use of fluocinolone in Portugal

Only a conference abstract (AAO 2018) [46] of this study is available. Patients were monitored for 12 months before and 12 months after fluocinolone injection. Baseline VA ranged from 60 to 70 ETRDS letters. 44% of 77 patients/113 eyes (abstract – Alimera submission says 93 eyes in 68 patients) were phakic, but some underwent cataract extraction at the same time as, or shortly after fluocinolone implantation. Details are sparse. The abstract describes the patients as having had an inadequate response to prior therapies but gives few details of prior anti-VEGF treatment. Most received shorter-acting steroids (not specified) before fluocinolone. 20% received anti-VEGF injections in the year before fluocinolone insertion compared to 10% afterwards. Does this imply that some were still responsive to anti-VEGF treatment? Ideally they would have been given fluocinolone and then been randomised to have additional other treatments.

In the year before fluocinolone, there was a decline in mean VA from 60 to 58.5 letters, but in the 12 months afterwards, an increase of 9 letters. Follow-up (so far) is only for 12 months, and no details are given of cataract development, presumably because longer observation is required, and because an unspecified proportion of the baseline phakics had immediate cataract extraction. Little useful data can be gleaned from this study at present.

The ICE-UK study.

One paper from the Iluvien Clinical Evidence study in the UK (ICE-UK) is included in the Alimera submission (Page 28). This is the clinical effectiveness article by Holden et al.[47] Most patients (82%) recruited to their study had had anti-VEGF drugs in the year before, but their visual acuity had declined, so they match the population in the NICE scope. 63% had had prior laser. Recruitment was from April 2013 to April 2015. The duration of DMO was not recorded in ICE-UK, so the proportion with chronic DMO is not known, but can be assessed from the number of previous treatments. In

Table 1 of Holden 2017 [47], median time from first laser treatment is reported to be 3.8 years (IQR 2.1 to 6.1 years) in the 63% who had had laser, and 1.2 years (IQR 0.8 to 2.5 years) in the 82% that had had prior anti-VEGF treatment. About two thirds of those having previous anti-VEGF treatment had had more than six prior treatments for DMO. So it seems reasonable to assume chronicity of DMO. Most (70%) previous anti-VEGF treatment was with ranibizumab, with bevacizumab in the rest. Aflibercept was used in only one patient.

Of the 233 eyes, 89% were pseudophakic at baseline, and others had cataract extraction shortly after fluocinolone insertion. So few match the phakic group in this appraisal. Most (82%) had had anti-VEGF treatment (70% ranibizumab, 32% bevacizumab) with a median of 5 injections. 43% had had prior steroid injections, mainly triamcinolone. The study was designed by Alimera. The study is a before and after one, comparing costs in the 12 months before fluocinolone with costs in the 12 months after fluocinolone. Because costs are only collected for 12 months after, some later costs of adverse effects such as cataract and glaucoma are not included. Only two patients had glaucoma surgery. 15% started IOP-lowering treatments after the fluocinolone was inserted.

Interestingly, fellow eyes had also often been treated: 55% with laser, 47% with anti-VEGF drugs, and 19% with steroids (mainly triamcinolone).

The cost comparison includes savings from reductions in anti-VEGF injections, but it could be argued that those should be stopped when they became ineffective. However after implantation of fluocinolone, 32% received anti-VEGF injections, and a few received laser (10%) or non-fluocinolone steroids (7%), implying that response to fluocinolone alone was insufficient.

VA improved by 5 or more letters in 44% of patients, by 10 or more in 30% and by 15 or more in 18%. Some of this would be due to natural recovery, improved glycaemic control or the additional treatments.

Holden was also first author of a paper written with colleagues from Alimera[48] that compared results from FAME with results from 13 UK centres taking part in ICE-UK. Because the ICE-UK patients were assumed to have chronic DMO, their results were compared with the chronic subgroup from FAME. A few other adjustments were made to improve matching. The ICE-UK patients had poorer baseline VA and foveal thickness. Holden and colleagues assume that most ICE-UK patients would have had retinal thickness >400 mcm at some time, and the baseline mean foveal thicknesses was 482. A much higher proportion of the ICE-UK group were pseudophakic compared to FAME patients (89% vs 37%), as would be expected in view of the NICE guidance and its effect on UK practice.

After 12 months, the ICE-UK group had gained 3.8 letters and the FAME chronic group had gained 5.0 letters. There is no split by phakic and pseudophakic. In the 12 month follow-up periods, a higher proportion of ICE-UK patients (33%) received additional treatment with anti-VEGF drugs than the FAME patients (a few % estimated from graph[27]). Holden et al attribute the poorer results in the ICE-UK study, compare to those in the FAME trial, to the more advanced DMO at baseline.

Given the small number of phakic patients in ICE-UK, the main value of these papers for our purposes is to provide further evidence of the efficacy of fluocinolone in patients previously treated with anti-VEGFs.

The IRISS study

The Iluvien Registry Safety Study (IRISS, NCT01998412)[49] has collected data on 593 eyes treated with intravitreal fluocinolone in 31 centres in UK, 11 in Germany and 5 in Portugal. All centres contribute data to a central registry. It includes pseudophakic and phakic. The proportions are 83% pseudophakic (compared to 36% in FAME) and 16% phakic (FAME 64%). All patients were considered to have chronic DMO insufficiently responsive to previous treatment. The focus of the study was on adverse effects, especially raised IOP.

The study is part of the response to the EMA requirement for post-regulatory surveillance.

Intra-ocular pressure did not vary between the phakic or pseudophakic subgroup. Rises in IOP were small (2.2 mmHg for the phakic subgroup) but 23% of patients did require drugs to lower IOP, though only 14 of the 593 eyes required surgery for glaucoma. (No details regarding proportion phakic.)

The results seen in IRISS were similar to those in FAME, with slightly greater gains in VA in the phakic at baseline group, but not statistically significant difference. Mean VA gains were modest, 3.7 letters at 12 months, but data on the spread of results such as proportion gaining more than 10 letters, are not provided.

Interestingly, by 24 months, 69% of patients had not required any additional treatment. Those who did received laser (10%), anti-VEGF (22%), or other steroids (7%).

Other studies

Two studies cited by Alimera are much too small to be of value. The RESPOND study [50] had only 12 patients, of which only four were phakic. The study by Massin et al [51] had 16 patients but only 5 phakic eyes.

The PALADIN study[52], is not yet relevant. It is only available as a conference abstract, is still recruiting, and only data to six months on about a sixth of the intended recruitment has been presented. The abstract does not say how many phakic patients have or will be recruited. Patients in PALADIN have previously been treated with intravitreal steroids (not specified) without developing raised IOP

The USER study by Eaton et al[53] is a retrospective study from four centres in the USA, reporting the effectiveness of fluocinolone in patients that have had previous treatment with anti-VEGFs (77%), other steroids (56%) or laser (50%). It reports a reduced need for other treatments after fluocinolone insertion but only 23% of the recruits were phakic at baseline, and their results are not separately reported.

The Retro-IDEAL study by Augustin and colleagues[54] is published only as a conference abstract. Most (75%) patients were pseudophakic and 50% had had vitrectomy before the insertion of fluocinolone. Most patients had been treated with ranibizumab but had sub-optimal responses (details not given). Results for the phakic subgroup are not given separately but with only about 12 (abstract) or 20 (submission) phakic eyes, cannot contribute much to this appraisal. (The abstract says 70 of 82 eyes were pseudophakic, so only 12 were phakic. Submission says 20 phakic.) According to the submission (Table B2.20), [redacted] phakic eyes have been followed up for 36 months after fluocinolone insertion. Their mean BCVA is [redacted] letters, but we are not told what their baseline BCVA was. In the whole group, baseline BCVA was [redacted]. With no control group, we do not know how much of any improvement is spontaneous, or due to improved glycaemic control.

RISE and RIDE

These two studies, reported by Brown et al [22], compared ranibizumab with sham injection. Alimera uses these studies for data on the efficacy of ranibizumab in fellow eyes. After 24 months, the sham group could cross over to treatment with ranibizumab, and in their submission, Alimera make a case (pages Table B2, and 84-85) that the cross-over group can be used to provide data on chronic DMO;

"Patients treated with ranibizumab 0.5 mg after 2 years on sham were considered representative of patients with chronic DMO, and therefore results for this group were included in this review".

This is a reasonable assumption, but because they are new to ranibizumab, they may not match the population in the NICE scope in which eyes being considered for fluocinolone treatment are unresponsive to anti-VEGF treatment. DMO is often bilateral and responses may be similar in both eyes. Alimera therefore add a caveat on page 84,

"This period from month 24 to months 36 corresponds to treatment for chronic DMO; however, they are not classified as unresponsive to treatment "

To overcome this problem, Alimera (page 121) assume that that fellow eyes might also be insufficiently responsive, that half the patients were assumed not to respond and for them, the efficacy in the sham arm of FAME was used. The assumption that half of eyes that are naïve to anti-VEGF treatment, will not respond to it, is reasonable given the data from ranibizumab trials.

However, the issue is complicated by the number of sham patients who did not cross-over and so did not receive ranibizumab. The ITT analysis of the whole group, using last observation carried forwards gives a letter gain of only 2 letters (from 2.5 at 24 months to 4.5 at 36 months).

One problem is that the results of the sham cross-over arm are not entirely clear. Only 190 (70%) of the 257 sham patients crossed over to ranibizumab. At 24 months, the whole sham group had a mean gain of 2.5 letters.ranibizumab. But we do not have 24 month data separately for the 190 patients that crossed over, and we are not told what their mean BCVA was at month 24. At month 36, in ITT analysis, the whole sham group, including cross overs and non-cross-overs, had mean gain from baseline of 4.5 letters. However that mean includes LOCF from the non-cross-overs. So it would be incorrect to say that the cross-over group gained only a further 2 letters, because, firstly, we don't know what their BCVA was at month 24, and secondly, the 4.5 letters gain from baseline (average of RISE and RIDE) at month 36 includes the non-cross overs. In the text (Brown et al) we are told that those who crossed over and had at least one injection of ranibizumab (most had more, with mean of 10 injections) had mean gains of 7.5 letters (RIDE) and 7.8 letter (RISE) at month 36, from baseline. However, this presumably means compared to the whole sham group at baseline, not the subgroup that crossed over. If we use the letter gain of 2.5 in the whole sham group at the cross-over point at 24 months, this would suggest a gain of about 5.2 letters.

The Nguyen and Brown papers report three outcomes for sham and 0.5mg ranibizumab patients: a] the proportion of patients that achieved a gain of at least 15 ETDRS letters and loss of < 15 letters at various time points (24 months Nguyen, and 36 months Brown); b] the mean change in ETDRS letters from baseline plotted at 2 monthly intervals with error bars to 24 months in Nguyen and 36 months in Brown; c] Brown (Table 3) provided mean and SD change in letters together with corresponding proportion

gaining at least 15 letters at 12 months after start of ranibizumab administration (for the sham arm this was for the 12 months following the option for cross over at 24 months of the trial).

Different odds ratios can be derived depending on which time points from which data are taken.

The Sham patients after 12 months on ranibizumab (i.e. following 24 months on Sham) can be regarded as chronic patients. This is argued by Alimera and the ERG agrees. However the 0.5 mg active arm at 12 months may not have reached a chronic state, so the comparison may be inappropriate. RIDE and RISE trials data for the 0.5mg ranibizumab arm for gain in 15 letters is available at 12 months (Brown Table 3), at 24 months (Nguyen Figure 3A) and at 36 months (Brown Table 2 and Fig 1). The OR for gain versus Sham (i.e. sham 12 months after cross over) using 12 month data (Brown Table3) is 5.9; using 24 month data for 0.5mg arm the OR is 9.66 (0.738399/0.076426); using 36 month 0.5mg data the OR is 9.04 (0.6913 / 0.076426).

OR for loss of 15 letters or more (0.5 mg vs. sham prior to cross over) at 24 months is 0.318 . At 36 months (including 12 months on ranibizumab for the sham arm) the OR is 0.368.

The key point is that results in the group that crossed over from sham to ranibizumab at 24 months are poorer than those in the group that received ranibizumab at the start, two years earlier. Chronic DMO responds more poorly. After 12 months on ranibizumab, the cross-over group had a mean BCVA letter gain of 2.8, compared to the gain of 11 letters in the ranibizumab 0.5mg arm after 12 months. The proportions gaining 15 or more letters after 12 months were 7.3% in the cross-over group and 31.7 in the ranibizumab arm.

……

There is some overlap amongst the centres in the Medisoft, ICE-UK and IRISS, but this is not a problem. Only four of 14 Medisoft centres were also amongst the 13 ICE-UK centres. About seven Medisoft centres were also amongst the 31 UK centres in IRISS, which also had 11 in Germany and five in Portugal.

3.3 Conclusions of clinical effectiveness review

- The FAME trial showed that fluocinolone is effective in people with chronic DMO that has not responded sufficiently to laser, though patients may have had only one laser treatment, so it is not fully clear whether patients recruited to FAME were truly unresponsive to laser. Mean baseline retinal thickness was 461.8 microns, making it less likely for laser to be effective. However, FAME cannot provide evidence on effectiveness in DMO that has not responded to anti-VEGF treatment, which is the key group for this appraisal, because the anti-VEGFs were not widely available when FAME was carried out.
- The most relevant analysis from the FAME trial is that by Yang et al [39] which reports results in people phakic at baseline. This analysis suggests that those phakic at baseline may gain 11 letters by 3 years. However that improvement include both the fluocinolone effect and the changes seen in the sham group, so the net effect will be less – possibly 9 letters.
- There are some important unknowns including;

1. In patients not responding to anti-VEGFs, what is the marginal gain with fluocinolone compared to continuing anti-VEGFs?
2. What is the cost per QALY of continuing anti-VEGFs in people who don't respond to these drugs? We suspect this would be high unless bevacizumab was used. And perhaps high even then.
3. In patients not responding to anti-VEGFs, what is the marginal gain with fluocinolone compared to no treatment?

- Ideally we would have had a trial in people with DMO not responding to anti-VEGF treatment, randomising to continued anti-VEGF (on the grounds that while not showing improvement, it might be preventing deterioration) or to fluocinolone.
- Alimera cite evidence from a number of observational studies. These show that in people with DMO that has failed to respond to anti-VEGFs, there is a response to fluocinolone though usually with a smaller mean gain in letters than seen in the trials.
- The most useful of the observational studies are in patients with chronic DMO that has not responded to previous treatment, including with anti-VEGF drugs. The improvements in BCVA are not dramatic – a mean 5.3 letters at 24 months in the Medisoft study, and a mean of about 3 letters at 12 months in the ICE-UK study. However these results are mainly in pseudophakic patients. Some studies reported a decline in VA in the year prior to fluocinolone treatment. So stability without improvement may be a useful outcome. In trials, it may be that people try harder, or have more time, when their BCVA is being measured. If so, results in observational studies in routine care may not be quite as good as in trials.

4 COST EFFECTIVENESS

Note that within this chapter all references to FAME data are restricted to the FAME phakic at baseline subgroup unless otherwise stated

The modelling uses some FAME data that is not restricted to those who were phakic at baseline. Transitions for those who become pseudophakic are estimated from pooling their pseudophakic transitions with the transitions among the FAME pseudophakic at baseline subgroup.

4.1 ERG comment on company's review of cost-effectiveness evidence

The objective of the cost effectiveness review was to identify relevant cost-effectiveness evaluations of treatments for DMO.

4.1.1 Search strategy

The databases searched were: Ovid MEDLINE(R), EMBASE (Ovid), Cochrane library, and the Centre for Reviews and Dissemination (CRD) database. The searches were last updated on 16 October, 2018. The search identified 509 records after removal of duplicates, and 486 were excluded after screening. The ERG also did its own independent searches.

4.1.2 Inclusion criteria

Studies considering the UK were included, and no additional geographical restrictions were imposed. The search strategy and selection process of relevant studies are detailed in appendix G.

4.1.3 Included studies

Alimera included 8 studies in Table B3.1 but the ERG did not consider these to be relevant;

- The Alimera 2012 analysis is not of phakic patients
- The previous manufacturers' submissions, Novartis 2013, Bayer 2012, and Allergan 2014 were of treatment-naïve patients
- Mitchell 2012 was also about treatment-naïve patients from the RESTORE trial of ranibizumab versus laser
- Beiderbeck 2017 (from Alimera) was an abstract and examined the cost-effectiveness of fluocinolone compared to anti-VEGF and laser, implying that it was not in eyes that did not respond to those treatments

- Regnier 2015 was also in new patients, and compared aflibercept and ranibizumab
- Royle 2015 was about diabetic retinopathy, comparing laser pan-retinal photocoagulation at severe non-proliferative retinopathy versus waiting till proliferative DR developed

Alimera did send on two other economic studies, not included in their review above, but neither seem relevant. Haig was from Novartis about ranibizumab versus laser in Canada, in treatment-naïve patients based on RESTORE. Hutton was about ranibizumab vs laser in proliferative retinopathy, not DMO. [55]

Alimera did not include the economic study by the Cardiff group, Holden, Currie and Owens, which is odd because Alimera designed and supported it. There are several papers from this study, two of which focused on the clinical effectiveness and were described earlier, first authors Currie and Holden.[47, 56] The economic paper, Holden et al 2017[57], compared costs in the year before fluocinolone insertion with costs in the year after. Most (82%) of patients had had prior anti-VEGF treatment, though the median number of injections is reported here as 3 (1 to 6) whereas in the clinical effectiveness paper it was reported as 5 (2-17). The costs after implantation of fluocinolone included 141 anti-VEGF injections in 233 eyes, compared to 649 in the year before, yielding considerable savings (Table 3 of Holden et al [57]. However because of the high cost of the fluocinolone implant, total costs in the year after were £1.6 million compared to £627,058 in the year before. Longer follow-up spreading the fluocinolone cost (about £1.3 million) over 3 years would have provided a more favourable picture. 88% were pseudophakic at baseline, and another 7% had cataracts removed at the same time as fluocinolone was implanted.

So the study by Holden et al cannot tell us much about the economics in phakic patients.

Another study omitted is Pochopien 2018 from Alimera[58] on the cost-effectiveness of fluocinolone in UK patients with chronic DMO, though reference 38, the abstract by Beiderbeck and colleagues has most of the authors from the Pochopien article, and may be an earlier form of it. Pochopien and colleagues report the result for phakics separately, estimating an ICER of £28,751. This does not match the dominance in the Alimera base case. A key difference is that the cost of "usual care" is lower than the fluocinolone arm in Pochopien (Table 3) but not in the Alimera submission, Table B3.42.

A study by Neubauer et al[59] was published very recently (after Alimera had sent their submission), in German but with an English abstract. It assessed the costs over 3 years of fluocinolone in patients with DMO insufficiently responsive to anti-VEGFs, compared to continuing ranibizumab or aflibercept. The conclusion was that fluocinolone was less expensive, by about 5,000 euros compared

to aflibercept and about 7,000 euros compared to ranibizumab, with the main difference being in the frequency of administration and hence the costs of the drugs. The option of using bevacizumab was not included. Clinical effectiveness does not seem to have been reported. The authors do consultancy work for Alimera.

Treatment

A key issue in cost analysis is how phakic patients who are insufficiently responsive are treated if fluocinolone is not available. Will anti-VEGF injections be continued? In some patients, if fluid remains after anti-VEGF treatment, but retinal thickness is less than 400 microns, laser could then be used as per NICE guidance. There appears to be a lack of consensus on stopping rules for anti-VEGF treatment. The NICE guidance on ranibizumab and aflibercept does not include stopping rules.

Alimera believe that many ophthalmologists will continue anti-VEGF treatment even in poorly responsive eyes. This is very unlikely to be cost-effective with ranibizumab and aflibercept due to their high costs (even with confidential discounts) but data are lacking.

4.1.4 Conclusions

The main problem in the cost-effectiveness literature is that, as in clinical effectiveness, we lack studies based on trials of fluocinolone in eye that have not responded sufficiently to anti-VEGF treatment. Another problem is that some of the studies such as Holden[57] and Neubauer [59] look only at short-term costs.

Of the cost effectiveness papers, based on the company summary of table B3.1. the Allergan analysis that compares fluocinolone with dexamethasone is perhaps the most relevant. While the fluocinolone PAS limits the relevance of the estimated ICER, it can be noted that the net gain from fluocinolone over dexamethasone in patients without adequate response to corticosteroid therapy is only 0.079 QALYs over a 15 year time horizon. This is somewhat less than the company estimate of the current submission for the comparison with usual care, though the patient group obviously differs to an extent.

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

4.2.1 NICE reference case checklist

Table 4. NICE reference case checklist

Attribute	**Reference case and TA Methods guidance**	**Does the *de novo* economic evaluation match the reference case**
Comparator(s)	Therapies routinely used in the NHS, including technologies regarded as current best practice. The scope specifies "*Established clinical management without fluocinolone acetonide intravitreal implant*".	Fluocinolone is compared to a composite comparator based upon the balance between treatments prior to fluocinolone implant in the ICE-UK study: • 28% laser / no treatment • 63% ranibizumab • 9% bevacizumab An immediate issue is that these treatments are assumed not to have provided sufficient improvement, so it is questionable whether they should be continued.
Patient group	As per NICE scope. " *People with chronic diabetic macular oedema that is insufficiently responsive to available therapies who have phakic lenses*".	The company presents results mainly based around an analysis of the FAME trial data, restricted to those who were phakic at baseline. The company also presents the base case analysis split by the subgroups of those with cataract at baseline and those without cataract at baseline.
Perspective costs	NHS & Personal Social Services	Yes.
Perspective benefits	All health effects on individuals	Yes.
Form of evaluation	Cost-effectiveness analysis	Cost utility analysis.
Time horizon	Sufficient to capture differences in costs and outcomes	30 years, which given the baseline age of 64 years is effectively a lifetime horizon.
Synthesis of evidence on outcomes	Systematic review	The company mainly relies upon the phakic subset of FAME.

		During years 1-3 of the model, for fluocinolone BCVA changes are based upon the change from baseline in the FAME fluocinolone arm. For the comparator arm it is assumed that there is no change in BCVA from baseline. During years 4-6 of the model, for the fluocinolone arm a proportion equal to the FAME fluocinolone arm proportion who improved by at least 15 letters receive a 2nd implant. Among these patients the FAME fluocinolone arm month 4-36 effectiveness data is reapplied and there are additional gains in BCVA. Those who discontinue fluocinolone retain the gains of years 1-3. In the comparator arm it is assumed that there is no change in BCVA.
Outcome measure	Quality adjusted life years	Yes.
Health states for QALY	Described using a standardised and validated instrument	The base case uses the NEI-VFQ-25. FAME collected NEI-VFQ-25 data at baseline, 24 months and 36 months during FAME. Rentz et a[60] developed an algorithm from a subset of 6 elements of the NEI-VFQ-25 to estimate the quality of life values for 8 of the possible 15,625 health state values for the NEI-VFQ-25. The resulting algorithm was applied by the company to the FAME data to estimate quality of life values for the bilateral health states. Scenario analyses that use the Brazier et al [61] BCVA to NEI-VFQ-25 mapping and the BCVA to EQ-5D mapping are also presented.

Benefit valuation	Time-trade off or standard gamble	Time trade off.
Source of preference data for valuation of changes in HRQL	Representative sample of the public	Unclear. Rentz et al recruited 607 members of the general Australian, Canadian, US and UK general public. This was mostly through newspaper adverts, though some of the 152 UK participants were recruited due to having participated in a previous study. The mean age of 43 was lower than the 64 years of FAME.
Discount rate	An annual rate of 3.5%.	Yes.
Equity	QALYs have the same weight.	Yes.
Probabilistic modelling	Probabilistic modelling	Yes. Though the majority of coefficients in the clinical effectiveness regression is not sampled.
Sensitivity analysis		A wide range of univariate sensitivity analyses are included. A range of scenario analyses are presented.

4.2.2 Model structure

The model independently simulates a cohort of study eyes and a cohort of a fellow eyes. These are then combined into bilateral health states.

The distributions across the health states for both the study eyes and the fellow eyes are taken from ICE-UK. Due to the limited number of patients and conflicting evidence on the impact of cataract removal, it was assumed that those with and without cataracts had identical distributions at baseline.

Table 5. Model health states and baseline patient distributions

Health state	HS1	HS2	HS3	HS4	HS5	HS6	HS7	HS8
BCVA letters	100-86	85-76	75-66	65-56	55-46	45-36	35-26	25-0
Study eye:								
No cataract	1%	2%	18%	27%	19%	15%	6%	12%
With cataract	1%	2%	18%	27%	19%	15%	6%	12%
Fellow eye:								
No DMO no cataract	13%	30%	13%	18%	25%	0%	0%	3%
No DMO with cataract	13%	30%	13%	18%	25%	0%	0%	3%
DMO no cataract	7%	19%	24%	14%	10%	9%	5%	12%
DMO with cataract	7%	19%	24%	14%	10%	9%	5%	12%
Pseudophakic	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.: Not applicable. All fellow eyes are assumed phakic at baseline.								

For the study eye the balance between those with and without cataract at baseline is taken from the Retro-IDEAL study: 50:50.[54] The Retro-IDEAL proportion with cataracts is considerably higher than that of FAME.

For the fellow eye the balance between those with and without DMO at baseline is assumed to be the percentage of fellow eyes in ICE-UK[47] with a history of treatment for DMO: 77%. It is assumed that among fellow eyes with DMO at baseline the balance between those with and without cataract is as for the study eye: 50:50. It is also assumed that all fellow eyes are phakic at baseline.

The model structure for the cohort of study eyes is presented below. This does not show death which is possible from all health states and is modelled as occurring at the start of period (SoP) prior to the transition probability matrices (TPMs) being applied which then lead to the end of period (EoP) at which point transitions between the patients subsets occur.

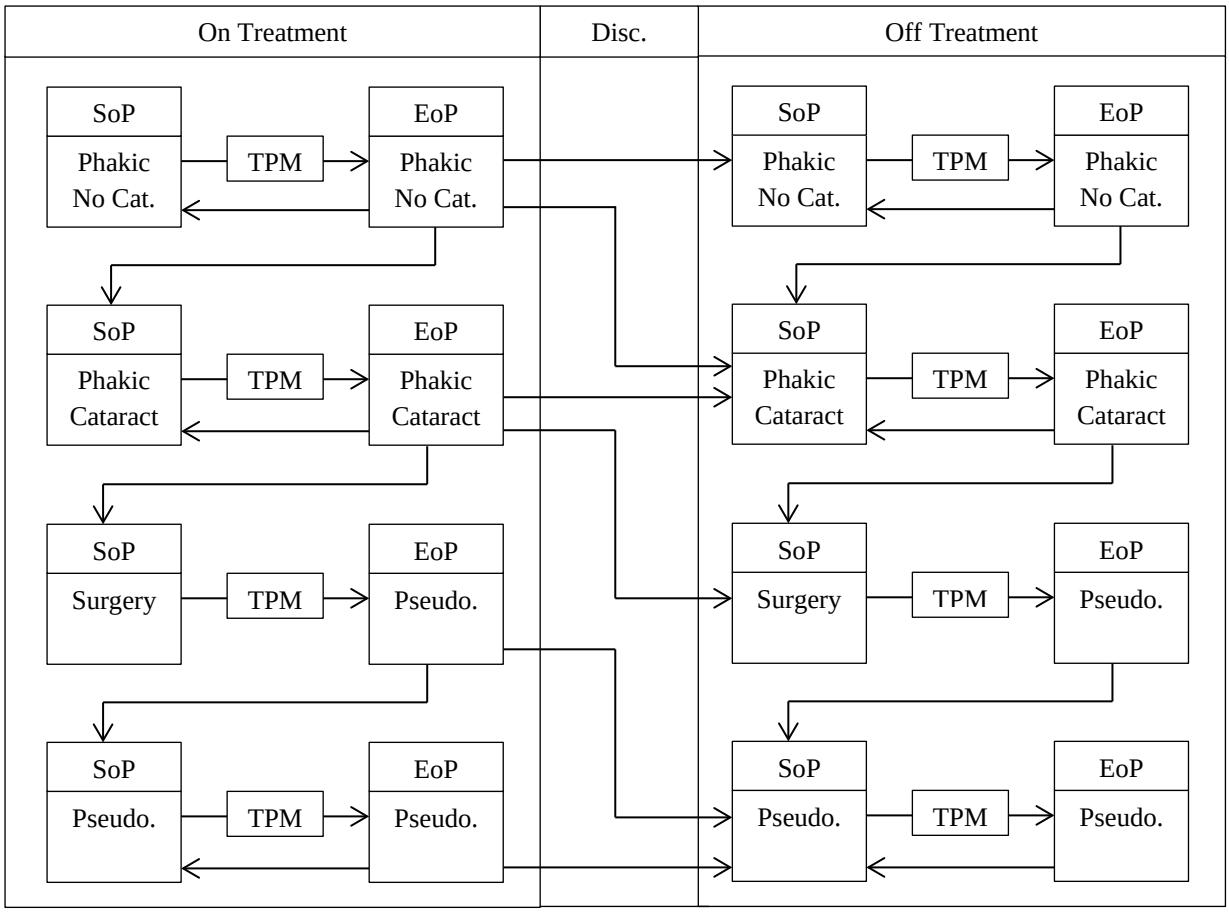


Figure 4. Model structure: cohort of study eyes

All study eyes start the model on treatment split 50:50 between phakic with cataract and phakic without cataract.

For each of the four patient subsets:

- Phakic with cataract,
- Phakic without cataract,
- Undergoing surgery, and
- Pseuodophakic

movements between the eight BCVA health state are determined by applying the patient subset and treatment specific transition probability matrix (TPM). After the TPMs have been applied treatment specific probabilities of:

- developing cataracts, and
- having cataract surgery

are applied. Cataract surgery is a tunnel health state lasting one model cycle with patients then moving into the pseudophakic subset.

The electronic model permits discontinuations for those in the usual care arm during each model cycle. But all company modelling assume none discontinue usual care until the end of year 6, when all discontinue.

Due to fluocinolone being an implant none discontinue and the implant lifespan is assumed to be 3 years. At the end of 3 years a 2nd implant is possible and is assumed to occur to 36% of patients, the proportion who gained at least 15 letters in the FAME fluocinolone arm. As a consequence, 64% are assumed to discontinue fluocinolone treatment at the end of year 3 and all are assumed to discontinue fluocinolone at the end of year 6.

The model structure for the cohort of fellow eyes is within the electronic model very involved due to the electronic model attempting to permit prevalent and newly incident fellow eye DMO to be treated differently between the arms. But this aspect of the electronic model is not applied within the company submission and the model structure is essentially as below. Patients at baseline are either with or without bilateral DMO in their fellow eye: 33:77. Those without bilateral DMO may subsequently develop it, and are assumed to have the same transition probabilities as those who were bilateral at baseline. Those with bilateral involvement receive usual care. For present purposes the model structure for the fellow eye is as presented below.

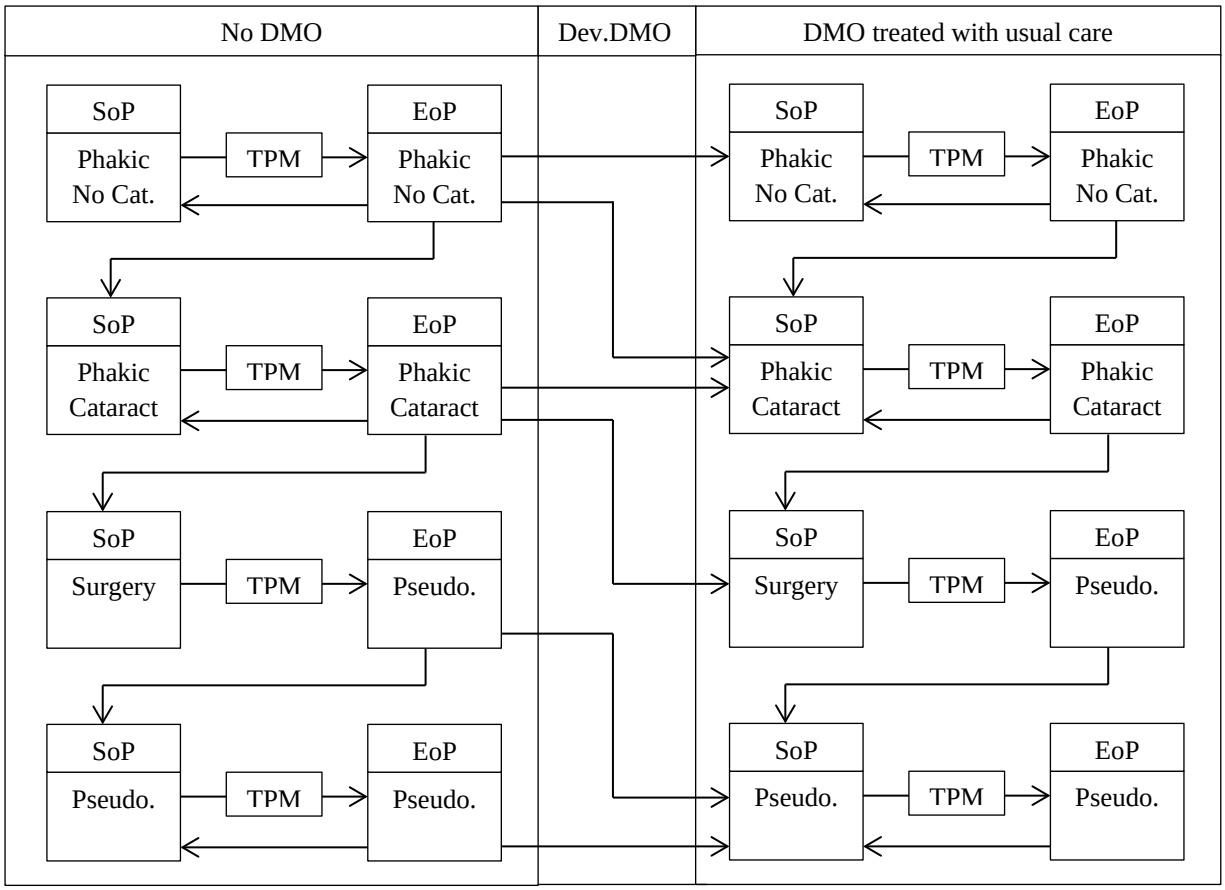


Figure 5. Model structure: cohort of fellow eyes

After year 6 it is assumed that all cease treatment, and in effect the On Treatment usual care section switches into an Off Treatment natural history section. Since the modelling of the fellow eye is common to both arms and largely nets out the ERG does not particularly dwell upon it. Its main effect is to determine the proportion of patients whose bilateral vision falls into blindness, which is in turn determined by the natural history worsening that is assumed to apply from year 6.

The distributions of the study eyes and the fellow eyes can be combined into a bilateral health states using either of two approaches.

- The distributions are independent; i.e. having a fellow eye in HS8 has no bearing upon the likelihood of the study eye being in HS1.
- The distribution of the study eye is conditional upon the health state of the fellow eye; i.e. if the fellow eye is in HS8 it is more likely that the study eye will be in HS8 than in HS1.

The company base case assumes the distributions are independent. This choice determines how likely it is that study eye is the better seeing eye (BSE) and how likely it is that the study eye is the worse seeing eye. This matters because changes to the BCVA of the BSE have a larger effect upon patient QoL, and the company does not model bilateral treatment with fluocinolone.

4.2.3 Population

The stated population is phakic patients who have shown insufficient response to previous therapies. 61% are male based upon ICE-UK. The company base case also assumes that at baseline 50% are free of cataract and 50% have cataract. This is based upon Retro-IDEAL. The proportion in FAME was very much lower due to clinically significant cataracts being an exclusion criterion.

Quite what the population is is complicated by the company also stating in its clarification response "*In clinical practice, 0.2 µg/day fluocinolone acetonide implant would only be considered for use in phakic patients in the following instances: a) in patients with pre-existing cataract and who would require cataract surgery in the next 1-3 years; b) in a small group of patients who are contraindicated for first line therapies or are needle phobic where the benefit of protecting the retina outweighs the risk of cataract formation*".

The patient population distribution, TPMs and probabilities of cataract surgery are mainly based upon the FAME trial. As with TA301, FAME provides the vast majority of the clinical inputs to the economic model. FAME was conducted prior to the general availability of anti-VEGFs, and as a consequence the company has to make a number of assumptions to relate the data from the FAME trial population to each of the three comparators which make up the composite comparator outlined below.

4.2.4 Interventions and comparators

Fluocinolone is compared to a usual care composite comparator of:

- 28% no drug treatment, labelled as laser treatment
- 63% ranibizumab, and
- 9% bevacizumab.

The balance of treatments within the composite comparator is taken from ICE-UK. The ERG assumption is that this is not restricted to the subgroup of ICE-UK who were phakic at baseline due to their limited number (26/208).

4.2.5 Perspective, time horizon and discounting

The perspective and discounting as per the NICE reference case. The time horizon is 30 years, which given the baseline age of 64 is effectively a lifetime horizon.

4.2.6 Treatment effectiveness and extrapolation

Transition probability matrices (TPMs) for transitions within the four patient subsets

The treatment effectiveness in the study eye of each arm is largely driven by quarterly transition probability matrices (TPMs) that are applied, but the probabilities of developing cataract and having cataract surgery as reviewed in the next subsection below are also important.

The TPMs determine how patients of each of the four patient subsets of the model:

- No cataract
- With cataract
- Undergoing cataract surgery, a tunnel health state lasting one model cycle
- Post cataract surgery

move between the eight BCVA health states. Each patient subset has a subset and treatment specific TPM applied.

The TPMs are estimated from FAME data using a random effects logistic regression using Mixed Model for Repeated Measures (MMRM) and the entire FAME data set. The regression splits the data by arm, fluocinolone or sham, by patient subset and by whether the quarterly transitions are during the 1st 3 months of FAME or whether the quarterly transitions are during months 4-36 of FAME. The TPMs derived for the 1st 3 months of FAME are applied within the 1st model cycle, while those of months 4-36 of FAME are applied within the model up to the end of year 6 when all are assumed to cease treatment. This assumes that, for a given patient subset and treatment arm, that the probabilities of moving between health state during, say, the 2nd quarterly cycle is the same as those during the 12th quarterly cycle. After the end of year 6 all are assumed to have a constant 3.5% quarterly probability of worsening by one health state.

The MMRM analysis is not restricted to the phakic at baseline to maximise the number of patient transitions for the pseudophakic, these transitions being applied to the post cataract surgery #262patient subset. The submission presents no detail of the logistic regressions or how the final regression was selected, so the ERG presents the company models in Appendix 4 and Appendix 5. At clarification the ERG asked how lost to follow-up was handled. The company clarified that "*Patients who were lost-to-follow-up were included in the analysis*" but provided no further details of this. The AIC and BIC of the company models are presented below, for models that permitted differing probabilities between the 1st quarter and subsequent quarters, (A), and for the corresponding models that permitted different probabilities at all visits (B). It appears that models between these two

extremes, such as permitting 1st quarter and thereafter annual time interaction effects, were not explored. The company chose model 1 (A) for the cost effectiveness modelling.

Table 6 FAME patient transitions regressions' information criteria

Model	1 (A)	2 (A)	3 (A)	4 (A)	5 (A)	6 (A)	7 (A)	8 (A)	9 (A)
AIC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BIC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Model		2 (B)	3 (B)	4 (B)	5 (B)	6 (B)		8 (B)	9 (B)
AIC		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]
BIC		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]

From the clarification response it appears that there are not results for two model (B)s due to one not being estimated and the other not converging. As a consequence it is not clear to the ERG whether the company did not estimate model 1(B) or did not estimate model 7(B). Given the company choice of model 1 (A) it may a concern if the company did not estimate model 1 (B).

At clarification the company stated that the (A) models that allowed for differing probabilities between the 1st quarter and subsequent quarters had superior AIC and BIC compared to the (B) models that permitted different probabilities at all visits, so the company preferred the former. Within these, model 9 has the lowest AIC and BIC, by some margin. The company states that "*The regression model used in the CEA model is model 1, which also includes an interaction between visit and BCVA level at last visit. The reason for keeping this interaction is based on the selection of regression model for a cost-effectiveness analyses previously performed for pseudophakic patients.*[58] *The visual inspection of BCVA curves over time had suggested that this interaction should be included in the model, at the predicted curves better fitted the curves directly obtained from the trial*". It is not clear whether the visual inspection of BCVA relates to the previous modelling of pseudophakic patients or to the current modelling. No detail of this visual inspection is provided by the company.

It can be noted that models 5, 6, 7 and 8 also included the interaction term for visit and BCVA level at last visit, that models 6(A) and 8(A) have lower AIC than model 1(A) and that model 8(A) also has a lower BIC than model 1(A). The company has provided the coefficients of model 9(A) at clarification, but the ERG does not have access to the coefficients of model 6(A) or model 8(A) so cannot explore their possible effects on the cost effectiveness estimates. Company estimates show that for the company base case applying model 9(A) affects net costs by 3% and net QALYs by 1%.

The logistic regression provides a set of four fluocinolone TPMs, one for each patient subset, for and another set of four sham TPMs.

The company constructs a set anti-VEGF TPMs by applying an odds ratio to the probabilities of improving in the corresponding sham TPMs. The odds ratio that is applied of 1.54 is derived from the ranibizumab RISE/RIDE trials. This odds ratio is applied to the sham TPMs for the phakic without cataract, the phakic with cataract and the pseudophakic to derive the corresponding anti-VEGF TPMs. The anti-VEGF TPM for cataract surgery is assumed to be the same as the sham TPM for cataract surgery.

A set of usual care TPMs is then constructed by weighting the probabilities of the sham TPMs by 28% and the probabilities of the anti-VEGF TPMs by 72% and then summing these. These usual care TPMs are applied to usual care in the fellow eyes. The company does not apply these usual care TPMs to the study eyes.

For the study eyes of the usual care arm the company base case assumes no change in BCVA and so applies the identity matrix. Scenario 1 of the company applies the sham TPMs to the study eyes of the usual care arm.

The TPMs that are applied in the company base case are outlined below in Table 7.

Table 7. TPMs applied in the company base case years 1-6

Eyes	Fluocinolone: study eyes		Usual care: study eyes		Fellow eyes	
Treatment	On Tx	Off Tx	On Tx	Off Tx	No DMO	On Tx
TPMs	FAME FLU	Identity	Identity	Identity	Identity	Usual care

The main elements to take from the above are:

- Patients getting fluocinolone have all the benefits observed in the FAME fluocinolone arm during years 1,2 and 3.
- Patients getting a 2nd fluocionolone implant have additional absolute benefits during years 4, 5 and 6.
- Patients getting usual care have none of the benefits observed in the FAME sham arm

A natural history TPM is also applied from year 6 which assumes 3.5% of patients worsen by one health state each quarter. This appears to be based upon the ranibizumab TA274. This mean that the

net benefit modelled at the end of year 6 for fluocinolone over usual care is extrapolated for the next 24 years,

The TPMs determine how patients transition between the eight BCVA health states within each patient subset. The TPMs do not determine how they discontinue treatment which is by assumption as outlined above with 64% of fluocinolone patients discontinuing at the end of year 3, due to not having achieved a 15 letters gain, with the remaining 36% of fluocinolone patients discontinuing at the end of year 6 and 100% of usual care patients discontinuing at the end of year 6. The TPMs also do not determine how they move between the four patient subsets, the key transitions being the probabilities of developing cataract and the probabilities of having cataracts removed.

The probability of developing cataracts

The probability of developing cataracts among those modelled as cataract free at baseline is based upon data from phakic chronic FAME data. The Kaplan Meier data underlying these calculations was supplied at clarification and is presented graphically below[1]. The ERG assumes that this data relates to the study eye, but this is confused by Table B3.12 of the company submission only referring to the probability of developing cataract in the fellow eye. The probabilities of developing cataract in the study eye in the model are as per Table B3.12.

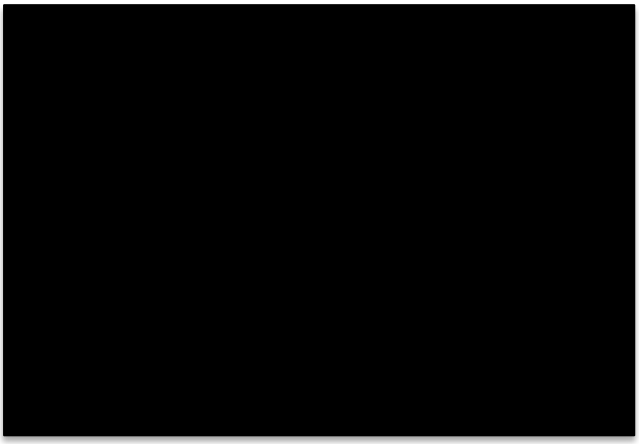


Figure 6. *Kaplan Meier proportion and modelled proportion remain cataract free*

The analysis is based upon [redacted] patients in the fluocinolone arm being cataract free at baseline, but a considerably lower number in the sham arm, only [redacted], due to FAME recruiting 375 patients to the

[1] It is more correct to present the points of the Kaplan Meier as a step function, but the figure conveys the essence of the data.

0.2µg fluocinolone arm but only 185 to the sham arm. By the end of 36 months the Kaplan Meier plot for the proportion remaining free of cataracts is [redacted] compared to [redacted] for sham.

The company does not adopt the usual method of fitting, say, an exponential and more simply totals the number of events and divides these by the sum of each quarter's number at risk to arrive at quarterly estimates for the probability of cataract of 18.1% for fluocinolone and 6.8% for sham. These are little different from assuming an exponential based upon the 36 month proportions.

The probability among those with cataracts of having them removed

The quarterly probability of having cataracts removed is based upon a regression analysis of the chronic phakic FAME data. The company constructed two sets of models: (A) a set with all variables as main effects and with no interactions, and (B) a set that included an interaction effect between the treatment variable and the other variables. A generalised estimating equation model was used with health state and treatment arm as explanatory variables. The company states that the evolution of the models was based on sequentially removing variables with a p-value of more than 0.05. This resulted in the following information criteria, QIC and QICu, though note that unlike the TPM regression there is no immediate vertical read through between the two sets of models (Table 8 and Table 9).

Table 8. Probability of cataract surgery regressions' information criteria

Models (A)	1	2	3	4	5	6	7
QIC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
QICu	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Models (B)	8	9	10	11	12	13	14
QIC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
QICu	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

In its clarification response the company acknowledges that model 14 has the lowest information criteria. But it states that this led to an overfitting of the model to the data as it gave rise to the following probabilities of cataract removal by health state, split by gender and treatment, The ERG pools these values across gender based upon a weighted average according to the ICE-UK population.

Table 9. Probabilities of cataract surgery model 14

Health state	HS1	HS2	HS3	HS4	HS5	HS6	HS7	HS8
Fluocinolone (male)	[redacted]							
Fluocinolone (female)	[redacted]							
Fluocinolone (pooled)	[redacted]							
Sham (male)	[redacted]							
Sham (female)	[redacted]							
Sham (pooled)	[redacted]							

Bearing in mind that the model does not differentiate by gender the values of model 14 would have to be pooled regardless. Other than the high probability of cataract removal for HS1 in the sham arm which will relate to very few patients in both the data and the model, it is not so obvious to the ERG that the pooled values are particularly over fitted when viewed alongside the company preferred model 1 as tabulated below.

Of the models without interaction terms the company notes that model 6 has the lowest information criteria but opted for model 1 which excluded gender as a variable due to it having a p-value of [redacted] in model 6. But it can also be noted that for model 1 the treatment effect coefficient has a p-value of [redacted], yet is retained. The p-values for the treatment effect in models 2 and 3 worsen to [redacted]. Model 7 excludes both gender and the treatment effect and has a lower QIC and QICu than model 1 (see Table 10).

Table 10. FAME quarterly probability of cataract removal by health state: Models 1 and 7

Health state	HS1	HS2	HS3	HS4	HS5	HS6	HS7	HS8
Including Tx effect:								
Fluocinolone	0.0%	4.3%	7.4%	27.9%	22.9%	28.3%	26.1%	45.8%
Sham	0.0%	3.5%	6.0%	23.8%	19.4%	24.2%	22.2%	40.6%
Excluding Tx effect:	0.0%	4.0%	6.9%	26.0%	21.6%	26.5%	25.4%	45.6%

The company retained the fluocinolone versus sham treatment effect, and assumed that the anti-VEGF treatment effect would be that of sham. The company justified this at clarification by noting that "*the model with treatment was deemed more appropriate from a clinical perspective, since patients treated with a 0.2 µg/day fluocinolone acetonide implant have a faster progression of cataract and may therefore be treated earlier*". In other words, cataracts arising from fluocinolone use are more

troublesome than cataracts arising in the usual care arm, possibly because they are more likely to be posterior sub capsular, and need to be and are removed more quickly. But the company statistical analyses do not demonstrate this, the p-value for the treatment effect is [redacted] and the models without a treatment effect have better quasi-likelihood criteria.

Adverse events: Treatment related

The probabilities of adverse events for fluocinolone and sham/laser are taken from the TA301 submission, which were in turn drawn from FAME data (see Table 11). For the anti-VEGFs they were taken from the ranibizumab trials' data. The reference cited, is incorrect, being the T2 diabetes guideline[62] which doesn't discuss treatment of DMO. The ERG expects that the reference should be to the RISE and RIDE trials.[22]

Table 11. Annual probabilities of adverse events: Treatment related

	Fluocinolone			Sham/laser			Anti-VEGF	
	Yr1	Yr2	Yr3	Yr1	Yr2	Yr3	Yr1+Yr2	Yr3
IOP	26.3%	13.6%	9.6%	5.3%	5.3%	5.3%	16.4%	7.9%
Retinal detachment	4.1%	4.1%	4.1%	4.1%	4.1%	4.1%	0.4%	0.2%
Endophthalmitis	0.8%	0.8%	0.8%	0.8%	0.8%	0.8%	0.8%	0.4%
Vitreous haemorrhage	1.0%	1.0%	2.3%	2.6%	2.6%	2.6%	0.8%	0.4%
Glaucoma	0.5%	2.1%	5.7%	0.0%	0.0%	0.0%	2.8%	1.4%

Adverse events: Complications of cataract surgery

Cataract surgery results in 0.17% having retinal detachment and 0.34% having endophthalmitis, based upon Norregaard[63] and Chan.[64]

Extrapolation during years 3 to 6

Extrapolation for years 3 to 6 applies the same TPMs, probabilities of developing cataract and probabilities of having cataracts removed as summarised above. Most notably, and as graphed in the model validation section below, those receiving a 2nd fluocinolone implant not only retain the gains in BCVA of the 1st 3 years but are modelled as having additional gains in their BCVA during years 4-6 that are of a similar magnitude as the modelled gains of the 1st 3 years.

Extrapolation during years 6 to 30

From the end of year 6 all patients have discontinued treatment in their study eye and a natural history TPM is applied which assumes 3.5% of patients worsen by 10 letters every quarter. This extrapolates

the same worsening of BCVA in both arms. As a consequence it maintains the year 6 net gain in BCVA from fluocinolone over usual care for the remaining 24 years of the model at no additional treatment cost.

Mortality

A mortality multiplier for DMO of 1.95 is reportedly taken from Preis et al[65], and is applied to mortality probabilities taken from England and Wales life tables.

An additional mortality multiplier of 1.23 for blindness taken from Christ et al[66] is applied in addition to this for eyes falling into HS7 or HS8; i.e. less than 35 letters. The blindness mortality multiplier is applied independently to the study eye and to the fellow eye, though the company model notes that it should only be applied to the better seeing eye (BSE).

This approach yields estimates for the number of study eyes that are surviving and the number of fellow eyes that are surviving. Patient survival is modelled as the average of the surviving study eyes and fellow eyes.

4.2.7 Health related quality of life

There are three quality of life elements:

- The quality of life associated with each of the 8*8=64 bilateral health states, which comprises the vast majority of the QALY calculation.
- The quality of life decrements for adverse events, including those arising from cataract surgery.
- For the anti-VEGF treatments the quality of life decrement for anxiety about injections.

Quality of life for the bilateral health states

In addition to bilateral BCVA data, FAME collected NEI-VFQ-25 data at baseline, month 24 and month 36. The company used the NEI-VFQ-25 algorithm developed by Rentz et al[2 60] to estimate quality of life values for bilateral BCVA health states within FAME. Rentz et al was a time trade off study among 607 respondents who valued 8 NEI-VFQ-25 visions related health states out of the possible 15,625, scored using a subsample of 6 of the 25 elements. The company estimated bilateral quality of life values were pooled across the time points by inverse variance weighting. These values

[2] Sponsored by Allergan

were estimated across all FAME patients including those in the 0.5µg fluocinolone arm, with baseline observations [redacted] and the following total number of observations for the bilateral health states.

Table 12. Company VFQ-25 number of observations

		Better Seeing Eye							
		HS1	HS2	HS3	HS4	HS5	HS6	HS7	HS8
Worse Seeing Eye	HS1	[redacted]							
	HS2	[redacted]	[redacted]						
	HS3	[redacted]	[redacted]	[redacted]					
	HS4	[redacted]	[redacted]	[redacted]	[redacted]				
	HS5	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
	HS6	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]		
	HS7	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
	HS8	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

The raw NEI-VFQ-25 data with the Rentz algorithm applied and then pooled across time points resulted in the following.

Table 13. Company VFQ-25 QoL values

		Better Seeing Eye							
		HS1	HS2	HS3	HS4	HS5	HS6	HS7	HS8
Worse Seeing Eye	HS1	[redacted]							
	HS2	[redacted]	[redacted]						
	HS3	[redacted]	[redacted]	[redacted]					
	HS4	[redacted]	[redacted]	[redacted]	[redacted]				
	HS5	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
	HS6	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]		
	HS7	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	
	HS8	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Given anomalous values which are higher than another health state which would be anticipated to be at minimum no worse, the company smoothed the values by further averaging as below.

Table 14. Company VFQ-25 QoL values averaged over anomalous values

		Better Seeing Eye							
		HS1	HS2	HS3	HS4	HS5	HS6	HS7	HS8
Worse Seeing Eye	HS1	0.914							
	HS2	0.914	0.867						
	HS3	0.914	0.849	0.820					
	HS4	0.914	0.826	0.785	0.747				
	HS5	0.914	0.800	0.759	0.716	0.701			
	HS6	0.914	0.786	0.714	0.687	0.683	0.669		
	HS7	0.721	0.684	0.684	0.684	0.680	0.634	0.516	
	HS8	0.688	0.684	0.684	0.684	0.680	0.603	0.481	0.484

These can be graphed as below.

[redacted]

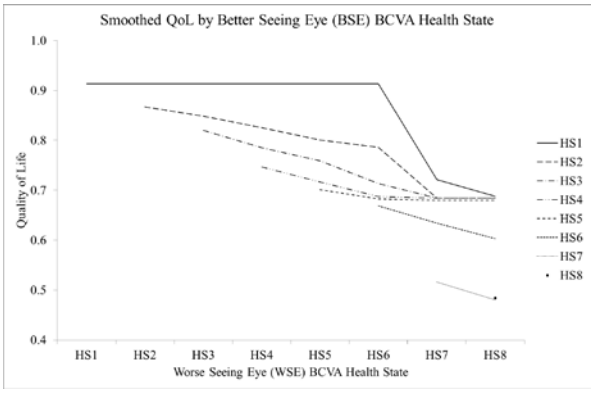


Figure 7. Weighted QoL values and their subsequent smoothing: Rentz et al[60]

In the above (Figure 7), provided the better seeing eye (BSE) is in the best BCVA health state, changes to the BCVA of the worse seeing eye (WSE) has little effect on quality of life until the WSE falls to quite a low BCVA. This is not unreasonable and mirrors the results of Brown[67] as reviewed later in the ERG critique.

Perhaps more striking and more difficult to provide an intuitive account of is that the quality of life values for the better seeing eyes in health states HS1 through to HS5 converge as the worse seeing eye falls into the lower health states.

Blindness, both eyes in either HS7 or HS8, causes a considerable reduction in quality of life.

The company also provides two scenario analyses which apply the quality of life values derived from Brazier et al[61] for bilateral health states based upon (A) the EQ-5D and (B) the VFQ-UI. Brazier et al[3] used data from 1,320 patients with DMO in the VISTA/VIVID trials. Quality of life data was collected using both the NEI-VFQ-25 and the EQ-5D. The EQ-5D values were transformed to quality of life values using the UK social tariff, while the NEI-VFQ-25 values were transformed to quality of life values using the Rentz et al algorithm. These quality of life values were then regressed on the VIVID/VISTA logs of BCVA in the BSE, BCVA in the WSE and in some models on the product of these two elements to provide an interaction terms, as well as various other patient characteristics. This provides mapping functions from bilateral BCVA health states to quality of life values without a requirement for NEI-VFQ-25 data or EQ-5D data. Brazier et al note that the mapping function for the NEI-VFQ-25 provides a better fit to the quality of life values derived from the Rentz et al algorithm than the EQ-5D mapping function to the quality of life values derived from the UK social tariff. But in itself this does not address whether the quality of life values from the Rentz et al algorithm for the NEI-VFQ-25 responses results in superior quality of life estimates than those derived from applying the UK social tariff to the EQ-5D data.

Brazier et al present a number of models and it is unclear which the company has used. The company values derived from Brazier et al are graphed below (Figure 8).

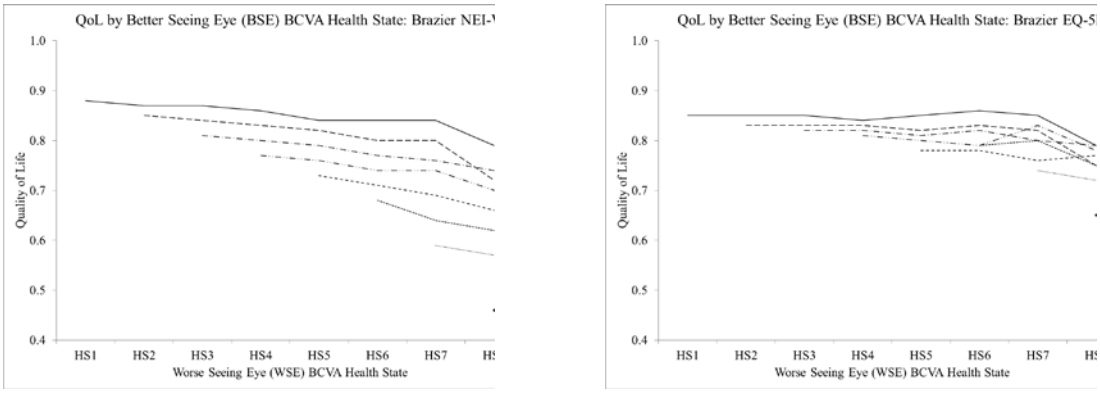


Figure 8. Company derived QoL values from Brazier et al[61]

The QoL values the company derives from the Brazier et al NEI-VFQ algorithm shows a somewhat stronger effect from changes in both the BSE BCVA and the WSE BCVA than those derived from Brazier et al's EQ-5D algorithm, and there are also fewer anomalous values.

[3] Sponsored by Bayer.

Quality of life decrements of adverse events

Cataract surgery and adverse events are assumed to have quality of life decrements that persist for 3 months. The values are reportedly mainly drawn from the ERG report of TA346[24], with the disutility for cataract surgery being taken from a dexamethasone STA.

Table 15. Adverse event quality of life decrements and QALYs lost per event

Event	QoL	Months	QALY
Cataract surgery	-0.003	3	-0.001
Post-cataract surgery comps.	-0.044	3	-0.011
Cataract surgery total			-0.001
IOP	..	..	..
Retinal detachment	-0.130	3	-0.033
Endophthalmitis	..	..	..
Vitreous haemorrhage	-0.020	3	-0.005
Glaucoma procedure	..	..	..

As these have relatively little impact upon results, as shown in the summary below, due to time constraints they have not been examined by the ERG.

Quality of life for injection anxiety

The company includes a quality of life decrement of -0.071 for an average of 2.5 days prior to each anti-VEGF injection to give a decrement of -0.0005 QALYs per injection. As shown in the results summary below, this has very little impact and has not been examined by the ERG. Its inclusion or exclusion does not affect the cost effectiveness estimate.

4.2.8 Resources and costs

Study Eye: Fluocinolone drug costs and administration costs

The fluocinolone implant cost at list prices is £5,500. All results within the company submission and this document are based upon the with PAS fluocinolone price of [redacted].

The company assumes one fluocinolone implant at baseline, and an additional one implant at the end of year 3 among those being retreated. A £108 administration cost is applied in the model.

Study Eye: Anti-VEGF drug costs and administration costs

All results within the company submission and this document are based upon the list price of ranibizumab of £551. The ERG supplies an additional confidential appendix that applies both the fluocinolone PAS and the ranibizumab PAS.

The company submits data from ICE-UK which shows that among the 124 patients who received any drug therapy in the year before the fluocinolone implant, this had fallen to 37 patients (30%) in the year after the fluocinolone implant. Furthermore, the average number of treatments among those receiving any drug therapy fell from 3.98 to 3.02 as outlined below in Table 16.

Table 16. ICE-UK Anti-VEGF and aflibercept treatments pre and post fluocinolone implant

	Before implant		After implant		Drug cost
	N	n	N	n	..
Ranibizumab	114	4.1	33	2.9	£551
Bevacizumab	9	2.4	0	..	£243
Aflibercept	1	5.0	4	4.0	£816
Total/Mean	124	3.98	37	3.02*	
				0.90**	

*Among those receiving any drug treatment post implant

**Among those who received any drug treatment pre implant

Rather than apply the post-implant costs in the fluocinolone arm and the pre-implant costs of these treatments in the usual care arm, the company nets out these figures to derive a net number of administrations 3.98 - 0.90 = 3.08 and applies this in the usual care arm.

Treatmetn discontinuations prior to the end of year 6 are not modelled in the usual care arm. The company assumes that prior to all discontinuing at the end of year 6, the number of drug administrations in the usual care arm falls linearly towards zero, resulting in an average annual number of injections of 1.61; i.e. roughly half the 3.08.

When coupled with the 28% receiving no drug treatment, 63% receiving ranibizumab and 9% receiving bevacizumab this results in an annual direct drug cost for the composite comparator of £547. Costs per administration of £108 for all treatments result in an annual administration cost of £161.

The reason for the above approach is not obvious. It may help avoid acknowledging more explicitly after their fluocinolone implant a substantial minority of patients continued to receive anti-VEGF treatments much as before. Perhaps more pertinently for the economics, it also avoids anti-VEGF monitoring costs among those with a fluocinolone implant.

Assuming all patient reduce their use of anti-VEGF as opposed to assuming some maintain their use while others discontinue also means higher ongoing anti-VEGF monitoring in the usual care arm.

Study Eye: Laser costs

ICE-UK data is used which suggests in the year after fluocinolone implant [redacted] patients received some laser treatment, with an average of 1.2 laser administrations in the year following fluocinolone implant.

In the fluocinolone arm it is assumed that there are no laser administrations.

In the usual care arm it is assumed that all the 28% of patients in the composite comparator who are not receiving anti-VEGF require 1.2 laser administrations annually. These are costed at the common £108 cost per administrations.

Study Eye: Treatment cessation and off-treatment drug costs and administration costs

Those who discontinue fluocinolone at the end of year 3 are in effect for costing purposes assumed to cross over to usual care and to have the same annual drug and administration costs as those on treatment with the composite comparator; £547 and £161 respectively.

Study eye: Ongoing monitoring costs

The following resource use for the outpatient (OP), optical coherence topography (OCT) and fluorescein angiography (FA), with unit costs of £91, £86 and £138 respectively, is applied for fluocinolone and each of the individual comparators[4].

[4] There is a small additional £2 cost associated with monitoring of previous laser treatments for those not receiving usual care / laser, but this is incidental and can be ignored for current purposes.

Table 17. Monitoring resource use

	Year 1			Year 2			Year 3		
	OP	OCT	FA	OP	OCT	FA	OP	OCT	FA
Fluocinolone	5.6	3.3	1.0	2.8	1.7	1.0	3.0	3.0	1.0
Usual care / laser	4.0	4.0	1.0	4.0	4.0	0.0	2.6	2.6	0.0
Ranibizumab	12.0	12.0	1.0	6.3	6.3	0.0	4.0	4.0	0.0
Bevacizumab	12.0	12.0	1.0	6.3	6.3	0.0	4.0	4.0	0.0

This results in the following annual monitoring costs for each of the comparators, and for the usual care arm when weighted by the proportion of each comparator within the composite comparator.

Table 18. Monitoring costs

	Year 1	Year 2	Year 3	Average
Fluocinolone	£929	£537	£667	n.a.
Usual care / laser	£843	£706	£459	
Ranibizumab	£2,255	£1,111	£706	
Bevacizumab	£2,255	£1,111	£706	
Composite comparator	£1,855	£997	£636	£1,163

In the fluocinolone arm the model applies the £929, £537 and £636 annual monitoring costs to years 1, 2 and 3 respectively, and carries forward the £636 year 3 monitoring cost to years 4, 5 and 6.

In the usual care arm the model averages the £1,855, £997 and £636 annual monitoring costs of years 1, 2 and 3 to arrive at an average annual monitoring cost of £1,163. This £1,163 average is applied to years 1, 2 and 3. Rather than carry over the year 3 £636 cost, for years 4, 5 and 6 the model carries over the £1,163 average.
In both arms, those who cease treatment incur the £1,163 monitoring cost.

When reviewing the above it should be borne in mind that the comparator arm patients are not newly incident anti-VEGF patients but by definition have been receiving anti-VEGF for some time. The above anti-VEGF monitoring schedule and its rapid tailing off in terms of annual visits and annual costs is for newly incident anti-VEGF patients.

It should also be borne in mind that in ICE-UK 30% of patients continued to receive anti-VEGF treatment after their fluocinolone implant and at much the same level as prior to their fluocinolone

implant. But due to the model construction these fluocionolone arm patients are not modelled as incurring any of the higher anti-VEGF monitoring costs.

At the end of year 6 all patients are assumed to cease treatment in their study eye and also to incur no further monitoring costs. It may be unlikely that all monitoring will cease after year 6 but provided that there is no difference in survival between the arms these costs would net out between the arms. But if fluocinolone results in fewer falling into blindness so results in a better survival, not including ongoing care and monitoring costs after year 6 biases the model in favour of fluocinolone.

Fellow Eye: Drug costs

The fellow eye drug costs largely net out between the arms, so the ERG does not dwell on this aspect. In both arms, those with bilateral DMO are assumed to receive the same balance of treatments as in the study eye usual care arm but with an annual number of anti-VEGF injections of 3.23 as drawn from ICE-UK.

This does not particularly distinguish between those developing bilateral DMO during the model and those with bilateral DMO at baseline, but as already noted these costs largely net out between the arms.

Fellow Eye: Administration costs

The administration cost for a fellow eye is based upon much the same calculation as for the study eye, but assuming the 3.12 anti-VEGF injections, resulting in an annual administration cost of £278. If there is the possibility of bilateral treatment at a single outpatient visit this will overestimate the administration costs in the usual care arm.

Fellow Eye and bilateral monitoring costs

Fellow eyes with bilateral DMO are assumed to receive usual care. The cost of monitoring usual care in a fellow eye is assumed to be 10% of the annual average cost of monitoring usual care in a study eye: £116. This is due to the possibility of bilateral monitoring at the same OP appointment. This results in total monitoring costs among those who have bilateral DMO as below in Table 19.

Table 19. Total annual monitoring costs [5] for bilateral patients

Arm	Fluocinolone arm			Usual care arm		
Eye	Study Eye	Fellow Eye	Bilateral	Study Eye	Fellow Eye	Bilateral
Treatment	Fluocinolone	Compo. UC	Total	Compo. UC	Compo. UC	Total
Year 1	£929	£116	£1,045	£1,163	£116	£1,279
Year 2	£537	£116	£653	£1,163	£116	£1,279
Year 3, 4, 5 & 6	£667	£116	£783	£1,163	£116	£1,279
Compo. UC: Composite Usual Care comparator of 28% laser, 63% ranibizumab and 9% bevacizumab						

The cost of monitoring bilateral patients in the fluocinolone arm varies between roughly 50% and 80% that of the usual care arm, and is only around 60% for the majority of the 6 years that these costs are incurred in the model.

Other event costs

The other events within the model are costed as follows (Table 20), each outpatient (OP) visit costing £86.

Table 20. Other event costs

	Description	Cost	OP	Total
Cataract removal	BZ34A-C: Phacoemulsification day case	£895	3	£1,153
Endophthalmitis	BZ86B: Non-elective long stay	£1,659	6	£2,175
Retinal detachment	BZ87A: Non-elective short stay	£730	4	£1,067
Vitreous haemorrhage	BZ87A: Non-elective short stay	£549	2	£894
IOP medication	Medication	£66	6	£611
IOP / glaucoma surgery	Code not stated: Day case	£581	6	£1,126

The codes stated above reflect the company submission. It is unclear why the reported BZ87A costs differ in the above.

[5] Again, ignoring the inconsequential additions for laser monitoring.

Costs after year 6

From year 6 the only costs that are applied are the costs arising from cataract surgery and the costs of blindness.

4.2.9 Cost effectiveness results

The company deterministic base case estimates the following.

Table 21. Company base case deterministic results: Costs[6]

	Study Eye				Fellow Eye		
	Tx	Admin.	Monitor.	AEs		Blind	Total
Fluocinolone	£5,256	£1,687	£4,589	£3,289	£5,886	£1,825	£22,532
Usual care	£2,950	£2,175	£6,302	£2,739	£5,886	£2,811	£22,863
Net	£2,306	-£487	-£1,713	£550	£0	-£987	-£330

Higher treatment costs are offset mainly by lower ongoing monitoring costs. But savings in the costs of blindness and costs of administration also help offset the higher treatment costs. A net saving of £330 results.

Table 22. Company base case discounted life years survival and QALYs

		QALYs			
	LYs	HSs	AEs	Anxiety	Total
Fluocinolone	11.480	8.493	-0.010	-0.004	8.479
Usual care	11.436	8.261	-0.012	-0.006	8.244
Net	0.043	0.232	0.002	0.002	0.236

The gains in discounted survival of 0.043 life years provide some of the net QALY gains, with adverse events and anxiety providing some additional net QALY gains. But the large majority of the net gain of 0.236 QALYs arises from the improvement in the BCVA distribution of the study eyes.

Given the cost savings and net QALY gains fluocinolone is estimated to dominate usual care.

[6] The totals reported here are the totals of the individual items reported to their left. These totals differ slightly from those reported in Table B3.42 but the net amounts correspond.

The probabilistic modelling yields similar central estimates: a net saving of £301 and a net gain of 0.25 QALYs, the following CEAC and a likelihood of fluocinolone being the most cost effective at a willingness to pay of £20k per QALY of 93%.

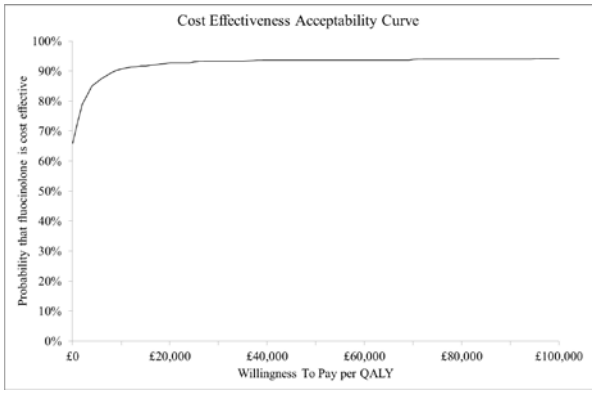


Figure 9. Company base case CEAC

4.2.10 Sensitivity analyses

Sensitivity analyses

The company performs a large number of univariate scenario analyses, these typically varying the parameters through their 95% confidence interval where these are available; i.e. ±1.96 standard errors. The exception to this is the treatment effect for the probability of cataract removal among those with cataract, which is only varied by ±1 standard errors. At clarification the company stated that it adopted this approach because it viewed the 95% confidence interval as being too wide to be plausible, with it resulting in estimates of cataract surgery among those with cataract being lower in the fluocinolone arm than in the standard care arm. The 10 variables that are explored that resulted in the largest spread of ICERs are summarised below, and are presented as a tornado diagram in Figure B3.4 of the company submission.

Table 23. Company sensitivity analyses: 10 most influential variables and their ICERs

	Lower	Upper
Base case	-£1,400 (Dominant)	
RANI - OP visits	£1,331	-£4,131
RANI - OCT visits	£1,173	-£3,973
FLUO - OP visits	-£3,617	£818
RANI - Drug Cost	£2,208	-£1,400
FLUO - OCT visits	-£2,939	£140
Usual care Tx proportions	-£1	-£2,520
Baseline age: no cataract subgroup	-£2,527	-£55
Baseline age: with cataract subgroup	-£2,527	-£55
Usual care - N drug admins	-£304	-£2,496
Discount rate	-£2,477	-£483

At clarification the company supplied additional sensitivity analyses which showed that the lower confidence interval for the treatment effect for the probability of cataract caused fluocinolone to no longer be cost saving and to be associated with an ICER of £3,681 per QALY.

Scenario analyses

The company also presents a number of scenario analyses.

- Scenario 1: Assuming the same balance of treatments for usual care as the base case for costing but applying the FAME sham arm transition probabilities resulted in fluocinolone costing a net £523 and resulting in a net gain of 0.036 QALYs, hence an ICER of £14,753 per QALY.
- Scenario 2: Assuming only laser costs for usual care and applying the base case clinical effectiveness assumption of no change in BCVA for usual care resulted in fluocinolone costing a net £3,763 and resulting in a net gain of 0.238 QALYs, hence an ICER of £15,842 per QALY.
- Scenario 3: Labelled watchful waiting which the company describes as fluocinolone treatment being possible in the usual care arm after cataract surgery. The ERG has not managed to replicate this scenario. The submission reports a net saving of 27p, a net gain of 0.145 QALYs hence dominance for using fluocinolone at baseline rather than subsequent to cataract removal.
- Scenario 4: Fluocinolone use in the fellow eye is very briefly reported as resulting in dominance. Note that this element of the electronic model is not used for anything else within the submission and the ERG has not parsed or rebuilt it. A cursory examination of it suggests

some peculiar feedback loops, and it can also be noted that accounting for newly incident bilateral disease in each cycle over the 30 year model time horizon is challenging.

- Scenarios 5 and 6: Using Brazier et al[61] quality of life values is very briefly reported as still resulting in dominance. ERG analyses suggest that the Brazier NEI-VFQ values cause the net gain to fall slightly from 0.236 QALYs to 0.217 QALYs, and that the Brazier EQ-5D values cause it to more than halve to 0.109 QALYs.
- Scenario 7: Applying the minimum difference between the eyes, which the ERG takes to mean assuming that the distribution of the BCVA of study eyes is not independent of the BCVA of the fellow eye, is very briefly reported as resulting in dominance. An ERG analysis suggest the net saving increases to £657 but the net gain falls to 0.229 QALYs.

Subgroup analyses

The company presents subgroup analyses.

- For those without cataract at baseline a net saving of £81 and a net gain of 0.190 QALYs, so dominance for fluocinolone.
- For those with cataract at baseline a net saving of £579 and a net gain of 0.282 QALYs, so dominance for fluocinolone.

4.2.11 Model validation and face validity check

The company has provided the following data at clarification for the model estiamtes of BCVA over the first 36 months for fluocinolone and for sham, and how this compares with those of FAME. Note that within this sham is not the same as the company usual care arm. The ERG has added the model estimates of BCVA for the company usual care arm, this effectively tracking the horizontal axis (Figure 10).

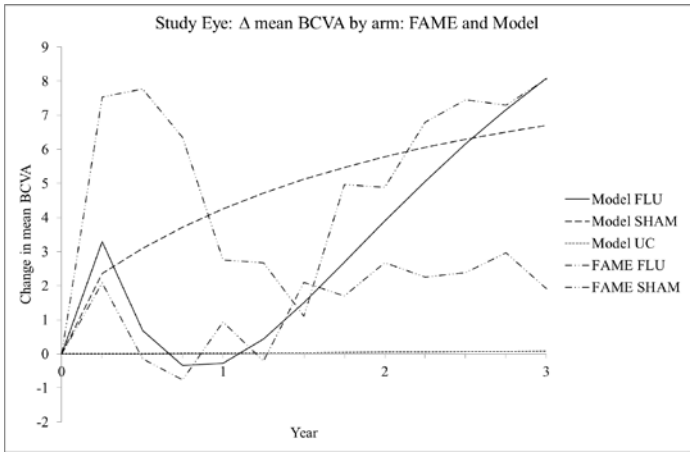


Figure 10. Model validation versus FAME trial

For fluocinolone the model appears not to particularly match the FAME trial until around 18 months, after which the fit improves. For sham the model appears to predict somewhat larger gains in BCVA than were observed during FAME.

The modelled changes in mean BCVA over the 30 year time horizon are presented below in Figure 11.

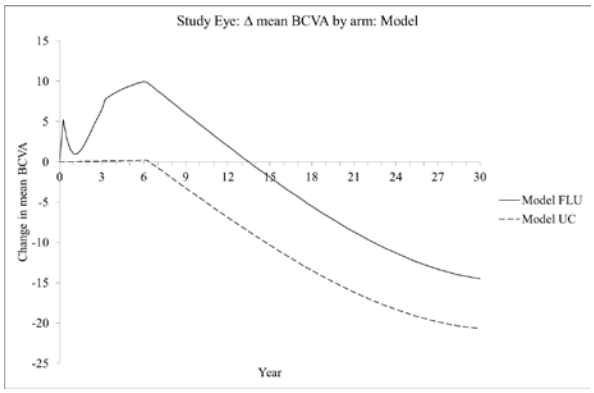


Figure 11. Modelled mean BCVA over 30 year horizon

The absolute net gain of around 10 letters at the end of year 6, when all treatment ceases, is modelled as being broadly maintained for the rest of the time horizon.

Between years 3 and 6 the model anticipates further improvements in the mean BCVA in the fluocinolone arm. The reason these gains tail off is because 64% of patients stop fluocinolone and receiving usual care. Among the 36% who receive a 2nd fluocinolone implant the mean BCVA continues the steep upward trend as shown below (Figure 12).

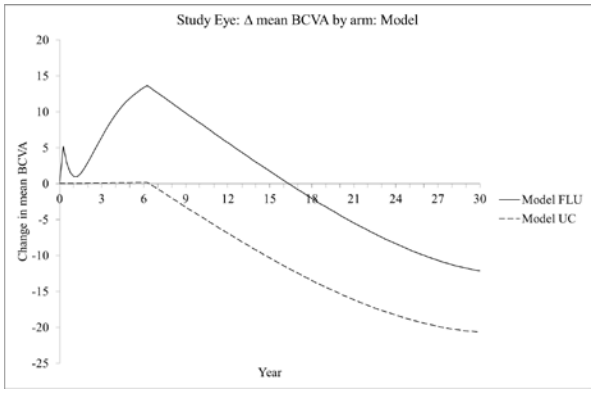


Figure 12. Modelled mean BCVA over 30 year horizon: 100% fluocinolone retreatment

Fellow eye modelling

Modelling of the fellow eyes cohort flow is quite involved, and the formulae involved differ in the fluocinolone cohort flow from the formulae in the usual care cohort flow. Despite this the proportion surviving and the mean BCVA in the fellow eyes are virtually identical between the arms, the latter being presented below in Figure 13.

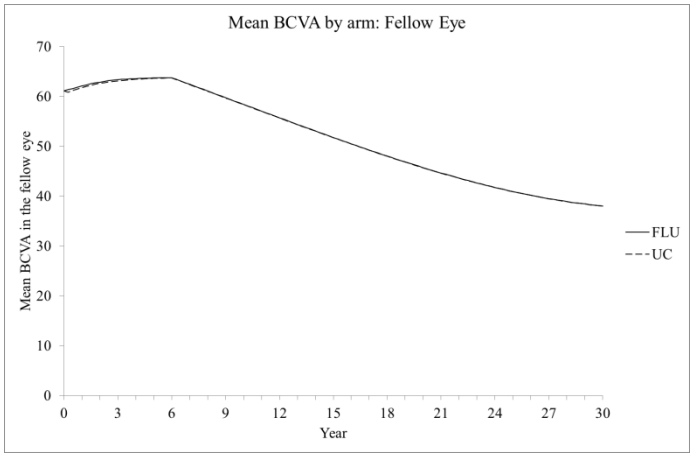
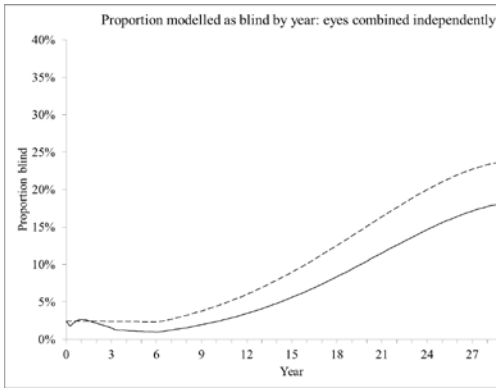


Figure 13. Mean BCVA in the fellow eye

The initial mean BCVA in the fluocinolone fellow eyes is slightly better than that of the usual care fellow eyes. But the difference is minimal, and thereafter the mean BCVAs converge. In the opinion of the ERG given the company assumptions the cohort flow for the fellow eyes should be identical between the arms. Due to the greater complexity of the formulae in the modelling of the fellow eyes in the fluocinolone arm, the ERG has revised the cohort flow for the fellow eyes to be that of the usual care arm in both arms. This has minimal effect upon the model outputs.

Proportion with and going blind

The other output of the model that requires examination, given the large quality of life impact and costs of being blind, is the proportion modelled as falling into blindness. This can be presented for the modelling that combines the distribution of the study eyes and the distribution of the fellow eyes independently, and that combines them as being interdependent (see Figure 14 and Figure 15).



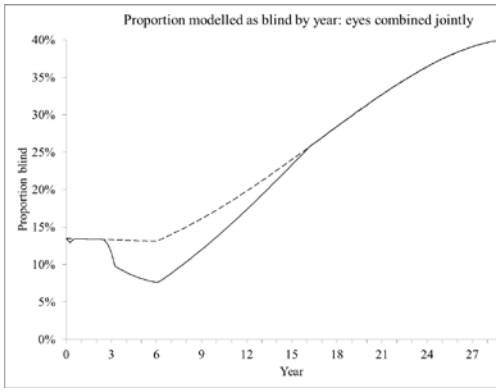
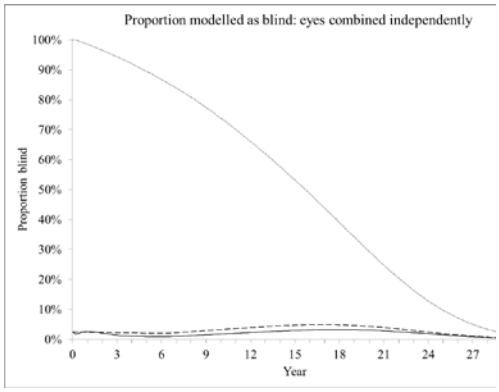


Figure 14. Proportion modelled as blind among those surviving



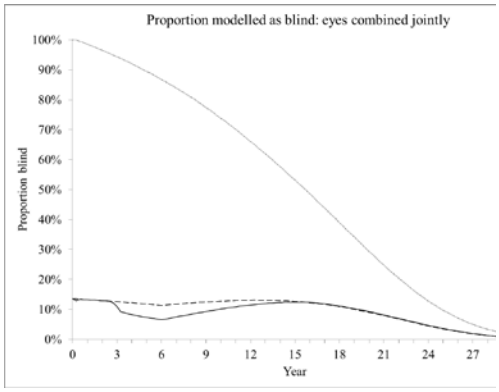


Figure 15. Patients modelled as blind against overall survival

If the distributions of eyes are combined independently only a very small proportion of patients are modelled as being blind at baseline. Over the next 3 years there are some small gains from fluocinolone, with these expanding a little more over the next 3 years due to the additional BCVA gains among the 36% receiving a 2nd fluocinolone implant. But the gains in the proportion who are blind occur mainly during the period of extrapolation, becoming ever greater as time passes and the extrapolation increases.

If the distributions of eyes are combined jointly, in essence largely assuming that the fellow eye is the better seeing eye, the proportion in blindness at baseline is modelled as being much higher. By year 6 the maximum net difference between the arms is reached, with this then waning until by year 18 it seems that it is purely changes in the distribution of the fellow eye that are determining the proportion in blindness in both arms.

4.3 ERG cross check and critique

4.3.1 Base case results and model structure

The company model is quite involved, with 80 health states for the study eye and 200 health states for the fellow eye. But cross checking the modelling of the fellow eye can be simplified because the company does not concentrate upon blilateral treatment with fluocinolone.

The ERG has not cross checked the unused portion of the modelling of the fellow eye which does permits discontinuations from usual care and treatment with fluocinolone. But the ERG would like to stress that this aspect of the model may not be reliable. There are strange feedback loops within it, and it is not obvious that it properly accounts for newly incident bilateral DMO as the model progresses.

Within the company model there are some relatively minor modelling errors. For instance, in the fluocinolone arm it is possible for some to develop cataract and have these cataracts removed within a single cycle whereas in the usual care arm the development of cataracts and having them removed takes a minimum of two cycles. In the fluocinolone arm there are also some instances of the FAME 1st quarter transition probabilities being applied in the period after the 1st quarter. There are also minor discrepancies between the modelling of the fellow eye in the fluocinolone arm and in the usual care arm, but these are easily bypassed by simply applying the same fellow eye cohort flow in the fluocinolone arm as that modelled in the usual care arm. Other than these errors the ERG finds the model implementation to be well aligned with the intended model structure.

The ERG has rebuilt the company deterministic model using a scenario that tests the structure of the model somewhat more extensively than would be provided by a rebuild of the company base case. This rebuild assumes that:

- laser patients receive the FAME sham TPMs,
- anti-VEGF patients have the 1.54 odds ratio of improvement applied to the FAME sham TPMs,
- 10% discontinue per cycle in the usual care arm, and
- 20% of fellow eyes are pseudophakic at baseline.

These assumptions are made purely to test the model structure and result in the following (see Table 24).

Table 24. ERG cross check model rebuild

	Company model		ERG rebuild	
	Costs	QALYs	Costs	QALYs
Fluocinolone	£21,856	8.510	£21,863	8.412
Usual care	£21,389	8.522	£21,471	8.421
net	£467	-0.013	£392	-0.008

The absolute amounts show quite good correspondence but due to the limited differences between the arms the net costs show larger proportionate differences. The difference between the net costs mainly arises due to some differences in the costs of adverse events and the costs of blindness.

4.3.2 Correspondence between written submission and cited sources

Costs of blindness

Blindness within the model requires that both the study eye and the fellow eye fall into either HS7 or HS8. This proportion of patients have an annual £19,795 cost of blindness applied, derived from Meades & Hyde.[68]

The costs of the individual elements of this are presented in table B3.39, together with the proportion who are assumed to incur these costs. This in turn relies upon an annual cost of depression of £2,457 derived by the company from McCrone et al[69] coupled with unit costs largely derived from the PSSRU. The weighted sum of the elements of table B3.39 is not £19,795 but is rather £12,369.

Table B3.39 also does not adjust the cost of residential care for the 30% that Meades & Hyde estimate self-fund. Applying this within Table B3.39 results in an annual cost of £9,411.

This cost can be compared with that of £6,298 that was applied by the company in its 2012 submission for DMO, which when inflated by the HCSC increase of 9% to be in 2017/18 prices results in an annual cost of £6,835. It can also be compared with the £7,429 estimated by the company from Meades & Hyde, augmented with the costs of depression from McCrone et al [69], in its 2016 submission for BRVO.

The annual cost of residential care is given as £57,616 in table B3.38 but as £32,864 in table B3.39, with PSSRU 2017 being cited for both. The higher figure corresponds with local authority provided residential care, the lower figure with private sector residential care. The Competition and Marketing

Authority (CMA) 2017 analysis of care home provision[70] estimates that "*around 95% of beds are provided by the independent sector … LAs generally commission care services from independent care providers. We estimate that the average cost for a self-funder in 2016 was £846 per week … while LAs on average paid £621 per week*" and "*41% of residents in care homes fund themselves (self-funders) and 49% receive LA-funding (around a quarter of these pay top-ups). Even for those receiving LA-funding, nearly all income, such as pensions, is offset against state contributions*".

To the ERG the above argues for applying a 41% self-funding proportion, and to assume 95% of residential care is provided in the private sector. There may also be some double counting of the cost of residential care due to it appearing in both the cost of depression, which is a constituent of the cost of blindness, and separately in the cost of blindness. These assumptions result in an annual cost of blindness estimate of £8,107 if 41% self-fund. The ERG will conduct a sensitivity analysis that assumes that only 30% self-fund their residential care, which results in an annual cost of £9,276. These estimates are much better aligned with those of previous assessments, including those of the company, than the estimate of the current submission.

These may still be overestimates if top-up considerations mean that 49% of those receiving LA-funding do not receive the full £621 per week, but the wording of the CMA report is ambiguous about this aspect.

4.3.3 Correspondence between the written submission and electronic model

In the opinion of the ERG the written submission is a poor account of the electronic model structure and its inputs, but what is written is correct as far as it goes.

During the ERG model rebuild some minor model structure issues were identified, but these are not key drivers of results. They are revised[7] in the ERG exploratory analyses of section 4.4 below, coupled with a sensitivity analysis that does not apply these changes.

4.3.4 ERG commentary on model structure, assumptions and data inputs

FAME absolute and net treatment effect and the modelling

[7] These revisions are documented in the *ERG* worksheet of the ERG revised model, with full cell referencing.

Model validation data supplied at clarification gives the following evolutions of the FAME study eyes' average BCVA and average change from baseline (see Figure 16).

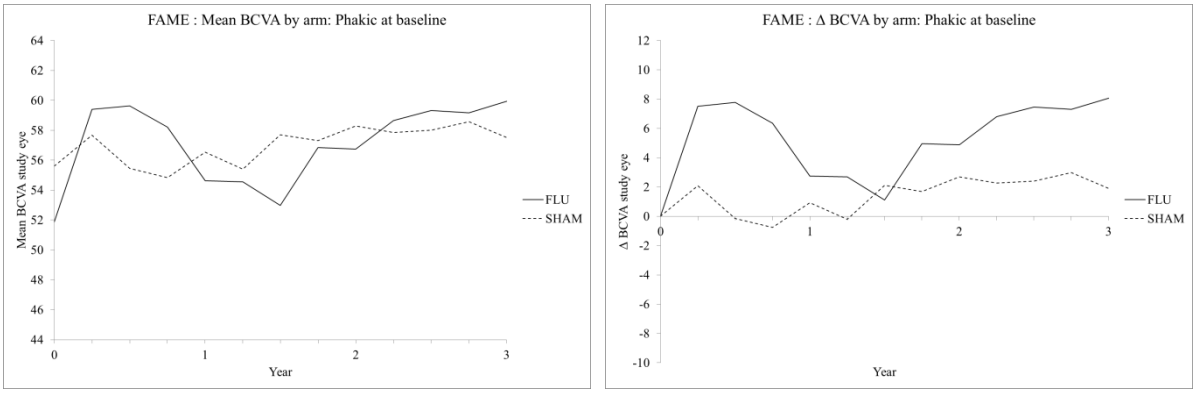


Figure 16. FAME mean BCVA and Δ BCVA among the phakic by arm

There may be a concern that for the phakic there was a reasonable difference in the mean BCVA at baseline between the arms, which might give rise to a different propensity to benefit. But for the economic modelling the main concern is how the FAME data will relate to the patient group under consideration.

The FAME data show the sham arm having an absolute gain of +2 letters at 36 months, implying a degree of natural recovery, perhaps due to improved glycaemic control. In the fluocinolone arm, the absolute gain of +8 letters at 36 months corresponds to a net treatment effect of +6 letters.

In the model, for the fluocinolone arm the FAME effect inclusive of natural recovery is applied. But for the usual care arm the FAME effect excluding natural recovery is applied. The usual care arm is assumed to have no change in BCVA.

To the ERG this seems unreasonable. Most dramatically, the model assumes a 50% prevalence of cataract at baseline and that in the fluocinolone arm the removal of cataracts improves vision. But while the removal of cataracts in the usual care arm incurs costs there is no improvement in vision.

A key question is whether the fluocinolone arm should be modelled as receiving the full benefits seen during FAME inclusive of natural recovery while the usual care arm should be assumed to exhibit no natural recovery. There is the argument that these are patients who have failed to respond to current treatments among whom any natural recovery effects will by now be exhausted, and so there will be no subsequent natural recovery. But if this is the case, to the ERG this argues for subtracting the natural recovery effect from the fluocinolone arm and only applying the net treatment effect observed from fluocinolone over sham during FAME. The closest that the model structure permits to this is

applying the FAME fluocinolone TPMs in the fluocinolone arm and the FAME sham TPMs in the usual care arm.

FAME treatment effects and the anti-VEGFs

The ERG argues above that the only the net treatment effect observed in FAME should be applied; i.e. the sham treatment effect should in a sense be subtracted from the fluocinolone treatment effect to give a net effect. It does not directly represent what a hypothetical RCT of fluocinolone against no drug treatment might look like in the patient group of interest. But if no further natural recovery can be anticipated, no further drug treatment might result in a flat BCVA much as assumed by the company with the net fluocinolone effect being added to this, as crudely hypothesised in the left hand figure (#1) below (see Figure 17). During the decision problem meeting the company hypothesised that no further drug treatment in the patient group of interest would see BCVA worsening. The fluocinolone net effect could be added to this, as more controversially hypothesised in the right hand figure (#2) below. But both figures maintain the net treatment effect at 36 months.

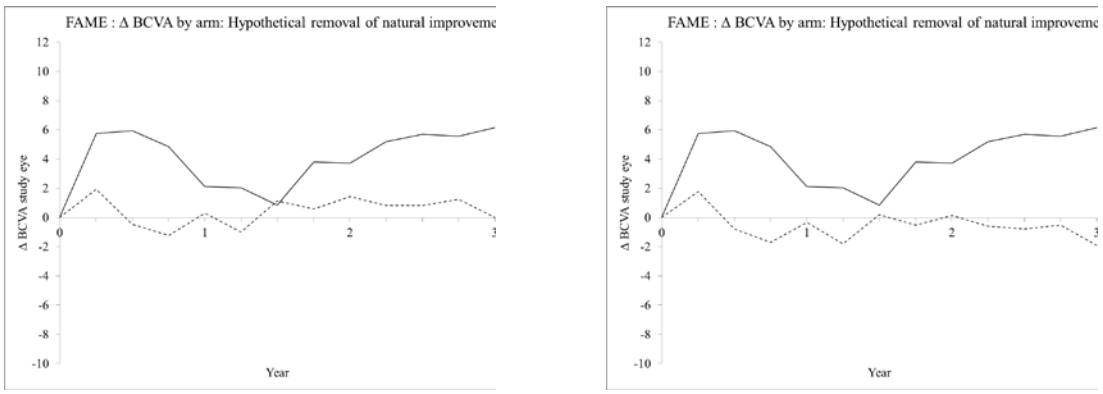


Figure 17. Hypothetical changes from baseline that maintain the FAME net treatment effect

At the decision problem meeting the company hypothesised that clinicians maintained anti-VEGFs to avoid a loss of vision that would occur from their withdrawal. This seems akin to the right hand figure with anti-VEGFs following the horizontal axis. In other words, the anti-VEGF would be superior to sham but inferior to fluocinolone.

In short, for the patient group of interest we do not know what an RCT would show for:

- The absolute change between baseline and 36 months in BCVA for sham.
- The net change between baseline and 36 months in BCVA for fluocinolone compared to sham.
- The net change between baseline and 36 months in BCVA for anti-VEGF compared to sham.

- The net change between baseline and 36 months in BCVA for fluocinolone compared to anti-VEGF.

Ethical or patient recruitment considerations might preclude a sham arm in any RCT, but there is no obvious bar to an RCT with a fluocinolone arm and an anti-VEGF arm.

Proportion with pre-existing cataracts

During the reasonably prolonged exchanges for clarification it transpired that the large majority proportion of patients with cataract at baseline in FAME that was originally reported was based upon fundus photography. Fundus photography estimated that 79% of those in the FAME fluocinolone arm had cataract at baseline, though results for the sham arm are less clear. All the economics is apparently based upon a different definition of cataracts as "*spontaneous reports of cataract as adverse events and in medical history*". Applying this definition causes the large majority of patients to be without cataract at baseline in FAME and only around 4% to be with cataract at baseline.

The ERG does not know what definition of cataract was used during Retro-IDEAL which gave rise to the company estimate of a 50% baseline prevalence of cataract.

Company estimate for anti-VEGF treatment effect relative to sham

The company submission is particularly opaque about how it derives and applies a treatment effect for the anti-VEGFs. The company uses data from the ranibizumab RISE and RIDE trials which suggests a mean gain at 24 months compared to baseline of 12.0 letters for ranibizumab and 2.5 letters for sham, hence a net treatment effect of 9.5 letters. The company then arbitrarily assumed that half of patients' fellow eyes would be insufficiently responsive to anti-VEGF and so would experience no net treatment effect, resulting in a halving of the overall net treatment effect to only 4.75 letters.

Within a model the ERG has not had access to, the company apparently experimented applying an odds ratio to the FAME derived sham probabilities of gaining letters and found that an odds ratio of 1.54 resulted in a gain of 4.75 letters. It is not know which FAME derived sham probabilities this was applied to when forming the 1.54 estimate. It can also be noted that the company does not estimate any odds ratio of losing letters as this is not possible using the stated company method.

For reasons that are not clear, as far as the ERG can ascertain an anti-VEGF treatment effect over sham is only ever applied in the fellow eye.

ERG review of the underlying RISE and RIDE data suggests the following odds ratio based upon Brown et al[22], and the following odds ratios if the company assumption that half of ranibizumab patients experience the sham effect rather than the observed ranibizumab effect (see Table 25).

Table 25. ERG anti-VEGF odds ratios of gaining and losing letters

	RISE&RIDE	Adjusted for 50%
OR: gaining letters	2.72	1.66
OR: losing letters	0.38	0.67

But as reviewed in the clinical effectiveness section, this data relates to the changes from baseline in RISE and RIDE. The most appropriate data to use for chronic patients to estimate the effect of anti-VEGFs among those not responding to non-pharmacological treatment may be the odds ratios of gaining and losing letters among sham crossing over to ranibizumab at 24 months. ERG note that combining the RISE and RIDE trial suggests a gain in the sham arm of around 2.5 letters at 24 months and around 4.5 letters at 36 months. This net gain of only 2 letters is somewhat less than the 4.75 letters that the company bases its construction of the 1.54 odds ratio. In the light of this the ERG will explore reducing the odds ratio pro rata to 1.23, and also the original company odds ratio of 1.54. But given the uncertainty around quite what the estimate should be these will only be explored in scenario analyses and not the ERG revised base case.

Composite comparator and subgroups

The company analysed retrospective treatment histories in the 12 months prior to recruitment to ICE-UK to estimate that 28% of patients received no anti-VEGF, with the remainder being divided 63% ranibizumab and 9% bevacizumab. These are combined into a single composite comparator.

No account of the reasons for treatments prior to ICE-UK is given and there may be a concern that different subgroups of patients are being combined. For instance, an immediate question is how or whether the 28% receiving laser differ from the phakic subgroup considered during TA301, given that the model relies upon FAME clinical effectiveness data.

It is also not typical NICE practice to have a composite comparator. The ERG acknowledges the situation here differs slightly in that the modelling is of patients ceasing their current treatment as

opposed to there being an obvious range of comparator treatments which are being considered for the patient.

But to the ERG this still argues for the composite comparator to be unpicked, even if only to illustrate what is driving the model. The initial pairwise comparison of fluocinolone with each single treatment of the composite comparator is presented below in Table 26, the pairwise comparison with laser corresponding with the company Scenario 3.

Table 26. Pairwise comparisons between fluocinolone and the individual comparators

	Fluocinolone		Comparator		Net change		
Comparator arm	Costs	QALY	Costs	QALY	Costs	QALY	ICER
Laser	[redacted]	[redacted]	£11,054	8.239	£3,763	0.238	£15,842
Bevacizumab	[redacted]	[redacted]	£21,610	8.245	-£492	0.235	FLUO Dominant
Ranibizumab	[redacted]	[redacted]	£28,764	8.245	-£2,417	0.235	FLUO Dominant

The minor differences in the total QALYs for each of the comparators arises solely due to the differing adverse event profiles.

In the stacked comparison ordered by increasing cost, there are three fluocinolone arms due to those who discontinue fluocinolone treatment at the end of year 3 being assumed to receive the comparator treatment: FLUO-Laser, FLUO-BEVA and FLUO-RANI.

Table 27. Stacked comparison between the individual comparators

	Costs	QALY	ICER
Laser	£11,054	8.239	
FLUO-Laser	£14,817	8.477	£15,842
FLUO-BEVA	£21,118	8.480	£1.8mn
Bevacizumab	£21,610	8.245	Dominated
FLUO-RANI	£26,347	8.480	Dominated
Ranibizumab	£28,764	8.245	Dominated

If the argument is that discontinuing anti-VEGF cannot be or should not be sensibly combined with remaining on anti-VEGFs into a composite comparator because the patients are from obviously distinct patient subgroups, it seems sensible to consider the individual comparators in isolation. But if the patient group is coherent and the treatments can be combined, it seems valid to disaggregate the composite comparator into its constituent parts.

As would be expected given the somewhat lower price of bevacizumab and the assumed equivalence with ranibizumab, options with bevacizumab dominate the corresponding option with ranibizumab.

Due the assumption of clinical equivalence between the anti-VEGFs and that the choice between ranibizumab and bevacizumab may be more due to geography and individual clinicians it may be reasonable to pool these as a comparator, coupled with scenario analyses around the balance between ranibizumab and bevacizumab. As a consequence and in order to avoid presenting too many sets of analyses the ERG will present three sets of pairwise analyses with a full set of sensitivity analyses for each:

- fluocinolone against laser / no ongoing drug treatment,
- fluocinolone against anti-VEGFs, and
- fluocinolone against the company composite comparator.

The base cases of the first two bullets will also be combined into a stacked analysis.

The probability of discontinuing fluocinolone

The model assumes fluocinolone is never discontinued during the three year implant lifespan. It is possible to remove the implant, for example if there is a steep and unmanageable rise in IOP. There are a few reports of removal or re-positioning if the implant migrates to the front of the eye where it could cause corneal damage but that seems rare and it may only occur in eyes that have had vitrectomy and are pseudophakic, so does not seem not relevant to this STA.

The model assumes that a proportion of patients equal to that which achieved a minimum of 15 letters gain: 36/114 = 36%, receives a 2nd fluocinolone implant. The model applies this proportion equally across the four patient subsets. In this regard it can be noted that [redacted] who remained phakic at 36 months had achieved a 15 letter gain, while [redacted] who had become pseudophakic by month 36 had achieved a minimum 15 letter gain.

It is assumed that a 3rd fluocinolone implant is not possible and that 100% of fluocinolone patients discontinue at the end of the 1st 6 years of the model.

Clinical effect of 2nd fluocinolone implant

The company re-applies the fluocinolone TPMs among those who receive a 2nd implant. As graphed in the model validation section 4.2.11 above, among these patients this results in an additional improvement in their BCVA between year 3 and year 6 similar to that modelled as occurring between baseline and 3 years. In the opinion of the ERG a 2nd implant occurs in order to prevent the fluocinolone vitreous concentration falling. ERG expert opinion suggests that this maintenance of the

fluocinolone vitreous concentration would not result in large additional gains in BCVA, but would rather maintain that gains observed by year 3.

The ERG revised base case will assume no additional benefit from a 2nd fluocinolone implant, and rather assume that BCVA will be maintained from this point.

Extrapolation of net BCVA benefit beyond year 6

As noted in the validation section, a net BCVA gain is modelling at the end of year 6. After year 6 all patients cease treatment and are assumed to have a 3.5% quarterly probability of declining to the next worse health state. As a consequence, the end of year 6 net BCVA gain is maintained for the next 24 years.

It is difficult to apply a waning of effect in the model. The closest possible approximation is to simply limit the time horizon, 18 years being a crude approximation to the net gain waning to zero by the end of the 30 year time horizon. The ERG will perform this scenario analysis but will not attempt to more formally model a waning of effect from year 6.

Quality of life studies and values

The company applies the algorithm developed by Rentz et al[60] to FAME NEI-VFQ data to derive the quality of life values for the company base case. It also provides scenario analyses that apply the algorithms developed by Brazier et al[61], which as already noted develops two algorithms to estimate quality of life values from bilateral BCVA data among DMO patients, one valuing NEI-VFQ-25 data using the Rentz et al algorithm and the other valuing EQ-5D data using the UK social tariff.

The company implementation of the Rentz et al algorithm at first sight looks a little peculiar. Individual patient quality of life values are not estimated by applying the Rentz et al algorithm, with these then being averaged. Rather, the company sums the categorical NEI-VFQ responses split by FAME arm, applies part of the Rentz algorithm to these summed amounts, averages this element to derive the "theta severity score" and applies the final step of the Rentz algorithm to this theta. Why the company adopts this approach is not known, or whether it introduces any biases. The ERG cannot explore this further. It could be argued that the Brazier et al data set is larger and less likely to overfit to individual health states, but the FAME NEI-VFQ-25 data means that the company can directly apply the Rentz NEI-VFQ-25 algorithm so does not require the Brazier mapping function for this aspect.

The values derived by applying the Rentz et al algorithm for baseline, 24 months and 36 months are pooled by weighting the means by their inverse variance, with these values being smoothed as graphed in section 4.2.7 above. There is a degree of arbitrariness and choice surrounding the final set of values. But ERG exploratory work weighting values by the numbers of observations and revising anomalous values through a variety of means suggests that the company base case net QALY estimate is not particularly sensitive to this.

Brown et al[67] estimated quality of life among 325 US patients with impaired vision using both Time Trade Off and standard gamble. Since this provides quality of life values from actual patients, this can be briefly examined. Of note, Brown et al estimated the effect of the BCVA of the WSE among the subset of patients with good vision of at least 20/40 in their BSE.

Table 28. Brown et al [67] effect of WSE BCVA among those with good BSE BCVA

BCVA in WSE	n	TTO	SG
20/40-20/50	18	0.860	0.930
20/70-20/100	12	0.900	0.960
20/200-20/400	13	0.950	0.940
≤ 20/800 (CF)	28	0.880	0.920
≤ 20/1600 (HM/NLP)	7	0.810	0.950
CF: Counting fingers HM: Detecting hand movement NLP: No light perception			

The above is supportive of the BCVA of the WSE having little effect upon overall quality of life if the BSE has a good BCVA, at least until the WSE is in a very poor health state indeed, as per the company BSE HS1values.

The ERG has not managed to replicate the values the company derives from the Brazier et al algorithms. The ERG has some concerns around the quality of life values derived from Brazier and that the bilateral health states' quality of life values may not be those for a representative patient, but may rather have been calculated by varying the patient characteristics across the bilateral health states. But again, the ERG cannopt examine this any further.

NICE ACs have in the past questioned whether the EQ-5D is sufficiently responsive to changes in patients' BCVA and in particular changes in the BCVA of the WSE. It is unclear to the ERG whether these concerns were around unresponsiveness at the patient level or at the population level. It is easy

to imagine relatively small changes in the BCVA of the WSE not causing many patients to transition between the broad EQ-5D no problems, some problems, major problems categories and the NEI-VFQ being more responsive at the patient level. It is less obvious that the EQ-5D will be unresponsive at the population level.

Previous NICE assessments have tended to rely upon the experimental lenses study of Czoski-Murray[71], with an assumption that the WSE BCVA affects quality of life by 15% of the amount of the BSE BCVA. In the light of the Brazier et al results, an exploration of the WSE quality of life effect being 30% that of the BSE is also warranted, though this is to some extent to mix apples and pears. The resulting values are graphed below.

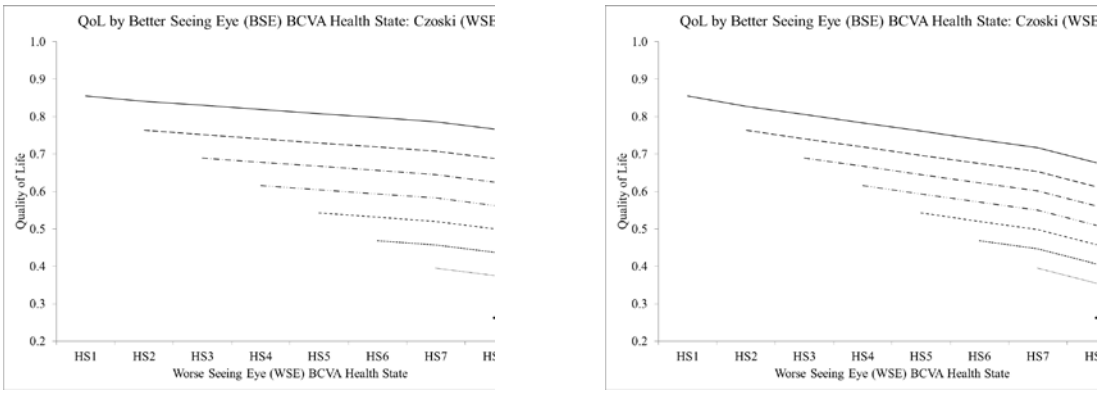


Figure 18. Czoski-Murray derived QoL values

In the light of the above the ERG will conduct analyses that apply:

- The company smoothed values derived from the Rentz et al NEI-VFQ-25 algorithm
- The company values derived from the Brazier et al EQ-5D algorithm
- ERG values derived from the Czoski-Murray algorithm, assuming the WSE has 15% the effect of the BSE and assuming the WSE has 30% the effect of the BSE[8].

Number of fluocinolone implants within 3 years

The company model assumes that patients only receive a single fluocinolone implant during the 1st 3 years of the model. Cunha-Vaz et al[38] report that during FAME among the chronic patients 76.1% of patients received 1 fluocinolone implant, 18.7% received 2 fluocinolone implants and 5.3% received

[8] Brazier et al note that the WSE BCVA coefficient is typically around 30% that of the BSE BCVA coefficient, but this is for BCVA on the log scale. For instance, if an eye with an initial BCVA of 40 letters improves to 50 letters and it is the BSE, the Brazier et al NEI-VFQ-25 model 1 suggests that quality of life will improve by an absolute 0.051. The corresponding improvement in quality of life for the same change in the WSE is around an improvement of 0.016: a ratio of 30%.

at least 3 fluocinolone implants. Assuming the 5.3% only received 3 implants yields an average of 1.3 per patient.

The ERG will apply the 1.3 fluocinolone administrations within its revised base case.

Cost of bevacizumab

The company applies the cost for a bevacizumab vial, as used for cancer treatment. The requirement for ophthalmic use is less, as was reviewed in some detail during the 2018 NICE economic modelling for NG82: Age-related macular degeneration. The ERG will apply the NG82 £49 cost.

Composite comparator drug costs

When calculating the on treatment drug costs for the usual care arm the model applies the weighted average number of administrations to the weighted average annual drug cost per administration. This is incorrect. Weighting each comparators annual drug cost increases the annual drug cost from £547 to £590.

Cost of administration

ERG expert opinion suggests that fluocinolone is a more involved procedure than the anti-VEGFs. It involves a bigger needle, so needs sub-conjunctival anaesthesia. An anti-VEGF administration might take 10-15 minutes, and a fluocinolone injection is likely to add perhaps an additional 5 minutes to this. In the light of this but there being no immediately obvious HRGs to fully account for this, the ERG will explore increasing the fluocinolone administration cost by 33%.

ERG expert opinion also notes that because anti-VEGFs are simpler to administer, administration is increasingly by nurses or hospital optometrists. The 2016-2017 NHS reference costs for non-admitted face to face follow-up visits are £86 if it is consultant led and £43 if it is non-consultant led. Not all anti-VEGFs will be administered by non-consultant staff and the OP visit costs are not immediately relatable to the BZ87A £108 outpatient procedure cost. But this does argue for an exploration of this by reducing the anti-VEGF administration cost by an admittedly arbitrary 25%.

Monitoring of fluocinolone: Resource use

The numbers of monitoring visits for fluocinolone are apparently drawn from FAME data. Subsequent to baseline, week 1 and week 6 visits fluocinolone was monitored quarterly within FAME. It appears that the company averages may be among baseline patient numbers rather than among those remaining in the trial e.g. year 3 had an average of three monitoring visits which roughly

corresponds to quarterly monitoring among the 79% of chronic patients remaining in FAME at month 36.

It may be more reasonable to apply the fluocinolone monitoring frequency among those who were followed up. The fluocinolone SmPC also states "*Following the procedure, patients should be monitored for potential complications such as endophthalmitis, increased intraocular pressure, retinal detachments, and vitreous haemorrhages or detachments. Biomicroscopy with tonometry should be performed between two and seven days after the implant injection. Thereafter it is recommended that patients are monitored at least quarterly for potential complications, due to the extended duration of release of fluocinolone acetonide, of approximately 36 months*". The ERG will apply quarterly monitoring for fluocinolone, and explore the company base case monitoring frequency in a sensitivity analysis.

The fluocinolone SmPC specifying *"biomicroscopy with tonometry"* also differs from the ranibizumab SmPC which only specifies that "*Monitoring … may include clinical examination, functional testing or imaging techniques (e.g. optical coherence tomography or fluorescein angiography*". Recalling that OCT and FA are costed separately in the model, this could imply that a fluocinolone monitoring visit is more involved and costly than an anti-VEGF monitoring visit. The ERG explores this in a sensitivity analysis by adding an arbitrary 50% to the unit cost of fluocinolone monitoring visits.

Monitoring of anti-VEGFs: Resource use

The company monitoring assumptions for those receiving ongoing anti-VEGF treatment appear to be for those starting anti-VEGF treatment rather than continuing with ongoing anti-VEGF treatment.

Among those with an insufficient response to anti-VEGF treatments, the annual average number of anti-VEGF injections based upon ICE-UK is roughly 4, or 1 per quarter. This average possibly conceals quite large differences between patients. There may be a distinction between patients who never "dry" and those with recurrences. ERG expert opinion suggests that monitoring visits may become less frequent as injections become less frequent, but that those treated less frequently than monthly are likely to have more monitoring visits than injections.

In the light of this and only 4 anti-VEGF administrations being administered among ICE-UK patients in the year prior to their fluocinolone implant, the ERG will reduce the number of monitoring visits for anti-VEGF to 8 in the 1st year, and perform sensitivity analyses of 4 visits and 12 visits.

Monitoring: Averaging annual cost and extrapolation

As outlined above in section 4.2.8 which summarises the company approach to resource use, for both fluocinolone and the anti-VEGFs the annual monitoring costs in years 1, 2 and 3 falls with time.

- For fluocinolone the company applies the individual costs for years 1, 2 and 3 and applies the year 3 costs in years 4, 5 and 6.
- For anti-VEGFs the company averages the costs of year 1, 2 and 3 and applies this average across all years. This carries forward the substantially higher year 1 monitoring costs into years 4, 5 and 6 which is not legitimate.

The ERG will revise the model to apply the individual costs for years 1, 2 and 3 and apply the year 3 costs in years 4, 5 and 6 for both fluocinolone and for the anti-VEGFs.

Bilateral monitoring

The company models treatment specific resource use for monitoring of the study eye. The elements of this such as outpatient visits as typically less for fluocinolone than for the anti-VEGFs or usual care.

But monitoring resource use for the fellow eye is assumed not to be affected by the monitoring resource use for the study eye. In the company base case this results in total bilateral monitoring costs in the fluocinolone arm typically being around 60% of that in the usual care arm, despite patients in both arms effectively being modelled as receiving anti-VEGF in their fellow eye.

ERG opinion is that it is more reasonable to assume that the monitoring required for bilateral treatment should be the maximum of that required for treatment of the study eye or of the fellow eye. For instance, if the study eye is receiving fluocinolone and only requires quarterly monitoring while the fellow eye is receiving anti-VEGF and requires bi-monthly monitoring, bilateral monitoring should be bi-monthly. There may be an argument for applying an additional percentage to the resulting bilateral monitoring costs, but this a side issue compared to the frequency of resource use that should be assumed.

The ERG will assume that bilateral monitoring visit frequency is the maximum required in either eye.

Costs of blindness

The model structure divides the annual costs of blindness by 4 to give cost per blind patient per quarterly cycle. But the model structure applies this to the life years lived in blindness per cycle; i.e.

the number of blind patients divided by 4. As a consequence, for a given annual cost of blindness the total costs of blindness are ¼ what they should be[9].

In the light of the previous ERG comment that the company annual cost of blindness estimate is double what it should be, these two points taken together imply that the estimated costs of blindness are around ½ what they should be.

Blindness mortality multiplier

Mortality is modelled a little oddly. Individual eyes are modelled as dying, and patient deaths are taken to be half the number of eyes that have "died". The mortality multiplier for blindness is applied to the probability of individual eyes dying. As a consequence, deaths due to blindness are overestimated.

In the opinion of the ERG any bias is likely to be greater from inclusion of the blindness mortality multiplier than its exclusion. The ERG excludes it from most of its analyses, but includes it in a sensitivity analysis.

Ongoing study eye costs years 6-30

If the blindness mortality multiplier is applied there is a survival benefit from fluocinolone. As a consequence, ongoing costs of study eye and fellow eye care from year 6 would, if included, not net out between the arms. Since they are not included, applying the blindness mortality multiplier results in further bias.

4.4 Exploratory and sensitivity analyses undertaken by the ERG

Key assumptions that must be explicit are the clinical effectiveness that should be applied for each of the individual treatments during years 1-3 of the model, what treatment effectiveness should be extrapolated for each of the individual treatments during years 4-6 of the model, and what net

[9] This is most easily seen by examining costs in the 1st cycle of the model when all are alive. The patient distributions suggest that 18% of study eyes are in either HS7 or HS8 at baseline, while for the fellow eye if it is in bilateral DMO the HS7 or HS8 proportion is 17% compared with only 3% for fellow eyes with no bilateral involvement. Weighting these by the 77% bilateral at baseline results in a proportion who are bilaterally in HS7 or HS8 of 2.4%, or 24 patients in a cohort of 1,000. An annual per patient cost of blindness of £19,795 suggests a quarterly cost of £4,948 which if incurred by 24 patients results in a quarterly cost of £120,666. The cost of blindness in the 1st quarterly cycle of the model is only £30,166, a quarter of £120,666.

treatment effect between the arms should be extrapolated during years 7-30 when all are assumed to have ceased treatment.

In the opinion of the ERG, for years 1-3 of the model the most reasonable clinical effectiveness estimates from those that are available are:

- The FAME fluocinolone TPMs for fluocinolone
- The FAME sham TPMs for laser / no drug treatment
- The FAME sham TPMs conditioned by an odds ratios of gaining letters, and preferably an odds ratio of losing letters, for the anti-VEGFs relative to laser

But the odds ratios of gaining and losing letters for the anti-VEGFs relative to laser for the patient group under consideration are highly uncertain. As a consequence, the ERG will apply an odds ratio of 1.00 for its base case, and perform scenario analyses that apply the company odds ratio of gaining letters of 1.54 and the speculative ERG 1.23 estimate to illustrate possible effects.

For reasons reviewed in more detail above the ERG presents pairwise analyses that compare:

- Fluocinolone against laser / no drug treatment
- Fluocinolone against the anti-VEGFs
- Fluocinolone against the company composite comparator

The ERG does not present a stacked analysis for the reasons given at the end of this section.

The ERG revises the company model along the following lines[10]:

- Apply the FAME sham arm effectiveness for laser.
- Assume an odds ratio for anti-VEGF relative to laser of 1.00, due more to a lack of data than a formal demonstration of the two being clinically equivalent.
- Apply the cataract removal regression which does not include the treatment effect, due to its inclusion resulting in a treatment coefficient with a p-value of [redacted].
- Remove the blindness mortality multiplier.
- Apply the same fellow eye cohort flow in both arms.
- Apply the FAME mean 1.3 fluocinolone injections.
- Apply the NG82 bevacizumab cost of £49.

[10] All ERG changes are documented with full cell referencing in the *ERG* worksheet of the ERG revised model. There are some additional relatively minor changes made to the cohort flow calculations, which are also fully documented in the model.

- For the composite comparator calculate the average drug cost based upon a weighted average of the individual comparator drug costs.
- Apply the year 1, 2 and 3 monitoring costs and use the year 3 costs for years 4, 5 and 6.
- Apply quarterly monitoring for fluocinolone in years 2 and 3 due to the SmPC.
- Apply bi-monthly monitoring for those on anti-VEGF due to this being established treatment rather than new treatment with anti-VEGF.
- Assume bilateral monitoring to be the most costly of that required for the study eye and the fellow eye.
- Revise the costs of blindness to be aligned with the cited reference and remove the incorrect quartering the costs of blindness.

The ERG conducts the following scenario analyses:

- SA01: Apply the company estimate of 1.54 for the odds ratio of gaining letters for anti-VEGF relative to sham and the ERG estimate of 1.23.
- SA02: All comparators have no effect upon BCVA while fluocinolone has the full change from baseline to 36 months of FAME applied.
- SA03: Assume no further improvements in BCVA during year 4, 5 and 6, other than those arising from cataract removal.
- SA04: Apply the cataract removal regression which does include the treatment effect.
- SA05: All with cataract at baseline, and none with cataract at baseline.
- SA06: All with a retinal thickness at baseline of less than 400µm, and all with a thickness of more than 400µm[11].
- SA07: Anti-VEGF, all receive ranibizumab and all receive bevacizumab.
- SA08: Natural history worsening per quarter by 2% and 5%.
- SA09: Applying quality of life values based upon the Brazier et al NEI-VFQ-25 mapping, the Brazier EQ-5D mapping, Czoski-Murray with the WSE having 15% the impact of the BSE and Czoski-Murray with the WSE having 30% the impact of the BSE.
- SA10: Combine study eyes and fellow eyes using the company joint distribution calculation.
- SA11: Fluocinolone administration costs 33% higher to account for the increase in time relative to anti-VEGFs.

[11] Note that the coefficient for this in the regression analysis affects the TPM calculations for fluocinolone and for sham in a like manner, though admittedly on the log scale. There is no allowance for a treatment effect interaction whereby e.g. the net effect for fluocinolone over sham is greater for those with thicker retinas than for those with thinner retinas.

- SA12: Anti-VEGF administration costs 25% lower to account for non-consultant administrations.
- SA13: Anti-VEGF year 1 monitoring visits of 4, and 12.
- SA14: 50% fluocinolone retreatment with this higher retreatment including additional improvements in BCVA in years 4, 5 and 6 among those retreated.
- SA15: Only 1.0 fluocinolone implant every 3 years for all patients
- SA16: Blindness 30% self-funding residential care.
- SA17: An 18 year time horizon to approximate a linear waning in the treatment effect after year 6.
- SA18: combining SA02 and SA03
- SA19: combining SA02, SA03 and SA17.

Pairwise comparison: Fluocinolone against laser / no drug treatment

The ERG scenario analyses for the pairwise comparison of fluocinolone and laser are as follows.

Table 29. Fluocinolone vs laser: ERG Analyses

Analysis	Δ Costs	Δ QALYs	ICER
Base case	£6,644	0.020	£334k
SA01a: 1.54 OR for anti-VEGF vs laser	..	..	..
SA01b: 1.23 OR for anti-VEGF vs laser	..	..	..
SA02: Fluocinolone net gain = FAME change from baseline	£5,110	0.200	£25,550
SA03: No additional BCVA in years 4, 5 and 6	£7,077	-0.030	Dominated
SA04: Cataract removal regression with treatment effect	£6,596	0.029	£226k
SA05a: 100% cataract at baseline	£6,484	0.035	£187k
SA05b: 0% cataract at baseline	£6,805	0.005	£1.3mn
SA06a: 100% retinas < 400μm at baseline	£6,737	0.015	£463k
SA06b: 100% retinas ≥ 400μm at baseline	£6,574	0.023	£289k
SA07a: anti-VEGF 100% ranibizumab	..	..	..
SA07b: anti-VEGF 100% bevacizumab	..	..	..
SA08a: Natural history 2% worsening per quarter	£6,753	0.022	£304k
SA08b: Natural history 5% worsening per quarter	£6,555	0.018	£366k
SA09a: Brazier et al NEI-VFQ-25 QoL algorithm	£6,644	0.019	£353k
SA09b: Brazier et al EQ-5D QoL algorithm	£6,644	0.007	£913k
SA09c: Czoski-Murray QoL with WSE 15% of BSE	£6,644	0.033	£198k
SA09d: Czoski-Murray QoL with WSE 30% of BSE	£6,644	0.030	£224k
SA10: Combine study eye and fellow eye joint distribution	£6,293	0.020	£321k
SA11: Fluocinolone administration costs +33%	£6,705	0.020	£337k
SA12: anti-VEGF administration costs -25%	..	..	..
SA13a: anti-VEGF 4 monitoring visits in year 1	..	..	..
SA13b: anti-VEGF 12 monitoring visits in year 1	..	..	..
SA14: 50% fluocinolone retreatment at end of year 3	£7,067	0.036	£195k
SA15: Only 1 fluocinolone implants every 3 years	£5,271	0.020	£265k
SA16: Blindness residential care 30% self-funded	£6,615	0.020	£332k
SA17: 18 year time horizon	£6,748	0.016	£425k
SA18: SA02 +SA03	£5,462	0.156	£34,947
SA19: SA02 + SA03 + SA17	£5,773	0.144	£40,224

Unfortunately the ERG model revisions cause the probabilistic analysis to not run. Time constraints mean that the ERG has not been able to find the reason for this, and as a consequence the ERG does not present any probabilistic analyses. The company probabilistic analysis resulted in similar central estimates to the deterministic model, though it has to be acknowledged that these did not sample the majority of the regression coefficients of the MMRM clinical effectiveness estimates.

The ERG revisions to the model suggest that fluocinolone has an extremely poor ICER relative to laser if it is assumed that fluocinolone only results in the net gain observed during FAME. The relevant company scenarios for this are:

- Scenario 1 which applied the sham clinical effectiveness but retained the costs of anti-VEGF which saw fluocinolone worsen from being dominant to an ICER of £14,753 per QALY
- Scenario 2 which removed the costs of anti-VEGF but retained the assumption that the net gain from fluocinolone over usual care would be the absolute change from baseline to 36 months in the FAME fluocinolone arm, which saw fluocinolone worsen from dominance to an ICER of £15,842 per QALY.

If the company had combined these two scenarios, as in the ERG revised base case above, the resulting ICER would also have been extremely poor.

But as shown in the model validation section above, the model estimates for when the FAME sham effectiveness estimates are inputted to the usual care arm result in gains in excess of those observed during FAME. Excluding these effects and assuming that the net gain from fluocinolone over laser is equal to the FAME fluocinolone change from baseline, SA02, results in a cost effectiveness estimate of £25,550 per QALY. But SA02 assumes that there are additional net gains in BCVA beyond year 3. If these gains are not applied as in SA03 it is estimated the fluocinolone actually results in a small QALY loss. This may arise due to the sham modelling not according particularly well to the sham arm of FAME, coupled with the additional BCVA gains in the usual care arm as patients progress through having their cataracts removed.

Whether there is a treatment effect upon the likelihood of those with cataract having them removed, SA04, has a reasonable impact.

Among the subgroup without cataracts the cost effectiveness worsens somewhat. It also worsens among those with thinner retinas, though in this regard it should be borne in mind that the company regression does not contain a treatment interaction effect for this coefficient.

The rates at which patients worsen after year 6 has some effect, this mainly affecting how quickly patients progress to blindness, particularly in the sham arm, with this appearing to mainly affect the costs of blindness which are incurred.

The Brazier et al EQ-5D quality of life algorithm somewhat reduces the net QALY gains, but it can be noted that the Czoski-Murray quality of life values which have often been used in previous appraisals improves them. This is not to argue that the Czoski-Murray quality of life values should be applied for the base case, but gives some read across to previous appraisals. The results for the Czoski-Murray quality of life values with the WSE having 15% the effect of the BSE are better than those where the effect is assumed to be 30%, which is counterintuitive.

Increasing the proportion who receive a 2nd implant improves the ICER, due to these patients being modelled as experiencing further gains in their BCVA during years 4, 5 and 6 and the resulting increase in the net BCVA at the end of year 6 being maintained during years 7-30 at no additional treatment cost.

If patients only receive 1 fluocinolone implant over a three year period rather than the 1.3 that occurred during FAME, the net costs improve by a reasonable margin as would be expected.

Restricting the time horizon to 18 years to approximate the net treatment effect waning after all treatments are stopped at the end of year 6 reduces the net QALY gain by a reasonable margin.

The net QALYs are relatively small over most of the scenario analysis and the ICERs are correspondingly erratic. The consistent result is that fluocinolone results in considerably higher costs than laser due to the cost of the implant, but also in part its requirement for ongoing quarterly monitoring. The ICER is sensitive to whether the net effect from fluocinolone over laser is modelled as the change between baseline and 36 months in the FAME fluocinolone arm or as the net gain at 36 months for the FAME fluocinolone arm compared to the FAME sham arm. This is complicated by the validation work suggesting that the model may overestimate the BCVA at 36 months when applying the FAME sham arm treatment effect.

Pairwise comparison: Fluocinolone against anti-VEGFs

Table 30. Fluocinolone vs anti-VEGF: ERG Analyses

Analysis	Δ Costs	Δ QALYs	ICER
Base case	£3,111	0.018	£176k
SA01a: 1.54 OR for anti-VEGF vs laser	£3,996	-0.163	Dominated
SA01b: 1.23 OR for anti-VEGF vs laser	£3,699	-0.071	Dominated
SA02: Fluocinolone net gain = FAME change from baseline	£1,626	0.196	£8,302
SA03: No additional BCVA in years 4, 5 and 6	£3,544	-0.033	Dominated
SA04: Cataract removal regression with treatment effect	£3,063	0.027	£113k
SA05a: 100% cataract at baseline	£2,951	0.032	£90,959
SA05b: 0% cataract at baseline	£3,272	0.003	£1.0mn
SA06a: 100% retinas < 400µm at baseline	£3,204	0.012	£259k
SA06b: 100% retinas ≥ 400µm at baseline	£3,040	0.021	£148k
SA07a: anti-VEGF 100% ranibizumab	£2,713	0.018	£153k
SA07b: anti-VEGF 100% bevacizumab	£5,853	0.018	£331k
SA08a: Natural history 2% worsening per quarter	£3,220	0.020	£161k
SA08b: Natural history 5% worsening per quarter	£3,022	0.016	£193k
SA09a: Brazier et al NEI-VFQ-25 QoL algorithm	£3,111	0.017	£188k
SA09b: Brazier et al EQ-5D QoL algorithm	£3,111	0.005	£614k
SA09c: Czoski-Murray QoL with WSE 15% of BSE	£3,111	0.031	£99,467
SA09d: Czoski-Murray QoL with WSE 30% of BSE	£3,111	0.027	£113k
SA10: Combine study eye and fellow eye joint distribution	£2,759	0.017	£159k
SA11: Fluocinolone administration costs +33%	£3,172	0.018	£179k
SA12: anti-VEGF administration costs -25%	£3,281	0.018	£185k
SA13a: anti-VEGF 4 monitoring visits in year 1	£2,855	0.018	£161k
SA13b: anti-VEGF 12 monitoring visits in year 1	£3,308	0.018	£187k
SA14: 50% fluocinolone retreatment at end of year 3	£3,117	0.034	£92,233
SA15: Only 1 fluocinolone implants every 3 years	£1,738	0.018	£98,241
SA16: Blindness residential care 30% self-funded	£3,082	0.018	£174k
SA17: 18 year time horizon	£3,214	0.014	£235k
SA18:SA02 + SA03	£1,992	0.151	£13,176
SA19: SA02 + SA03 + SA17	£2,347	0.138	£16,985

Because the ERG revised base case does not apply an odds ratio for anti-VEGF relative to laser the net QALYs are much the same as for the comparison with laser, only differing due to the adverse events profiles. There are still reasonably large net costs due to the increased number of fluocinolone implants. The revised ERG monitoring cost implementation and assumptions also somewhat reduces the cost offsets from this element. But costs in the comparator arm are considerably higher than for the comparison with laser and the ICER improves as a result.

Applying odds ratios for anti-VEGF relative to laser causes the anti-VEGFs being estimated to dominate fluocinolone. These analyses are only illustrative, and it should be borne in mind that these are applied in the context of the sham arm TPMs, which as reviewed above and in the model validation section appear to overestimate the effectiveness of sham compared to that observed in FAME. These concerns are again reflected in the results of SA04. But if the sham TPMs are not the basis of the comparison and it is assumed that the net gain from fluocinolone is the change between baseline and 36 months in the FAME fluocinolone arm, both the net costs and the net QALYs improve yielding an ICER of £8,302 per QALY.

As in the comparison with laser, the cataract removal model that includes a treatment effect has a reasonable effect and improves the ICER.

Cost effectiveness is also estimated to improve among those with cataracts at baseline and those with thinner retinas. Among the latter is should be recalled that this is applying the FAME sham arm TPMs so is not really specific to anti-VEGF.

Among patients in the comparator arm who receive bevacizumab the net cost of fluocinolone is predictably higher and the ICER somewhat worse at £331k per QALY.

Results show a similar sensitivity to the quality of life values that are applied as in the comparison with laser.

Varying monitoring visits costs and the number required for anti-VEGF in year 1 has some effect, but given the other ERG changes to the implementation of monitoring costs and the net drug costs these are not in themselves major drivers. Within the scenario analyses around costs, a key determinant is whether patients will during a three year period only ever get 1 fluocinolone implant rather than the 1.3 of the FAME fluocinolone arm.

Approximating a linear decline in the net treatment after year 6 due to all treatments having ceased has a reasonable effect upon the net QALY gain, with the ICER worsening accordingly.

Pairwise comparison: Fluocinolone against composite comparator

Table 31. Fluocinolone vs composite comparator: ERG Analyses

Analysis	Δ Costs	Δ QALYs	ICER
Base case	£4,084	0.018	£223k
SA01a: 1.54 OR for anti-VEGF vs laser	£4,855	-0.117	-£41,436
SA01b: 1.23 OR for anti-VEGF vs laser	£4,548	-0.047	-£97,281
SA02: Fluocinolone net gain = FAME change from baseline	£2,610	0.196	£13,300
SA03: No additional BCVA in years 4, 5 and 6	£4,517	-0.032	-£141k
SA04: Cataract removal regression with treatment effect	£4,036	0.028	£146k
SA05a: 100% cataract at baseline	£3,924	0.033	£119k
SA05b: 0% cataract at baseline	£4,245	0.004	£1.2mn
SA06a: 100% retinas < 400µm at baseline	£4,177	0.013	£322k
SA06b: 100% retinas ≥ 400µm at baseline	£4,013	0.021	£190k
SA07a: anti-VEGF 100% ranibizumab	..	..	..
SA07b: anti-VEGF 100% bevacizumab	..	..	..
SA08a: Natural history 2% worsening per quarter	£4,193	0.021	£204k
SA08b: Natural history 5% worsening per quarter	£3,995	0.016	£245k
SA09a: Brazier et al NEI-VFQ-25 QoL algorithm	£4,084	0.017	£237k
SA09b: Brazier et al EQ-5D QoL algorithm	£4,084	0.006	£718k
SA09c: Czoski-Murray QoL with WSE 15% of BSE	£4,084	0.032	£128k
SA09d: Czoski-Murray QoL with WSE 30% of BSE	£4,084	0.028	£145k
SA10: Combine study eye and fellow eye joint distribution	£3,733	0.018	£207k
SA11: Fluocinolone administration costs +33%	£4,145	0.018	£226k
SA12: anti-VEGF administration costs -25%	£4,242	0.018	£232k
SA13a: anti-VEGF 4 monitoring visits in year 1	£3,926	0.018	£214k
SA13b: anti-VEGF 12 monitoring visits in year 1	£4,225	0.018	£231k
SA14: 50% fluocinolone retreatment at end of year 3	£4,208	0.034	£122k
SA15: Only 1 fluocinolone implants every 3 years	£2,711	0.018	£148k
SA16: Blindness residential care 30% self-funded	£4,055	0.018	£221k
SA17: 18 year time horizon	£4,188	0.014	£293k
SA18: SA02+SA03	£2,974	0.152	£19,612
SA19: SA02+SA03+SA17	£3,327	0.139	£23,998

As would be expected, for the composite comparator the ICERs lie between the corresponding ICERs of the pairwise comparison with laser and the pairwise comparison with anti-VEGF, tending to be closer to those of the anti-VEGF comparison due to the 28:72 split in the composite comparator between laser and anti-VEGF.

This also carries through to the scenario analyses that apply an odds ratio for anti-VEGF relative to laser, which result in the comparator arm being estimated to be both cheaper and more effective than the fluocinolone arm. Quite what odds ratio should be applied is extremely uncertain, but these scenario analyses illustrate that if only the net treatment effect of fluocinolone over sham in FAME should be applied, any treatment effect for anti-VEGF over laser could have a marked effect upon results.

The scenario analyses around the number of anti-VEGF monitoring visits have the opposite effect to those in the pair-wise comparison with the anti-VEGF. The ERG has not sourced the reason for it, but it may be due to the effect upon the net cost of fellow eye monitoring.

Stacked comparison: Fluocinolone, laser / no drug treatment and anti-VEGFs
Given the results above, the ERG does not present a stacked analysis in part because it is unlikely to affect conclusions. It is also not presented because of differing treatment costs in the fellow eye, the fellow eye being assumed to incur the treatment costs of the study eye usual care, which means that a stacked analysis would to some extent mix apples and pears. The ERG has not had time to revise the model to assume the same treatment in the fellow eye for the pair wise comparison of fluocinolone and laser as for the pairwise comparison of fluocinolone and anti-VEGF.

4.5 Conclusions of the cost effectiveness section

The main issues relate to how well the available clinical effectiveness estimates match the decision problem.

- Is the patient group homogeneous? If so, is it reasonable to consider a single composite comparator or should the composite comparator be split into its constituent parts? If not, are those with insufficient response to laser a distinct subgroup from those with insufficient response to anti-VEGF and should they be considered separately?
- For those on non-drug treatment receiving laser what is the clinical effect of remaining on current treatment? What would be the net effect of switching to fluocinolone? Does the FAME data

used in the model match this? To what extent should this comparison be expected to differ from the assessment of the chronic phakic during TA301?

- For those on anti-VEGF what is the clinical effect of remaining on current treatment? What would be the net effect of having anti-VEGF withdrawn? What would be the net effect of switching to fluocinolone? How should the odds ratios for ranibizumab estimated by the company and the ERG be viewed within this? How well does the FAME data used in the model inform this?

In the opinion of the ERG the FAME sham arm reflects a degree of natural recovery, and the FAME fluocinolone arm will also include this natural recovery element. Given the patient group this natural recovery element may have been previously exhausted so may not apply. To the ERG this argues for the model removing the natural recovery element from the usual care arm and the fluocinolone arm, leaving the net treatment effect observed during FAME. But the model structure does not permit this. The best that can be achieved is to retain natural recovery element in both arms in order to model the FAME net treatment effect. This is achieved by applying the FAME sham arm effectiveness for usual care and the FAME fluocinolone arm effectiveness for fluocinolone. The model validation section suggests that the model inputs and structure overestimate the treatment effectiveness of the FAME sham arm. As a consequence, the resulting clinical effectiveness estimates are probably biased against fluocinolone.

In the opinion of the ERG it is not suitable to model additional net benefits from ongoing treatment beyond the end of year 3, and it is more reasonable to assume that the net benefits that are realised at the end of year 3 are maintained but do not increase further. The ERG revised modelling only applies this as a scenario analysis due to the validation problem when sham effectiveness is assumed for the comparator arm, as outlined above. But if the modelling of the comparator arm can be better aligned with the FAME sham arm effectiveness at 36 months, the ERG is of the opinion that additional net gains in BCVA beyond this should not be modelled, other than those arising from cataract removal.

A key consideration is whether patients will only ever receive 1 fluocinolone implant during each 3 year period or will receive more, 1.3 being received during the 3 years of FAME. In the opinion of the ERG the number observed during FAME should be applied. There are also questions around the proportion of those who receive a fluocinolone implant who continue to receive anti-VEGFs after implant, and in both arms what number of anti-VEGF administrations should be extrapolated among those receiving them.

Another issue is whether those with cataract at baseline should be considered separately from those without cataract at baseline. Similarly, should those with thinner retinas be considered separately from

those with thicker retinas, does the company regression analysis sufficiently explore this and what should be assumed for the anti-VEGFs in this regard.

The company implementation of monitoring costs is in the opinion of the ERG biased and overestimates the savings that will result from this element. It is unreasonable to average anti-VEGF monitoring costs over years 1, 2 and 3 and carry forward this average, rather than carry forward the year 3 monitoring cost. It also seems more reasonable to the ERG to model bilateral monitoring costs as the more expensive of the monitoring cost of the study eye and the monitoring cost of the fellow eye.

The company estimate of the cost of blindness is biased and is roughly double the estimate that the company has used in previous assessments. But the company model incorrectly quarters this estimate meaning that the costs of blindness are roughly half what they should be.

After year 6 the net gain in BCVA between the arms is maintained at no additional treatment cost for 24 years. If this extrapolation is not reasonable the model is biased in favour of fluocinolone, and a waning of the net gain in BCVA during the 24 years should be applied.

The company submission concentrates upon treatment in the study eyes. It does not particularly consider differentiating treatment in the fellow eyes by arm, and the company may not have faith in the model structure for this. The ERG has not rebuilt the elements of the model that would be used for this, but has some concerns about them. Treatment with fluocinolone in the fellow eyes also requires assumptions about their chronic status and phakic status, and where within the treatment pathway they fall. But this might result in a higher proportion of chronic, phakic eyes that are treated with fluocinolone being the better seeing eye. This might improve the cost effectiveness estimate for fluocinolone.

5 DISCUSSION

5.1 Problems with evidence

The main problem is this assessment is the lack of trial evidence of effectiveness of fluocinolone in people with DMO that has not responded well to anti-VEGF drugs. The only trial data comes from the FAME trials which started before the anti-VEGFs came into common use.

In 2014, the FDA [37] said:

"It is recommended that a new study with at least 12 months of follow-up be submitted in which patients who have failed to respond to a three month or more course of anti-VEGF therapy are randomized between your drug product and continued anti-VEGF therapy."

No such trial has been done. Instead, Alimera has supported some observational studies, described earlier in this report.

Other problems include uncertainties around the definition of chronic, and "insufficiently responsive".

5.2 Defining lack of response.

We have noted that response is usually taken to be a gain of 10 letters, or in some studies 15 letters. However it can be argued that in a condition where slow deterioration is common, it can be argued that sustained stability (e.g. gains or losses of less than five letters) is also a benefit. We know from observational studies that anti-VEGF is sometimes continued even when BCVA is declining. Presumably the treating ophthalmologists believe that deterioration is being slowed. Without a trial, this cannot be proved or disproved, nor its cost-effectiveness assessed.

5.3 Economic assumptions

The economic analysis by Alimera is largely based on the assumption stated in their Summary, that

"Clinical data show that first-line treatment with anti-VEGFs is insufficient in approximately 40% of patients following monthly administration, and as a result of current NICE recommendations, restricting second-line therapies, DMO patients with phakic eyes and an insufficient response to first-line therapy may continue to receive expensive anti-VEGF injections."

The statement about the high proportion that does not respond sufficiently (such as 10 or more letters) to anti-VEGF drugs is correct. These drugs are a major advance in DMO but are far from a complete solution. The reference to "expensive" applies to ranibizumab and aflibercept but less to

bevacizumab, though that still require frequent attendances for injections and monitoring. There is evidence that in some areas, the NHS has problems providing optimal delivery of anti-VEGF services [72] with patients not being seen in time, so reducing the number of attendances would help with that. The Royal College of Ophthalmologists [73] has expressed concern about the number of consultant ophthalmologists being insufficient to meet rising demand. (Though it should be noted that in many centres, injections are given by nurses or optometrists.)

Alimera may be correct in saying that despite an insufficient response, patients may continue to receive expensive anti-VEGFs. However, it is very unlikely that in such patients, the anti-VEGFs would be cost-effective. Ophthalmologists may feel that although no significant improvement in vision has been achieved, the anti-VEGFs may be preventing or slowing deterioration, but there is no evidence for this. Unfortunately, the NICE guidances on the two approved anti-VEGF drugs TA274 and TA 346 [24] [23] do not include stopping rules.

Stopping anti-VEGFs drugs that are having an insufficient effect may require difficult conversations with, and possibly removing hope from, patients, especially when the treating ophthalmologists do not know if the drugs are still having some effect, even if only slowing decline of vision.

The Alimera submission assumes that in patients not treated with fluocinolone, no further deterioration occurs. This is optimistic, and disadvantages fluocinolone. If oedema remains present in the macula, sooner or later it would be expected that a drop in vision will occur in some patients. However in DMO, sight decrease is slow in many patients. DMO waxes and wanes. In the 2-year follow-up in RISE and RIDE, 28% of patients in the sham group did not need rescue therapy with laser.

5.4 Economic analysis

Within the economics a key difference between the ERG and the company is whether the fluocinolone should be modelled as having a net effect over the comparators equal to the FAME net effect at 36 months or as having a net effect over the comparators equal to the absolute change between baseline and 36 months in the FAME fluocinolone arm. In the opinion of the ERG, the FAME sham arm shows a natural recovery element which will also have been present in the FAME fluocinolone arm. To the ERG this argues for the modelled net gain being the FAME net effect at 36 months.

The only means of implementing the FAME net effect at 36 months within the company model structure is to apply probabilities estimated from the FAME fluocinolone arm to model fluocinolone and probabilities estimated from the FAME sham arm to model the comparators. But validation work shows that modelling using probabilities estimated from the FAME sham arm has a poor correspondence with the FAME sham arm, overestimating changes in BCVA and probably biasing the model against fluocinolone. This is not an argument against modelling the FAME net effect, just an observation that the company model does not do it very well.

So, is the most appropriate treatment effect estimate that is currently available the FAME net effect? If so, is the company model fit for purpose?

Turning to the most appropriate treatment effect estimate, FAME was not in the population of interest. FAME also did not consider anti-VEGFs as they were not available at the time. Uncertainty around the clinical effectiveness estimates, the appropriateness of the FAME data, and quite what patient group is being modelled also flows through to whether the company composite comparator that combines patients receiving laser and patients receiving anti-VEGFs is appropriate.

- Is the patient group homogeneous? If so, is it reasonable to consider a single composite comparator or should the composite comparator be split into its constituent parts? If not, are those with insufficient response to laser a distinct subgroup from those with insufficient response to anti-VEGF and should they be considered separately?
- For those on non-drug treatment receiving laser what is the clinical effect of remaining on current treatment? What would be the net effect of switching to fluocinolone? Does the FAME data used in the model match this? If the current assessment should rely upon FAME data, to what extent should this comparison be expected to differ from the assessment of the chronic phakic during TA301?
- For those on anti-VEGF what is the clinical effect of remaining on current treatment? What would be the net effect of having anti-VEGF withdrawn? What would be the net effect of switching to fluocinolone? How should the odds ratios for ranibizumab estimated by the company be viewed within this? How well does the FAME data used in the model inform this?

The other main economic questions and differences in approach between the company and the ERG are:

- Among those receiving another fluocinolone implant after 3 years, will this result in additional BCVA gains over and above those that occurred during the first 3 years. The company assumes that it will. The ERG thinks that another implant maintains the intravitreal concentration of fluocinolone and that this will maintain but not further improve vision.
- How many fluocinolone implants will occur on average over a three year period? The company assumes only 1. To the ERG FAME data suggests 1.3.
- Should there be more explicit acknowledgement of the large minority of ICE-UK patients who continued to receive anti-VEGF in the year after their fluocinolone implant, and is this appropriately modelled? Would those renewing their anti-VEGF treatment start with a loading dose?
- Does the company present convincing evidence that cataracts will be more quickly removed in the fluocinolone arm than in the comparator arm? The ERG understands the company arguments, but thinks that the company statistical analysis argues that a treatment effect should not be applied.
- Do fluocinolone monitoring visits fall to less than quarterly or does the SmPC require them to be quarterly? The company models them as falling below quarterly. The ERG prefers to assume quarterly monitoring based upon its reading of the fluocinolone SmPC.
- Among anti-VEGF patients on established treatment and receiving quarterly injections would monthly monitoring still occur or would it be less frequent. The company assumes monthly monitoring. The ERG thinks that in the patient group of interest anti-VEGF monitoring would be less frequent.
- How should monitoring be extrapolated during years 4, 5 and 6, using the average of years 1, 2 and 3 or using year 3 resource use? Should the approach differ between the arms? The company uses year 3 for the fluocinolone arm and the average of years 1, 2 and 3 for the comparator arm. The ERG prefers extrapolating using the same approach in both arms and using arm specific year 3 resource use.

- Will the modelled net gain at 6 years be maintained for the next 24 years at no additional treatment cost? The company assumes that it will. The ERG is more sceptical of this and thinks waning may occur.
- Is the company cost of blindness estimate reasonable? It is more than double the estimates the company has used in previous fluocinolone submissions, despite the source paper being the same. The ERG prefers the estimates of the previous company submissions.
- Should the blindness mortality multiplier be applied? The company modelling applies it. The ERG thinks that including it results in more bias than excluding it.

Other more general questions for consideration include:

- Should those with cataracts at baseline be considered separately from those without cataracts at baseline?
- Should those with thinner retinas be considered separately from those with thicker retinas? How reliable is the FAME data for the anti-VEGFs in this regard?
- Is it necessary to differentiate treatment in the fellow eyes by arm, and if so how is this best done? The company does not base its main argument upon this, though a scenario analysis of this is very briefly reported. The ERG thinks that not doing so may bias the analysis against fluocinolone, but while it has not fully parsed this aspect of the model it has some concerns about it. The ERG also thinks that there would need to be a more detailed consideration of aspects such as the chronic status, retinal thickness, phakic status, etc. of the fellow eyes leading on to a consideration of what the balance of treatments might be in the fellow eyes.

5.5 Real world evidence

The data on response to fluocinolone reported by Yit Yang et al [39], showed an improvement in visual acuity in the pseudophakic at baseline group (called CBI in the paper – cataract surgery before implant) of about six letters, compared to only two in the sham group. In the Medisoft study of real-world experience in the UK [41], Bailey and colleagues found an improvement of 5.3 letters in their mainly (96%) pseudophakic group, although the two groups are not exactly comparable because there was a much broader range of baseline VA (5 to 85 letters) in the Medisoft study than in FAME (19 to

68). So the real-world results are similar to the FAME results. We mention this because as mentioned earlier, the real-world results for ranibizumab [31] were poorer than in the trials.

5.6 Fluocinolone as first-line therapy

One issue not addressed by the NICE scope is whether there might be a place for fluocinolone implants in phakic people with DMO who have problems with the frequency of injections and monitoring visits required for anti-VEGF treatment. Some people may have problems with access, and the prospect of one injection that lasts three years, even if it means the risk of cataract extraction being necessary is increased, may appeal. However, after any steroid injections, patients will need to be followed-up to see if they respond, and because there is a risk of raised IOP and glaucoma. In addition, many with DMO have retinopathy, and the RCO guidelines recommend monitoring moderate or higher levels of retinopathy every 4-6 months. However, that is more convenient than monthly follow-up.

Alimera envisage some such use in their (b) group;

"b) in a small group of patients who are contraindicated for first line therapies or are needle phobic where the benefit of protecting the retina outweighs the risk of cataract formation."

Fluocinolone has to be given by injection but the number of injections is much less than with anti-VEGFs.

5.7 Dexamethasone

We note that the EURETINA guidelines [29] recommend that dexamethasone (the Ozurdex implant) should be used before fluocinolone. No reasons are given. Given the concern about increases in IOP, perhaps they thought it was better to test the risk with a shorter-acting steroid implant. Ozurdex releases dexamethasone for under 3 months, with peak effect at 90 days. If there was no rise in IOP, treatment could be continued with longer-acting fluocinolone. However, it is possible that IOP may not rise with dexamethasone but rise with fluocinolone. It is also possible that even using the same steroid, there may be no rise with the first injection but a rise with a second one. So at present there is no sound evidence base for a trial of a shorter-acting steroid as a predictor of raised IOP with fluocinolone. Ozurdex is not currently recommended by NICE (TA349)[25] for use in phakic eyes.

A previous study tried to use topical steroids to predict which patients would have marked rises in IOP with intravitreal steroids. This was not very successful (see Han [74] for review) with a negative

predictive value of only 60% - i.e. 40% of patients who had no rise in IOP after topical steroids would have a rise after intravitreal steroids.

5.8 Research need: does DMO change over time?

It has been suggested, by Augustin amongst others[75], that in the early stages, DMO is driven by VEGF, but that over time it becomes more of a chronic inflammatory condition. If so, there would be logic to using anti-VEGF drugs in the early stages and intravitreal steroids in more chronic DMO. However, it may be that some patients have inflammation at early stages (it is known to occur in diabetic retinopathy at an early stage. Steroids may work on different pathways from anti-VEGFs, and some patients may respond better to steroids than anti-VEGFs – we noted earlier that about 40% of eyes with DMO respond poorly to anti-VEGFs. Perhaps measuring the balance of VEGF versus cytokines in the eye (sample might be taken at time of first injection) might allow more personalised treatment. This is a research area meantime.

5.9 Conclusion

The evidence base for using fluocinolone in patients with chronic phakic DMO is not ideal, lacking data from an RCT in the patients identified in the NICE scope, who are insufficiently responsive to anti-VEGF drugs. However the ERG view is that data from several useful observational studies provide evidence that fluocinolone is effective in patients who have had a poor response to anti-VEGFs treatment.

6 REFERENCES

1. Becker C, Schneider C, Aballea S, Bailey C, Bourne R, Jick S, *et al.* Cataract in patients with diabetes mellitus-incidence rates in the UK and risk factors. *Eye (Lond)* 2018;**32**:1028-35. http://dx.doi.org/10.1038/s41433-017-0003-1

2. Scanlon PH, Foy C, Chen FK. Visual acuity measurement and ocular co-morbidity in diabetic retinopathy screening. *Br J Ophthalmol* 2008;**92**:775-8. http://dx.doi.org/10.1136/bjo.2007.128561

3. Dowler JG, Sehmi KS, Hykin PG, Hamilton AM. The natural history of macular edema after cataract surgery in diabetes. *Ophthalmology* 1999;**106**:663-8. http://dx.doi.org/10.1016/s0161-6420(99)90148-3

4. Feng Y, Zhu S, Skiadaresi E, McAlinden C, Tu R, Gao R, *et al.* Phacoemulsification Cataract Surgery With Prophylactic Intravitreal Bevacizumab For Patients With Coexisting Diabetic Retinopathy: A Meta-Analysis. *Retina* 2018; 10.1097/IAE.0000000000002221. http://dx.doi.org/10.1097/IAE.0000000000002221

5. Age-Related Eye Disease Study Research Group. The age-related eye disease study (AREDS) system for classifying cataracts from photographs: AREDS report no. 4. *Am J Ophthalmol* 2001;**131**:167-75.

6. NICE. *Glaucoma: diagnosis and management: NICE guideline [NG81]*. 2017. URL: https://www.nice.org.uk/guidance/ng81 (Accessed 15/01/2019).

7. Edmunds B, Thompson JR, Salmon JF, Wormald RP. The National Survey of Trabeculectomy. III. Early and late complications. *Eye (Lond)* 2002;**16**:297-303. http://dx.doi.org/10.1038/sj/eye/6700148

8. NICE. *Trabecular stent bypass microsurgery for open-angle glaucoma: Interventional procedures guidance [IPG575]*. 2017. URL: https://www.nice.org.uk/guidance/ipg575 (Accessed 01/03/2019).

9. Rolim de Moura C, Paranhos A, Jr., Wormald R. Laser trabeculoplasty for open angle glaucoma. *Cochrane Database Syst Rev* 2007; 10.1002/14651858.CD003919.pub2:Cd003919. http://dx.doi.org/10.1002/14651858.CD003919.pub2

10. El Sayyad F, Helal M, El-Kholify H, Khalil M, El-Maghraby A. Nonpenetrating deep sclerectomy versus trabeculectomy in bilateral primary open-angle glaucoma. *Ophthalmology* 2000;**107**:1671-4.

11. Burr J, Azuara-Blanco A, Avenell A, Tuulonen A. Medical versus surgical interventions for open angle glaucoma. *Cochrane Database Syst Rev* 2012; 10.1002/14651858.CD004399.pub3:Cd004399. http://dx.doi.org/10.1002/14651858.CD004399.pub3

12. Klein R, Klein BE, Moss SE, Cruickshanks KJ. The Wisconsin Epidemiologic Study of Diabetic Retinopathy: XVII. The 14-year incidence and progression of diabetic retinopathy and

associated risk factors in type 1 diabetes. *Ophthalmology* 1998;**105**:1801-15. http://dx.doi.org/10.1016/s0161-6420(98)91020-x

13. Stitt AW, Curtis TM, Chen M, Medina RJ, McKay GJ, Jenkins A, *et al.* The progress in understanding and treatment of diabetic retinopathy. *Prog Retin Eye Res* 2016;**51**:156-86. http://dx.doi.org/10.1016/j.preteyeres.2015.08.001

14. Ockrim Z, Yorston D. Managing diabetic retinopathy. *Bmj* 2010;**341**:c5400. http://dx.doi.org/10.1136/bmj.c5400

15. Moseley J. *Visual Function Endpoints: The Regulatory Perspective - EMA Scientific Advice*. 2011. URL: https://www.ema.europa.eu/documents/presentation/presentation-day-1-visual-function-endpoints-regulatory-perspective_en.pdf (Accessed 10/02/2019).

16. Williams R, Airey M, Baxter H, Forrester J, Kennedy-Martin T, Girach A. Epidemiology of diabetic retinopathy and macular oedema: a systematic review. *Eye (Lond)* 2004;**18**:963-83. http://dx.doi.org/10.1038/sj.eye.6701476

17. Yau JW, Rogers SL, Kawasaki R, Lamoureux EL, Kowalski JW, Bek T, *et al.* Global prevalence and major risk factors of diabetic retinopathy. *Diabetes Care* 2012;**35**:556-64. http://dx.doi.org/10.2337/dc11-1909

18. Idris I, Warren G, Donnelly R. Association between thiazolidinedione treatment and risk of macular edema among patients with type 2 diabetes. *Arch Intern Med* 2012;**172**:1005-11. http://dx.doi.org/10.1001/archinternmed.2012.1938

19. White NH, Sun W, Cleary PA, Danis RP, Davis MD, Hainsworth DP, *et al.* Prolonged effect of intensive therapy on the risk of retinopathy complications in patients with type 1 diabetes mellitus: 10 years after the Diabetes Control and Complications Trial. *Arch Ophthalmol* 2008;**126**:1707-15. http://dx.doi.org/10.1001/archopht.126.12.1707

20. Minassian DC, Owens DR, Reidy A. Prevalence of diabetic macular oedema and related health and social care resource use in England. *Br J Ophthalmol* 2012;**96**:345-9. http://dx.doi.org/10.1136/bjo.2011.204040

21. Romero-Aroca P. Targeting the pathophysiology of diabetic macular edema. *Diabetes Care* 2010;**33**:2484-5. http://dx.doi.org/10.2337/dc10-1580

22. Brown DM, Nguyen QD, Marcus DM, Boyer DS, Patel S, Feiner L, *et al.* Long-term outcomes of ranibizumab therapy for diabetic macular edema: the 36-month results from two phase III trials: RISE and RIDE. *Ophthalmology* 2013;**120**:2013-22. http://dx.doi.org/10.1016/j.ophtha.2013.02.034

23. NICE. *Ranibizumab for treating diabetic macular oedema: Technology appraisal guidance [TA274]*. 2013. URL: https://www.nice.org.uk/guidance/ta274 (Accessed 15/01/2019).

24. NICE. *Aflibercept for treating diabetic macular oedema: Technology appraisal guidance [TA346]*. 2015. URL: https://www.nice.org.uk/guidance/ta346 (Accessed 08/01/2019).

25. NICE. *Dexamethasone intravitreal implant for treating diabetic macular oedema: Technology appraisal guidance [TA349]*. 2015. URL: https://www.nice.org.uk/guidance/ta349 (Accessed 08/01/2019).
26. Virgili G, Parravano M, Evans JR, Gordon I, Lucenteforte E. Anti-vascular endothelial growth factor for diabetic macular oedema: a network meta-analysis. *Cochrane Database Syst Rev* 2018;**10**:Cd007419. http://dx.doi.org/10.1002/14651858.CD007419.pub6
27. Campochiaro PA, Brown DM, Pearson A, Ciulla T, Boyer D, Holz FG, *et al.* Long-term benefit of sustained-delivery fluocinolone acetonide vitreous inserts for diabetic macular edema. *Ophthalmology* 2011;**118**:626-35.e2.
28. Aiello LP, Beck RW, Bressler NM, Browning DJ, Chalam KV, Davis M, *et al.* Rationale for the diabetic retinopathy clinical research network treatment protocol for center-involved diabetic macular edema. *Ophthalmology* 2011;**118**:e5-14. http://dx.doi.org/10.1016/j.ophtha.2011.09.058
29. Schmidt-Erfurth U, Garcia-Arumi J, Bandello F, Berg K, Chakravarthy U, Gerendas BS, *et al.* Guidelines for the Management of Diabetic Macular Edema by the European Society of Retina Specialists (EURETINA). *Ophthalmologica* 2017;**237**:185-222. http://dx.doi.org/10.1159/000458539
30. Mitchell P, Bandello F, Schmidt-Erfurth U, Lang GE, Massin P, Schlingemann RO, *et al.* The RESTORE study: ranibizumab monotherapy or combined with laser versus laser monotherapy for diabetic macular edema. *Ophthalmology* 2011;**118**:615-25. http://dx.doi.org/10.1016/j.ophtha.2011.01.031
31. Egan C, Zhu H, Lee A, Sim D, Mitry D, Bailey C, *et al.* The United Kingdom Diabetic Retinopathy Electronic Medical Record Users Group, Report 1: baseline characteristics and visual acuity outcomes in eyes treated with intravitreal injections of ranibizumab for diabetic macular oedema. *Br J Ophthalmol* 2017;**101**:75-80. http://dx.doi.org/10.1136/bjophthalmol-2016-309313
32. Elman MJ, Qin H, Aiello LP, Beck RW, Bressler NM, Ferris FL, III, *et al.* Intravitreal Ranibizumab for Diabetic Macular Edema with Prompt versus Deferred Laser Treatment: Three-Year Randomized Trial Results. *Ophthalmology* 2012;**119**:2312-8. http://dx.doi.org/10.1016/j.ophtha.2012.08.022
33. Banaee T, Ashraf M, Conti FF, Singh RP. Switching Anti-VEGF Drugs in the Treatment of Diabetic Macular Edema. *Ophthalmic Surg Lasers Imaging Retina* 2017;**48**:748-54. http://dx.doi.org/10.3928/23258160-20170829-10
34. Ferris FL, 3rd, Maguire MG, Glassman AR, Ying GS, Martin DF. Evaluating Effects of Switching Anti-Vascular Endothelial Growth Factor Drugs for Age-Related Macular Degeneration and Diabetic Macular Edema. *JAMA Ophthalmol* 2017;**135**:145-9. http://dx.doi.org/doi:10.1001/jamaophthalmol.2016.4820
35. Mieno H, Yoneda K, Yamazaki M, Sakai R, Sotozono C, Fukui M. The Efficacy of Sodium-Glucose Cotransporter 2 (SGLT2) inhibitors for the treatment of chronic diabetic macular oedema in

vitrectomised eyes: a retrospective study. *BMJ Open Ophthalmol* 2018;**3**:e000130. http://dx.doi.org/10.1136/bmjophth-2017-000130

36. Campochiaro PA, Hafiz G, Shah SM, Bloom S, Brown DM, Busquets M, *et al.* Sustained ocular delivery of fluocinolone acetonide by an intravitreal insert. *Ophthalmology* 2010;**117**:1393-9.e3.

37. FDA Center for Drug Evaluation and Research. *Medical Review Iluvien*. 2014. URL: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/201923Orig1s000MedR.pdf (Accessed 12/02/2019).

38. Cunha-Vaz J, Ashton P, Iezzi R, Campochiaro P, Dugel PU, Holz FG, *et al.* Sustained delivery fluocinolone acetonide vitreous implants: Long-term benefit in patients with chronic diabetic macular edema. *Ophthalmology* 2014;**121**:1892-903.

39. Yang Y, Bailey C, Holz FG, Eter N, Weber M, Baker C, *et al.* Long-term outcomes of phakic patients with diabetic macular oedema treated with intravitreal fluocinolone acetonide (FAc) implants. *Eye (Basingstoke)* 2015;**29**:1173-80.

40. Campochiaro PA, Wykoff CC, Kapik B, Green KE. Sustained intraocular delivery of fluocinolone acetonide slows progression of diabetic retinopathy. *Investigative Ophthalmology and Visual Science* 2016;**57 (12)**:3283.

41. Bailey C, Chakravarthy U, Lotery A, Menon G, Talks J. Real-world experience with 0.2 mug/day fluocinolone acetonide intravitreal implant (ILUVIEN) in the United Kingdom. *Eye (Basingstoke)* 2017;**31**:1707-15.

42. West Essex CCG. *Pathway for treatment of adults with Diabetic Macular Oedema*. 2017. URL: https://westessexccg.nhs.uk/your-health/medicines-optimisation-and-pharmacy/high-cost-drug-policies-and-pro-formas/ophthalmology/diabetic-macular-oedema/218-diabetic-macular-oedema-pathway/file (Accessed 10/02/2019).

43. Bottoni F, Margaron P, Dunger-Baldauf C, Staurenghi G. Benefit of continuing ranibizumab 0.5 mg treatment in diabetic macular edema patients with delayed initial visual response. *Investigative Ophthalmology and Visual Science* 2015;**56 (7)**:1737.

44. Raman V. A cost analysis comparing continued 3-year aflibercept monotherapy versus a switch from aflibercept to the fluocinolone acetonide intravitreal implant in phakic patients with chronic diabetic macular edema. *Expert Review of Ophthalmology* 2018;**13**:299-307. http://dx.doi.org/10.1080/17469899.2018.1523720

45. Quhill F, Beiderbeck A. Cost Advantage of Fluocinolone Acetonide Implant (ILUVIEN®) versus Ranibizumab in the Treatment of Chronic Diabetic Macular Oedema. *Global & Regional Health Technology Assessment* 2017;**4**:grhta.5000268. http://dx.doi.org/10.5301/grhta.5000268

46. Carneiro A. A retrospective cohort study of the clinical outcomes of ILUVIEN in four ophthalmology centres in Portugal: The results of the ILUVIEN clinical evidence study in the Portugal (ICEPT). 2018

47. Holden SE, Currie CJ, Owens DR. Evaluation of the clinical effectiveness in routine practice of fluocinolone acetonide 190 micro g intravitreal implant in people with diabetic macular edema. *Current Medical Research and Opinion* 2017;**33**:5-17.

48. Holden SE, Kapik B, Beiderbeck AB, Currie CJ. Comparison of data characterising the clinical effectiveness of the fluocinolone intravitreal implant (ILUVIEN) in patients with diabetic macular edema from the real world, non-interventional ICE-UK study and the FAME randomized controlled trials. *Curr Med Res Opin* 2018; 10.1080/03007995.2018.1560779:1. http://dx.doi.org/10.1080/03007995.2018.1560779

49. Chakravarthy U, Taylor SR, Koch FHJ, Castro De Sousa JP, Bailey C. Changes in intraocular pressure after intravitreal fluocinolone acetonide (ILUVIEN): Real-world experience in three European countries. *British Journal of Ophthalmology* 2018.

50. Figueira J, Henriques J, Amaro M, Rosas V, Alves D, Cunha-Vaz J. A Nonrandomized, Open-Label, Multicenter, Phase 4 Pilot Study on the Effect and Safety of ILUVIEN in Chronic Diabetic Macular Edema Patients Considered Insufficiently Responsive to Available Therapies (RESPOND). *Ophthalmic Research* 2017;**57**:166-72.

51. Massin P, Erginay A, Dupas B, Couturier A, Tadayoni R. Efficacy and safety of sustained-delivery fluocinolone acetonide intravitreal implant in patients with chronic diabetic macular edema insufficiently responsive to available therapies: A real-life study. *Clinical Ophthalmology* 2016;**10**:1257-64.

52. Mansour S. Initial efficacy, anatomic and safety outcomes after fluocinolone acetonide 0.2 ug/day implant in patients with DME: The PALADIN real-world observational study. *Investigative Ophthalmology and Visual Science Conference* 2017;**58**.

53. Eaton A, Koh SS, Jimenez J, Riemann CD. The USER Study: A Chart Review of Patients Receiving a 0.2 microg/day Fluocinolone Acetonide Implant for Diabetic Macular Edema. *Ophthalmol Ther* 2018; 10.1007/s40123-018-0155-5. http://dx.doi.org/10.1007/s40123-018-0155-5

54. Augustin AJ. Retro-IDEAL study - results from real-world practice show that after substantial amounts of prior treatment with anti-VEGF and other therapies a single ILUVIEN (fluocinolone acetonide; FAc) implant leads to sustained improvements lasting up to 36 months. *Investigative Ophthalmology & Visual Science* 2018;**59**:4856-.

55. Haig J, Barbeau M, Ferreira A. Cost-effectiveness of ranibizumab in the treatment of visual impairment due to diabetic macular edema. *J Med Econ* 2016;**19**:663-71. http://dx.doi.org/10.3111/13696998.2016.1154566

56. Currie CJ, Holden SE, Berni E, Owens DR. Evaluation of the clinical effectiveness of fluocinolone acetonide 190 micro g intravitreal implant in diabetic macular edema: a comparison between study and fellow eyes. *Current Medical Research and Opinion* 2017;**33**:19-31.
57. Holden SE, Currie CJ, Owens DR. Health-economic evaluation of fluocinolone acetonide 190 micro g implant in people with diabetic macular edema. *Current Medical Research and Opinion* 2017;**33**:45-52.
58. Pochopien M, Beiderbeck A, McEwan P, Zur R, Toumi M, Aballea S. Cost-effectiveness of fluocinolone acetonide implant (ILUVIEN(R)) in UK patients with chronic diabetic macular oedema considered insufficiently responsive to available therapies. *BMC Health Serv Res* 2019;**19**:22. http://dx.doi.org/10.1186/s12913-018-3804-4
59. Neubauer AS, Haritoglou C, Ulbig MW. [Cost Comparison of Licensed Intravitreal Therapies for Insufficiently Anti-VEGF Responding Fovea Involving Diabetic Macular Edema in Germany]. *Klin Monbl Augenheilkd* 2019;**236**:180-91. http://dx.doi.org/10.1055/s-0043-125076
60. 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 Ophthalmol* 2014;**132**:310-8. http://dx.doi.org/10.1001/jamaophthalmol.2013.7639
61. Brazier J, Muston D, Konwea H, Power GS, Barzey V, Lloyd A, *et al.* Evaluating the Relationship Between Visual Acuity and Utilities in Patients With Diabetic Macular Edema Enrolled in Intravitreal Aflibercept Studies. *Invest Ophthalmol Vis Sci* 2017;**58**:4818-25. http://dx.doi.org/10.1167/iovs.17-21945
62. NICE. *Type 2 diabetes in adults: management: NICE guideline [NG28]*. 2015. URL: https://www.nice.org.uk/guidance/ng28 (Accessed 28/02/2019).
63. Norregaard JC, Thoning H, Andersen TF, Bernth-Petersen P, Javitt JC, Anderson GF. Risk of retinal detachment following cataract extraction: results from the International Cataract Surgery Outcomes Study. *Br J Ophthalmol* 1996;**80**:689-93.
64. Chan E, Mahroo OA, Spalton DJ. Complications of cataract surgery. *Clin Exp Optom* 2010;**93**:379-89. http://dx.doi.org/10.1111/j.1444-0938.2010.00516.x
65. Preis SR, Hwang SJ, Coady S, Pencina MJ, D'Agostino RB, Sr., Savage PJ, *et al.* Trends in all-cause and cardiovascular disease mortality among women and men with and without diabetes mellitus in the Framingham Heart Study, 1950 to 2005. *Circulation* 2009;**119**:1728-35. http://dx.doi.org/10.1161/circulationaha.108.829176
66. Christ SL, Lee DJ, Lam BL, Zheng DD, Arheart KL. Assessment of the effect of visual impairment on mortality through multiple health pathways: structural equation modeling. *Invest Ophthalmol Vis Sci* 2008;**49**:3318-23. http://dx.doi.org/10.1167/iovs.08-1676
67. Brown GC. Vision and quality-of-life. *Trans Am Ophthalmol Soc* 1999;**97**:473-511.
68. Meads C, Hyde C. What is the cost of blindness? *Br J Ophthalmol* 2003;**87**:1201-4.

69. McCrone P DS, Patel A, Knapp M, Lawton-Smith S. *Paying the Price: the Cost of Mental Health Care in England to 2026*. 2008. URL: https://www.kingsfund.org.uk/sites/default/files/Paying-the-Price-the-cost-of-mental-health-care-England-2026-McCrone-Dhanasiri-Patel-Knapp-Lawton-Smith-Kings-Fund-May-2008_0.pdf (Accessed 04/03/2019).

70. Competition and Markets Authority. *Care homes market study: summary of final report*. 2017. URL: https://www.gov.uk/government/publications/care-homes-market-study-summary-of-final-report (Accessed 04/03/2019).

71. 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. http://dx.doi.org/https://dx.doi.org/10.1111/j.1524-4733.2009.00527.x

72. Foot B, MacEwen C. Surveillance of sight loss due to delay in ophthalmic treatment or review: frequency, cause and outcome. *Eye* 2017;**31**:771. http://dx.doi.org/10.1038/eye.2017.1

73. Royal College of Ophthalmologists. *Workforce Census* 2018. URL: https://rcophth.ac.uk/publications/workforce-census-2018/ (Accessed 15/02/2019).

74. Han DP, Heuer DK. Intravitreal corticosteroid therapy: putting the problem of glaucoma in perspective. *Arch Ophthalmol* 2012;**130**:380-2. http://dx.doi.org/10.1001/archopthalmol.2011.1097

75. Augustin A, Yang Y, Diestelhorst M, Massin P. Changing Perspectives in Diabetic Macular Oedema – Recognising and Understanding Chronic Diabetic Macular Oedema. *European Ophthalmic Review* 2014;**8**:145–51. http://dx.doi.org/10.17925/EOR.2014.08.02.145

7 APPENDICES

7.1 Appendix 1. Risk of bias table for FAME study

Table 32. FAME study risk of bias

Criteria	Description	Judgement
Adequate sequence generation	Patients were randomized in a 1:2:2 ratio stratified by baseline BCVA and site. Patients were enrolled in the study using a computer generated schedule and an integrated voice recognition system.	Yes
Allocation concealment	Patients were enrolled in the study using a computer generated schedule and an integrated voice recognition system. Injectors and packaging identical	Yes
Masking	Double-blinding. To preserve masking, two investigators were used. One investigator performed the treatments and the other masked investigator performed all assessments and determined retreatment eligibility.	Yes
Incomplete outcome data addressed	Intention to treatment analysis was used for the primary endpoint and missing data were imputed by last observation carried forward. Adequate description of loss to follow-up, withdrawals and adverse events.	Yes
Free of selective reporting	All pre-specified outcomes were reported	Yes
Groups comparable at baseline		Yes
Sample size calculation	180:180:90 subjects were expected to provide 89% power to detect a difference of 16% including a projected 10% dropout rate.	Yes

One possibility is that the fluocinolone insert might be detected as a floater and thereby unmask patients in the fluocinolone group, but floaters are common in people of the age group recruited, and in those with diabetic retinopathy. So unmasking is unlikely to be a problem.

7.2 Appendix 2. Detailed tables from FAME trial for patients phakic at baseline

Table 33. Characteristics of participants in FAME

Baseline characteristics	**Sham Chronic**	**0.2 µg/day FAc Chronic Phakic - Pseudo**	**0.2 µg/day FAc Chronic Phakic – Phakic**
FAME A + B (n= NA)	(n=112)	(n=97)	(N=17)
Mean Age	62.9 years	60.6 years	63.6 years
Continent (%)			
North America	79	69	10
European Union	9	14	2
India	24	14	5
Mean Duration of DMO (SD)	5.4 (4.24) years	5.2 (3.29)	4.7 (3.10)
Diabetes Type (%)			
Type 1	8.9	9.3	5.9
Type 2	90.2	87.6	94.1
Uncertain	0.9	3.1	0.0
Duration of diabetes, mean (SD)			
Type 1	24.9 (8.32)	30.6 (10.68)	34.0 (NA)
Type 2	17.2 (8.02)	15.9 (8.0)	14.8 (7.17)
All	17.9 (8.28)	17.2 (9.20)	15.9 (8.36)
Pre-existing cataract (%)			
Yes	42.0	74.2	94.1
No	12.5	23.7	5.9
N/A	45.5	2.1	0.0
Baseline HbA1c, mean (SD)	7.7% (1.51)	7.9 (1.59)	8.4 (2.02)

HbA1c level (%)			
≤5%	0.0	0.0	0.0
>5-7%	30.4	30.9	29.4
>7-9%	42.9	43.3	29.4
>9-11%	8.0	12.4	23.5
>11-13%	2.7	5.2	11.8
>13%	0.9	0.0	0.0
Missing	15.2	8.2	5.9
Study eye lens status (%)			
Pseudophakic	41.1%	0.0	0.0
Phakic	58.9%	100.0	100.0
Years since last steroid injection (not fluocinolone)			
N	23	11	1
Mean (SD)	1.9 (0.85)	2.3 (0.65)	2 (NA)
Years since last Anti-VEGF treatment			
N	6	4	1
Mean (SD)	1.2 (0.41)	1.5 (1.0)	1 (NA)

Table 34. Proportion of patients with a DMO duration ≥3 years who had a ≥15 letter increase from baseline in BCVA

	0.2 µg/day FAc Chronic Phakic – Pseudo (N=97)[1]	**0.2 µg/day FAc Chronic Phakic – Phakic (N=17)[2]**	**Sham Chronic (N=112)**	**Difference[1] (95% CI)**	**P[1] value**	**Difference[2] (95%CI)**	**P[2] value**
	% with ≥15 Letter Improvement in BCVA						
Month 18	20.6	0.0	9.8	(-20.6, -1.0)	0.051	(4.3, 15.3)	0.168
Month 24	43.3	5.9 (1 patient)	13.4	(-41.6, -18.2)	<0.001	(-5.3, 20.4)	0.331
Month 30	46.4	5.9	10.7	(-47.1, -24.2)	<0.001	(-7.7, 17.4)	0.507
Month 36	42.3	0.0	13.4	(-40.6, -17.2)	<0.001	(7.1, 19.7)	0.105

Note: Differences and P-values relate to comparisons versus sham chronic group.

The ERG finds some figures in the above table puzzling. Column 5 gives differences that do not seem compatible with the other columns.

Table 35 Distribution of BCVA in the study eye at Month 36 by treatment group

Category	**Sham (N=112)**	**0.2 µg/day FAc Chronic Phakic – Pseudo (N=97)[1]**	**0.2 µg/day FAc Chronic Phakic – Phakic (N=17)[2]**	**Difference[1] (95% CI)**	**P1 value**	**Difference[2] (95% CI)**	**P2 value**
≥15 letter decrease, %	9.8	6.2	35.3	NA	NA	NA	NA
10-14 letter decrease, %	6.3	1.0	5.9	NA	NA	NA	NA
5-9 letter decrease, %	9.8	4.1	11.8	NA	NA	NA	NA
No change group (-4 to +4 letters) , %	27.7	9.3	29.4	NA	NA	NA	NA
5-9 letter increase, %	17.0	16.5	0.0	NA	NA	NA	NA
10-14 letter increase, %	16.1	20.6	17.6	NA	NA	NA	NA
≥15 letter increase, %	13.4	42.3	0.0	(-40.6, -17.2)	<0.001	(7.1, 19.7)	0.105

Note: Differences and P-values relate to comparisons versus sham chronic group. Not applicable (NA) as the original analysis summarized the distribution of BVCA, but did not include the p-values and CI.

Table 36. Summary of off-protocol treatments for DMO by type of therapy in subjects with chronic DMO (integrated FAME studies)

Type of Therapy	Treatment Group		
	Sham Chronic (N = 112)	**0.2 µg/day FAc Chronic Phakic – Pseudo (N=97)**	**0.2 µg/day FAc Chronic Phakic – Phakic (N=17)**
Intravitreal steroids	27 (24.1%)	9 (9.3%)	0 (0.0%)
P-value[i)]	NA	0.004	0.022
Posterior sub-Tenon's steroids	5 (4.5%)	1 (1.0%)	0 (0.0%)
P-value	NA	0.086	0.336
Anti-VEGF therapy	17 (15.2%)	4 (4.1%)	1 (5.9%)
P-value	NA	0.008	0.301
Vitrectomies	9 (8.0%)	4 (4.1%)	0 (0.0%)
P-value	NA	0.249	0.227

Table 37. Adverse events: Duration of Diabetic Macular Oedema ≥ 3 years

	Sham Chronic N=112	**0.2 µg/day FAc Chronic Phakic – Pseudo (N=97)**	**0.2 µg/day FAc Chronic Phakic – Phakic (N=NA)**
New cataract			
Study eye, %	30.4	80.4	29.4
Non-study eye, %	24.1	39.2	11.8
Posterior capsule opacification			
Study eye, %	3.6	8.2	0.0
Non-study eye, %	2.7	2.1	0.0
Vitreous haemorrhage			

Study eye, %	13.4	7.2	0.0
Non-study eye, %	13.4	13.4	0.0
Retinal detachment			
Study eye, %	2.7	1.0	0.0
Non-study eye, %	0.0	0.0	0.0
Cataract operation			
Study eye, %	21.4	100.0	0.0
Non-study eye, %	19.6	38.1	11.8

Table 38. IOP related events for duration of diabetic macular oedema ≥ 3 years

	Sham Chronic N=112	**0.2 µg/day FAc Chronic Phakic – Pseudo (N=97)**	**0.2 µg/day FAc Chronic Phakic – Phakic (N=17)**
Elevation considered an adverse event			
Study eye	16 (14.3%)	33 (34.0%)	4 (23.5%)
Non-study eye	13 (11.6%)	14 (14.4%)	0 (0.0%)
Elevation increase of 12 or more mmHG			
Study eye	12 (10.7%)	26 (26.8%)	3 (17.6%)
Non-study eye	4 (3.6%)	9 (9.3%)	0 (0.0%)
IOP lowering medications in study eye			
Any IOP lowering medication	17 (15.2%)	38 (39.2%)	4 (23.5%)
1 IOP lowering medication [1]	11 (9.8%)	20 (20.6%)	3 (17.6%)

2 IOP-lowering medications [2]	3 (2.7%)	9 (9.3%)	1 (5.9%)
≥3 IOP lowering medication [2]	3 (2.7%)	9 (9.3%)	0 (0.0%)
Elevation to over 25 mmHG			
Study eye	13 (11.6%)	33 (34.0%)	4 (23.5%)
Non-study eye	10 (8.9%)	15 (15.5%)	0 (0.0%)
Elevation to over 30 mmHG			
Study eye	6 (5.4%)	13 (13.4%)	2 (11.8%)
Non-study eye	3 (2.7%)	5 (5.2%)	0 (0.0%)
Trabeculoplasty			
Study eye	0 (0.0%)	1 (1.0%)	0 (0.0%)
Non-study eye	0 (0.0%)	1 (1.0%)	0 (0.0%)
Trabeculectomy			
Study eye	0 (0.0%)	4 (4.1%)	0 (0.0%)
Non-study eye	0 (0.0%)	0 (0.0%)	0 (0.0%)
Vitrectomy performed for elevated IOP			
Study eye	0 (0.0%)	0 (0.0%)	0 (0.0%)
Non-study eye	0 (0.0%)	0 (0.0%)	0 (0.0%)

Table 39. FAME Chronic Phakics*

Baseline (n= 180)		**3 years (n= 180)**	
		% with any cataracts	**% with cataract extraction**
No cataract (%)	NA	132 (73.3%)	121 (67.2%)

Note: [1]Investigators did not "grade" cataract, they only entered cataract data as adverse event data, and then, often did not enter the cataract formation as an adverse event, but only entered the cataract surgery.

7.3 Appendix 3. Calculation of duration

A9. Please clarify how duration of DMO was determined in the FAME trial?

Response:

In the analyses submitted to Health Authorities to support the registration of the fluocinolone implant an algorithm was employed to calculate the duration of DMO at entry into the phase III clinical trial. This algorithm used the date of diagnosis, provided by the investigator upon randomising a subject into the trial, and the date of randomisation into the trial. Based on these two dates, the duration of DMO at baseline was determined as:

Year of Randomisation minus Year of Diagnosis plus One

This algorithm addresses two goals, which have significant regulatory importance. First, it includes all subjects randomised, where, even if a subject were randomised in the same year as their diagnosis, the duration of DMO would still be one year. Therefore, the duration of disease would not be zero, and the subject would be included in the dataset. Secondly, it creates bias only toward longer duration disease.

While there is a significant amount of evidence linking severity of diabetic disease with duration of disease, it is not possible to definitively point to a specific time where the "balance" in the microenvironment shifts. That is, as the role of inflammation grows, one would expect that a point may be reached where "harm or damage" begins to accumulate in the microenvironment. This point will be different for every patient. Once the balance has been tipped, the accumulation of damage will be such that new factors become important for consideration in the treatment of the disease which may not have been as important earlier in the disease.

Efficacy Analysis Using an Alternate Method of Calculation of Duration of DMO

With this perspective regarding the algorithm employed to support the marketing application, it is relevant to consider another algorithm as the use of the FAc implant in clinical practice is initiated. This algorithm most closely reflects the exact date reported by retina specialists, and serves as a sensitivity analysis to the effects of methodology for assessing duration of disease.

Thus, based on this method, which accounts for every subject enrolled in the FAME studies, the median duration of diagnosis of DMO was 1.73 years.

Using the original method, the median duration of DMO was 3 years, with 416 subjects having a duration < the median, and 536 subjects having a duration ≥ the median. Based on the new approach, 475 subjects fall below the median and 477 subjects fall above or equal to the median. A concordance/ discordance analysis of subjects above and below the median using these two methods is presented in the table below.

Kappa is a measure of agreement ranging from -1 (complete discordance) to +1 (complete concordance). A value of 0.8508 represents very high agreement. The p-value confirms that we can reject the null hypothesis of no agreement, i.e., there is agreement. A significant number of subjects stayed in their original category. This indicates that the assignment to DMO subgroup is fairly insensitive to the method used in calculating the duration of DMO.

Concordance/Discordance Analysis of Subjects Above and Below the Median by Method of Calculation of Duration

Original Method	**New Method**	
	< 1.73 years	**≥ 1.73 years**
< 3 years	410 (43.1%)	6 (0.6%)
≥ 3 years	65 (6.8%)	471 (49.5%)
kappa		0.8508
p-value		<0.0001

Using this method for analysis of the primary outcomes in FAME A and FAME B, the relationships between sham and the fluocinolone implant groups for subjects with duration of DMO ≥1.73 years and <1.73 years at randomisation are the same as those found for using the initial method for calculation of duration of DMO.

7.4 Appendix 4. Regression analysis on the probability of cataract surgery

Methods

The binomial model assessing the **probability of cataract surgery** was estimated by the Generalized Estimating Equations (GEE) method using Proc Genmod in SAS.

First, we estimated a model including all available variables as main effects and no interactions. Additional models were then tested by removing non-significant variables one by one, starting from the variable with the highest p-value (Table 1). A variable was considered as non-significant if the associated p-value was >0.05.

Then, we applied the same approach starting from a model with all variables as main effects and interactions between treatments and all other variables (Table 2). Main effects were not removed when the p-value for the corresponding interactions were < 0.05.

Results

The model with the lowest QIC is the model 14, which includes the following independent variables:

- Gender
- Treatment
- BCVA level at last visit
- Interaction between treatment and gender
- Interaction between treatment and BCVA level at last visit

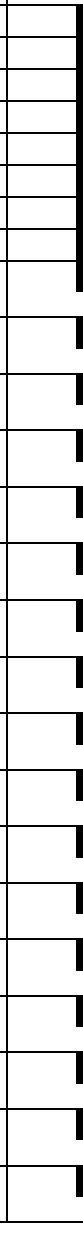
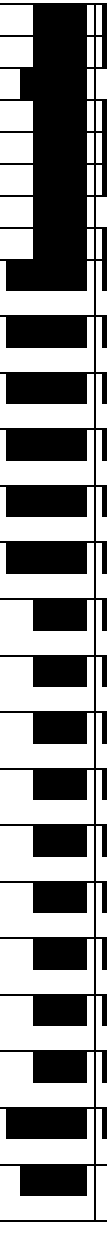
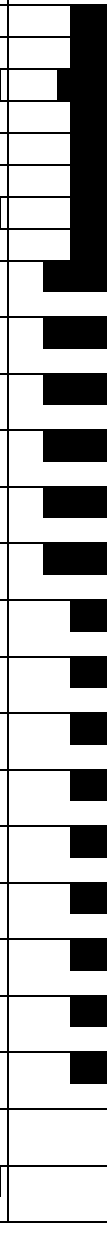
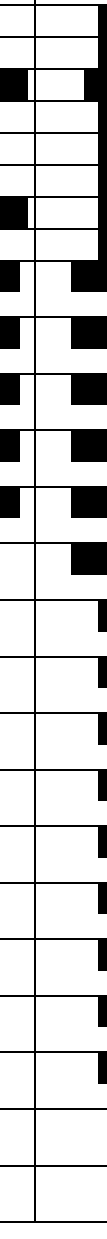
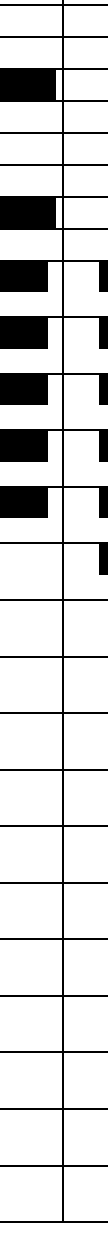
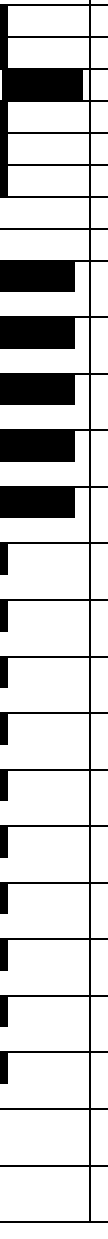
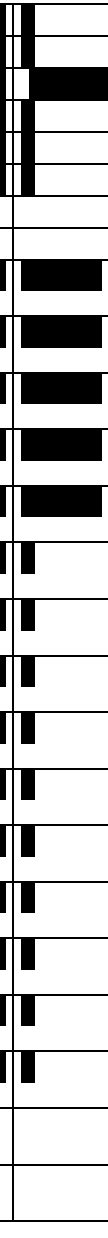
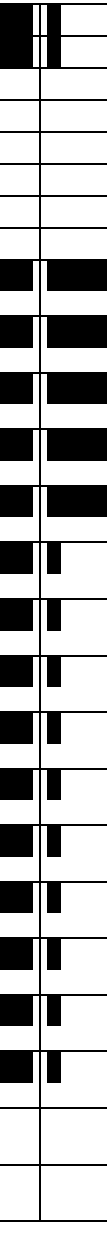
However, this model was not finally selected because it led to overfitting of data, as shown in Table 3 (probabilities by treatment group, gender and BCVA level at last visit). Therefore, we decided to select a model without interactions. The model 6 (with gender and BCVA at last visit) has the lower QIC, among models without interaction. However, we opted for model 1 because the effect of gender in model was not statistically significant (p=0.1792), and it seemed likely from a clinical perspective that an effect of treatment, and this effect could impact the results. Sensitivity analyses showed that CEA results were in fact similar whether we used model 1 or model 7(without treatment effect).

Table 1. Main effect models assessing the probability of cataract surgery (regression 1)

		model 1 (selected)		model 2		model 3		model 4		model 5		model 6		model 7	
QIC (smaller is better)															
QICu (smaller is better)															
VARIABLES	*level*	*estimate*	*p*	*estimate*	*p*	*estimate*	*p*	*estimate*	*p*	*estimate*	*P*	*estimate*	*p*	*estimate*	*p*
Intercept															
Sex															
Sex															
Center point retinal thickness at baseline															
Center point retinal thickness at baseline															
Age															
Treatment Group															
Treatment Group															
Previous BCVA level															
Previous BCVA level															
Previous BCVA level															
Previous BCVA level															
Previous BCVA level															
Previous BCVA level															
Previous BCVA level															
Previous BCVA level															
Study Visit Number															
Study Visit Number															
Study Visit Number															
Study Visit Number															
Study Visit Number															
Study Visit Number															
Study Visit Number															
Study Visit Number															
Study Visit Number															
Study Visit Number															
Study Visit Number															
Study Visit Number															

Table 2. Models with treatment interactions assessing the probability of cataract surgery (regression 1)

		model 8		model 9		model 10		model 11		model 12		model 13		**model 14**	
QIC (smaller is better)															
QICu (smaller is better)															
VARIABLES	*Level*	*estimate*	*p*	*estimate*	*p*	*estimate*	*p*	*estimate*	*p*	*estimate*	*p*	*estimate*	*p*	*estimate*	*p*
Intercept															
Sex															
Sex															
Center point retinal thickness at baseline (CPT)															
Center point retinal thickness at baseline (CPT)															
Age															
Treatment Group															
Treatment Group															
Previous BCVA level															
Previous BCVA level															
Previous BCVA level															
Previous BCVA level															
Previous BCVA level															
Previous BCVA level															
Previous BCVA level															
Previous BCVA level															
Study Visit Number															
Study Visit Number															
Study Visit Number															
Study Visit Number															
Study Visit Number															
Study Visit Number															
Study Visit Number															
Study Visit Number															
Study Visit Number															
Study Visit Number															
Study Visit Number															
Sex*Treatment Group															
Sex*Treatment Group															

		model 8	model 9	model 10	model 11	model 12	model 13	**model 14**
Sex*Treatment Group								
Sex*Treatment Group								
Treatment Group*CPT								
Treatment Group*CPT								
Treatment Group*CPT								
Treatment Group*CPT								
Age*Treatment Group								
Age*Treatment Group								
Treatment group*Previous BCVA level								
Treatment group*Previous BCVA level								
Treatment group*Previous BCVA level								
Treatment group*Previous BCVA level								
Treatment group*Previous BCVA level								
Treatment group*Previous BCVA level								
Treatment group*Previous BCVA level								
Treatment group*Previous BCVA level								
Treatment group*Previous BCVA level								
Treatment group*Previous BCVA level								
Treatment group*Previous BCVA level								
Treatment group*Previous BCVA level								
Treatment group*Previous BCVA level								
Treatment Group*Study Visit Number								
Treatment Group*Study Visit Number								

		model 8		model 9		model 10		model 11		model 12		model 13		**model 14**	
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												

		model 8		model 9		model 10		model 11		model 12		model 13		**model 14**	
Treatment Group*Study Visit Number	[redacted]	[redacted]	[redacted]												

Table 3. Probabilities of cataract surgery among phakic patients with cataract, as predicted by model 14.

Patient subgroup	BCVA level at previous visit							
	86-100 letters	76 -85 letters	66-75 letters	56-65 letters	46-55 letters	36-45 letters	26-35 letters	0-25 letters
Sham, females	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ILUVIEN, females	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sham, males	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ILUVIEN, males	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

7.5 Appendix 5. Regression analyses on transition probabilities between BCVA levels

Methods

The model assessing the **probability of gaining and losing letters (BCVA change) for all patients** was a multinomial model with random intercept. It was estimated using Proc Glimmix in SAS, with multinomial distribution, glogit (generalized logit) link function, random intercept with variance components matrix (VC) as covariance structure and Laplace estimation method (other methods did not converge).

The following table shows available variables.

Table 1. Variables in regression analysis

Variable	Label	Definition
BGROUP	BCVA level, num	1 = 86-100 letters, 2 = 76 -85 letters, 3 = 66-75 letters, 4 = 56-65 letters, 5 = 46-55 letters, 6 = 36-45 letters, 7 = 26-35 letters, 8 = 0-25 letters
TRT	Treatment Group, num	1 = Sham control, 2 = 0.2 μg/d FAc
CATARACT	Lens status, num	1 = phakic w/o cataract, 2 = phakic w/ cataract, 3 = cataract surgery within the last 90 days prior to BCVA assessment, 4 = pseudophakic/cataract surgery >90 days prior to BCVA assessment
CPT	Center point retinal thickness at baseline, num	1 = < 400μm, 2 = ≥400 μm
VISITNUM	Study Visit Number, num	1 = baseline 7 = month 3 8 = month 6 9 = month 9 10 = month 12 11 = month 15 12 = month 18 13 = month 21 14 = month 24 15 = month 27 16 = month 30 17 = month 33 18 = month 36

Variable	Label	Definition
LAG_BGROUP	Previous BCVA level, num	See BGROUP
CHANGE	Change from previous BCVA level, num	2 = improvement of 2 levels or more, 1 = improvement of 1 level, 0 = stable/no change, -1 = worsening of 1 level, -2 = worsening of 2 levels or more
AGE	Subject age, num	Continuous variable
DMSEX	Subject gender, char	'1 ' = Male, '2 ' = Female

First, we estimated a model including all available variables as main effects and no interactions. Additional models were then tested by removing non-significant variables one by one, starting from the variable with the highest p-value. Then, we estimated a model will all available variables as main effects and interaction terms including:

- Interaction between treatment and CPT
- Interaction between treatment and lens status
- Interactions between treatment and visit number / dummy for first visit
- Interaction between treatment and BCVA level at last visit
- Interactions between visit number / dummy for visit and BCVA level at last visit

Different variance-covariance structures were tested for random effects for selected models.

Results

Table 2 presents the p-values for models in which time is represented by a dummy variable only, allowing for different probabilities before and after 3 months. Table 3 presents the p-values for models in which time is represented by a categorical variable allowing for different probabilities at all visits. The models with a dummy variable only for time (Table 4) were better in terms of AIC and BIC.

The best model in terms of AIC was model 9, which includes the following variables:

- Age
- Treatment
- Lens status
- CPT (as binary variable)
- Visit (as binary variable: 3 months, after 3 months)

- Interaction of treatment with lens status
- Interaction of treatment with visit.

The regression model used in the CEA model is model 1, which also includes an interaction between visit and BCVA level at last visit. The reason for keeping this interaction is based on the selection of regression model for a cost-effectiveness analyses previously performed for pseudophakic patients.[12] The visual inspection of BCVA curves over time had suggested that this interaction should be included in the model, at the predicted curves better fitted the curves directly obtained from the trial.

We also tested different variance-covariance structures for random effects, and the model with variance components (VC) structure was best in terms of AIC (Table 5).

[12] Pochopien M, Beiderbeck A, McEwan P, Zur R, Toumi M, Aballéa S. Cost-effectiveness of fluocinolone acetonide implant (ILUVIEN®) in UK patients with chronic diabetic macular oedema considered insufficiently responsive to available therapies. BMC Health Serv Res. 2019 Jan 9;19(1):22.

Table 2.Models assessing BCVA change (regression 3) with study visit 7

	model selected (model 1)	model 2	model 3	model 4	model 5	model 6	model 7	model 8	**model 9**
AIC (smaller is better)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BIC (smaller is better)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Variables	*p-values presented for included variables*								
Age		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sex		[redacted]			[redacted]	[redacted]			
Treatment Group	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Study visit 7 (month 3)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Center point retinal thickness at baseline (CPT)	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*CPT					[redacted]	[redacted]	[redacted]		
Treatment Group*Lens status	[redacted]				[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment group*Previous BCVA level					[redacted]				
Study visit 7 (month 3)* Treatment group	[redacted]				[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Study visit 7 (month 3)* Previous BCVA level	[redacted]				[redacted]	[redacted]	[redacted]	[redacted]	

Table 3. Models assessing BCVA change (regression 3) with visit number

	model 10	model 11	model 12	model 13	model 14	model 15*	model 16	model 17
AIC (smaller is better)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BIC (smaller is better)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Variables	*p-values presented for included variables*							
Age	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]
Sex	[redacted]			[redacted]	[redacted]			
Treatment Group	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]
Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]
Study Visit Number	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]
Center point retinal thickness at baseline (CPT)	[redacted]	[redacted]		[redacted]	[redacted]		[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]		[redacted]	[redacted]
Treatment Group*CPT				[redacted]	[redacted]			
Treatment Group*Lens status				[redacted]	[redacted]		[redacted]	[redacted]
Study Visit Number* Treatment Group				[redacted]	[redacted]		[redacted]	[redacted]
Treatment group*Previous BCVA level				[redacted]				
Study visit number*Previous BCVA level				[redacted]	[redacted]		[redacted]	[redacted]

**The estimation algorithm did not converge for model 6 without interaction Treatment group*CPT*

Table 4. Estimates for selected model (model 1) and the best model in terms of AIC and BIC (model 9)

			Model 1 selected			**Model 9**		
Effect	level	Change from	Estimate	Standard Error	p	Estimate	Standard Error	p
Intercept								
Intercept								
Intercept								
Intercept								
AGE								
AGE								
AGE								
AGE								
Treatment Group								
Treatment Group								
Treatment Group								
Treatment Group								
Treatment Group								
Treatment Group								
Treatment Group								
Treatment Group								
Lens status								
Lens status								
Lens status								
Lens status								
Lens status								
Lens status								
Lens status								
Lens status								
Lens status								
Lens status								
Lens status								
Lens status								
Lens status								
Lens status								
Lens status								
Lens status								
Lens status								
Study visit 7								
Study visit 7								

			Model 1 selected			**Model 9**		
Effect	level	Change from	Estimate	Standard Error	p	Estimate	Standard Error	p
Study visit 7		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Study visit 7		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CPT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CPT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CPT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CPT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CPT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CPT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CPT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CPT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

			Model 1 selected			Model 9		
Effect	level	Change from	Estimate	Standard Error	p	Estimate	Standard Error	p
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment Group*Lens status	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

			Model 1 selected			Model 9		
Effect	level	Change from	Estimate	Standard Error	p	Estimate	Standard Error	p
Treatment Group*Lens status								
Treatment Group*Lens status								
Study visit 7 * Treatment group								
Study visit 7 * Treatment group								
Study visit 7 * Treatment group								
Study visit 7 * Treatment group								
Study visit 7 * Treatment group								
Study visit 7 * Treatment group								
Study visit 7 * Treatment group								
Study visit 7 * Treatment group								
Study visit 7* Previous BCVA level								
Study visit 7* Previous BCVA level								
Study visit 7* Previous BCVA level								
Study visit 7* Previous BCVA level								
Study visit 7* Previous BCVA level								
Study visit 7* Previous BCVA level								
Study visit 7* Previous BCVA level								
Study visit 7* Previous BCVA level								
Study visit 7* Previous BCVA level								
Study visit 7* Previous BCVA level								
Study visit 7* Previous BCVA level								
Study visit 7* Previous BCVA level								

			Model 1 selected			**Model 9**		
Effect	level	Change from	Estimate	Standard Error	p	Estimate	Standard Error	p
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			

			Model 1 selected			Model 9		
Effect	level	Change from	Estimate	Standard Error	p	Estimate	Standard Error	p
Study visit 7* Previous BCVA level	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]			

Table 5. Selected model parameters for selected and the best model assessing BCVA change (regression 3) without random effects and with different covariance structures

	Model 1					Model 9				
Final model	**without random effects**	**variance component (selected)**	**first-order autoregressive**	**unstructured**	**unstructured parameterized through Cholesky root**	**without random effects**	**variance component (selected)**	**first-order autoregressive**	**unstructured**	**unstructured parameterized through Cholesky root**
AIC (smaller is better)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BIC (smaller is better)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]