A Multicentre Phase III Double-masked Randomised Controlled Non-Inferiority Trial comparing the clinical and cost effictiveness of intravitreal therapy with ranibizumab (Lucentis) vs aflibercept (Eylea) vs bevacizumab (Avastin) for Macular Oedema due to Central Retinal Vein Occlusion (CRVO)

Statistical Analysis Plan

Version 5.4

EudraCT no: 2014-000272-26

ISCRN: 13623634

CONTENTS

ABBREVIATIONS:

BCVA	Best Corrected visual acuity
CFP	Colour Fundus Photography
eCRF	Electronic Case Report Form
CSRI	Client Service Receipt Inventory
CST	Central Subfield Thickness
CTU	Clinical Trials Unit
CVI	Certificate of Visual Impairment
DA	Disk areas
DMC	Data Monitoring Committee
DMO	Diabetic Macular Oedema
EMA	European Medicines Agency
ERM	Epiretinal Membranes
EQ-5D	Euro Quality of life questionnaire
ETDRS	Early Treatment Diabetic Retinopathy Study
FDA	Food and Drug Administration
FFA	Fundus Fluorescein Angiography
HEDMAP	Health Economic and Decision Modelling Analysis Plan
ICH	International Conference on Harmonization
IMP	Investigational Medicinal Product
ISRCTN	International Standard Randomised Clinical Trials Number
ITT	Intention to treat
LME	Linear mixed effects model
NPDR	Non Proliferative Diabetic Retinopathy
NVD	Neovascularisation Disc
NVE	Neovascularisation elsewhere
OCT	Optical Coherence Tomography
PDR	Proliferative Diabetic Retinopathy
PP	Per protocol
RCT	Randomised Control Trial
REC	Research Ethics Committee
SAR	Serious Adverse Reaction
SAP	Statistical Analysis Plan
SOP	Standard Operating Procedure
TSC	Trial Steering Committee
UK	United Kingdom
VA	Visual Acuity
VAS	Visual Analogue Scale
VEGF	Vascular Endothelial Growth Factor
VFQ-25	Visual Function Questionnaire
VMT	Vitreomacular Traction

INVESTIGATORS:

Trial Statistician:

Name: Joana Vasconcelos
Address: Imperial Clinical Trials Unit, Stadium House, 68 Wood Lane, Imperial College London, London W12 7RH
Email: joana.vasconcelos@imperial.ac.uk

Lead Statistician:

Name: Professor Toby Prevost
Address: Imperial Clinical Trials Unit, Stadium House, 68 Wood Lane, Imperial College London, London W12 7RH
Email: a.prevost@imperial.ac.uk

Chief Investigator:

Name: Mr. Philip Hykin
Address: Moorfields Eye Hospital. 162, City Road, London EC1V 2PD.
Email p.hykin@nhs.net

Co-lead:

Name: Professor Sobha Sivaprasad
Address: Moorfields Eye Hospital & UCL Institute of Ophtalmology. 162, City Road, London EC1V 2PD
Email: sobha.sivaprasad@nhs.net

Chair of Trial Steering Committee:

Name: Miss Susan Downes
Address: Oxford Eye Hospital, Oxford University Hospitals, West Wing, John Radcliffe Hospital, Headley Way, Oxford OX3 9DU
Email: susan.downes@ouh.nhs.uk

Chair of Data Monitoring Committee:

Name: Prof Sarah Walker
Address: Professor of Statistics,
Medical Research Council Clinical Trials Unit (MRC CTU) Aviation House, 125 Kingsway, London, WC2B 6NH
Email: asw@ctu.mrc.ac.uk

Sponsor:

Name: Ms Maria Hassard
Address: NIHR Moorfields Biomedical Research Centre, Moorfields Eye Hospital & UCL Institute of Ophthalmology 162 City Road, London, EC1V 2PD
Email: maria.hassard@moorfields.nhs.uk

Funder:

Name: NIHR HTA CET – National Institute for Health Research, Health Technology Assessment Programme, Clinical Trials and Evaluation Stream
Contact name: Simon Bevan
Email: s.bevan@soton.ac.uk

1.0 Introduction

1.1 Derivation of the statistical analysis plan

The present statistical analysis plan was derived from the trial protocol, by the trial statistician, Joana Vasconcelos, with the supervision of Professor Toby Prevost. The trial statistician is responsible for developing the SAP as well as for carrying out the statistical analysis for interim and final statistical reporting of the trial. The supervisor will revise the SAP and give an overall verification of the analysis throughout the study, in keeping with the Standardised Operating Procedures (SOPs) of the King's Clinical Trials Unit, including the SOP for developing the Statistical Analysis Plan.

The formation of this Plan has drawn on statistical guidance from: the ICH Harmonised Tripartite Guideline: Statistical Principles for Clinical Trial E9 and E3(1), the CONSORT statement for reporting trials(2), the Committee for Medicinal Products for Human Use (formerly known as the Committee for Proprietary Medicinal Products) report(3), from general issues in non-inferiority designs(4, 5) and from specific trial issues (6, 7).

The trial statistician will write the first version of the plan. After revision by the supervisor the plan will be filed as version number 2. The plan will then be discussed with the Principal Investigator for further input and filed as version number 3. The plan will then be sent to the DMC and TSC for final approvals and saved as version 4 and 5 of the plan, respectively.

1.2 Purpose and scope of the statistical analysis plan

The purpose of this Statistical Analysis Plan is to set out the study objectives and hypotheses, and the analytical approaches and procedures necessary to address these for the main trial paper and to provide guidance for further research reported in other papers, promoting consistent approaches and methods.

As there can typically be more than one analytical approach to address a hypothesis, there is the potential for different results to be produced from using alternative approaches, alternative methods, alternative outcome definitions and the alternative data that may be involved. These differences can be influential, for example, when results are of borderline statistical significance.

Therefore, this Plan records those decisions that can be made about study hypotheses, outcome definitions and statistical procedures, along with their basis and the appropriateness of the assumptions required for their use, in advance of the main trial analysis, while any access to unmasked follow-up data and to trial arm is prevented.
Changes within subsequent versions of the Plan prior to analysis will be dated, with the basis for the changes reasoned, and recorded within the plan.

Other analysis decisions may need to be made later, based on viewing the observed distribution of the data. Where possible these decisions will be made prior to access to trial arm, or where necessary from control arm data alone. The main place for this will be the "pre-analysis review" phase, blinded to treatment arm (ICH E9 (1)), taking place after final participant follow-up and prior to the study arm being made available. Prior to the first DMC

meeting with follow-up data, indications, such as on the need for data transformation, will be made at a point when baseline data only is available without access to arm.

Decisions will be supported by reasoning and justification, and these will be appended to the Statistical Analysis Plan, and dated, to provide a record of any post-analysis decisions and their basis.

It is not intended that the strategy set out in the plan should prohibit sensible practices. However, the principles established in the plan will be followed as closely as possible when analysing and reporting the trial.

2.0 Overview of the condition and treatment

2.1 Description of the condition and its importance/scale

Retinal Vein Occlusion (RVO) is a blockage of the small retinal veins by a blood clot. When the blood cannot drain away from the retina, there is an accumulation of pressure in the blood vessels resulting in leakage of fluid and blood causing macular oedema and ischemia. This condition can affect the central retinal vein (CRVO) (formed by the union of the four retinal veins (one retinal vein drains each quarter of the eye)) or a major branch retinal vein (BRVO), where blockage occurs somewhere along the course of one of the four retinal veins.CRVO is characterised by retinal haemorrhages, venous dilatation and tortuosity in all four quadrants of the retina and is typically more severe than BRVO(8, 9).

Approximately 6,860 people develop CRVO every year in England and Wales of whom 5,150 are potentially eligible for treatment. Once established, the visual impairment due to CRVO is typically profound with no tendency to improve spontaneously. Without intervention permanently impaired visual loss is likely to occur. In this study the focus will be on CRVO.

2.2 Description of the standard treatment (or placebo or current care)

Until 2011 no treatment was available to improve vision in people with CRVO. In 2011, NICE recommended the NHS an implant called “Ozurdex”, which although improving vision, could cause cataracts and glaucoma with repeated use, was difficult to administer and had only moderate uptake in the UK. In the meantime another treatment, more effective with fewer side effects, Ranibizumab (Lucentis, Novartis, & Genentech), was approved by the FDA and EMA for macular oedema due to CRVO. Ranibizumab is a humanized, affinity-matured VEGF antibody fragment that binds to and neutralizes all isoforms of VEGF-A and their biologically active degradation products and it was the first anti-VEGF therapy to demonstrate improved visual outcomes in patients with neovascular age related vascular degeneration. It is EMA licensed for use in wet age related macular degeneration, diabetic macular oedema and retinal vein occlusion and NICE recommended for all three.

2.3 Description of the investigational treatments

This study aims to determine if the two anti-VEFG agents Bevacizumab or Afilbercept are as effective as Ranibizumab in reducing visual loss from MO due to CRVO, whether they have

an equivalent side effect profile and whether either could be considered as a recommended NHS treatment based on non-inferior clinical effectiveness and superior cost-effectiveness.

Aflibercept or VEGF Trap-Eye (Eylea, Bayer/Regeneron), is a fusion protein of the key domains of VEGF receptors 1 and 2 and human IgGFc that blocks all VEGF-A isoforms and placental growth factor. Like ranibizumab, it is EMA licensed for nvAMD, DMO and RVO. It is FDA approved for CRVO and NICE has recommended this drug for MO due to CRVO (TA 305).

Bevacizumab (Avastin, Genetech/Roche), is a monoclonal antibody that inhibits vascular endothelial growth factor (VEGF), a mediator in the pathogenesis of common and disabling eye disorders including neovascular age related macular degeneration (nvAMD), diabetic macular oedema (DMO) and retinal vein occlusion (RVO). Bevacizumab is EMA licensed for the treatment of cancer but not for use in the eye. There is limited evidence regarding its use in central retinal vein occlusion (CRVO) with macular oedema (MO). To date, bevacizumab has been found to be non-inferior to ranibizumab for all visual acuity primary and secondary endpoints in nvAMD in the IVAN and CATT studies (10, 11).

2.4 Description of the motivation for the study / need to investigate the new treatment

This project will compare the relative clinical and cost effectiveness of the anti-VEGF agents bevacizumab, aflibercept, and ranibizumab in MO due to CRVO over 100 weeks which is of critical importance to the NHS in the next 10 years. The NICE Final Appraisal Document for ranibizumab in RVO has recommended that further research is required comparing ranibizumab and bevacizumab (12) and there are no comparisons to date of bevacizumab and ranibizumab with aflibercept. This trial will be the first well-powered Phase III trial exploring the relative effectiveness of these drugs and aflibercept in the management of MO in CRVO. This study will inform us of the potential use of the most clinically effective and cost-effective drug for this condition in the NHS by providing a better understanding of the economic and societal impact of RVO in the long-term and help decision makers evaluate and compare these medical interventions over the duration of the natural history of the condition.

3.0 Populations and Study Sample

3.1 Target Population

The *target population*, to which inferences from the end of this trial are intended to generalise, is the population of adult patients with MO due to CRVO.

3.2 Trial Population

The trial population, from which the study sample is drawn, is further defined to be adults aged 18 year or over, of less than 12 months duration who attend the 40 ophthalmology centres in the UK with expertise in retinal disorders and a proven track record in effective research.

Only one eye per patient will be included in the trial. In subjects with both eyes meeting the eligibility criteria, then the ‘worst seeing eye’ will be enrolled unless the patients preference is for the best seeing eye (see section 7.3).

3.3 Trial Samples

3.3.1 Intention To Treat (ITT)

The achieved trial sample comprises those patients who consent to participate and are actually randomised into this trial. These patients are the study subjects.

This randomised trial sample is also the trial Intention To Treat (ITT) population. The intention-to-treat principle states that every subject will be analysed according to the treatment group to which they were randomised. In this trial, subjects' data will be analysed according to the *Intention to Treat Strategy* (13), under which at least one analysis is recommended to be based on the ITT population.

The trial ITT population comprises all randomised participants, regardless of eligibility (inclusion/exclusion) error, post-randomisation withdrawal, and whether the correct study treatments were received, or other interventions received.

3.3.2 Per Protocol (PP)

Definition

A per protocol set of subjects will also be included. These will be defined as the subset of the found to be eligible at entry and who had minimal sufficient exposure to the treatment regimen, defined as 4 treatments correctly assessed and received during the first 6 visits up to week 20. For each of the first four visits, a correct treatment is defined as receiving the injection. For the 5^{th} and 6^{th} visits, a correctly assessed and received treatment is defined to be the receipt of an injection where this is indicated to be required by the retreatment criteria or the non-receipt of an injection where this is indicated by the retreatment criteria..

Rationale

The main reason for having a per protocol set comes from the fact that this is a non-inferiority trial and so the use of the full analysis set is generally not conservative (ICH E9 section 5.2.3 (1)). As Lesaffre 2008 (7) states, "*dropouts and a poor conduct of the study might direct the results of the two arms towards each other*". Although this can be interpreted as an indication that the per protocol analysis is the conservative choice for non-inferiority studies Garrett AD 2003 (6) state that "*The perceived conservative nature of the PP population appears to be much more a reflection of reduced patient numbers than the presence of bias, while bias can be in either direction depending on the pattern of violations*". Moreover, with two active treatments it may be more likely that any bias affecting both treatments would be reduced in comparison to a placebo-controlled trial.

Prominence

Non-inferiority will only be declared if both ITT and the PP analysis are supportive of a non-inferiority conclusion. This is supported by the Committee on Proprietary Medical Products Points-to-Consider (5) and several other papers (7, 14).

The requirement to declare noninferiority in both the ITT and the PP analyses promotes the adherence to treatment protocol and the minimisation of exclusions, maintaining power.

3.4 Inclusion and Exclusion criteria

The inclusion and exclusion criteria are the following:

Inclusion criteria:
1. Subjects of either sex aged ≥ 18 years.
2. Clinical diagnosis of centre-involving macular oedema (MO) due to CRVO
3. CRVO of ≤ 12 months duration.
4. Best corrected visual acuity in the study eye ≥ 19 and ≤ 78 ETDRS letters (approximate Snellen VA 3/60 to VA 6/9).
5. Best corrected visual acuity in the non-study eye ≥ 14 ETDRS letters (approximate Snellen VA ≥ 2/60.
6. SD-OCT central subfield retinal thickness (CST) > 320µm (Spectralis) predominantly due to MO secondary to CRVO in the study eye.
7. Media clarity, pupillary dilatation and subject cooperation sufficient for adequate fundus imaging of the study eye.
8. In cases of bilateral CRVO, if both eyes are potentially eligible, unless the patient prefers otherwise, the worst seeing eye will be recruited.

Exclusion criteria (applied to the study eye only and to the non study eye only where specifically stated):
1. Macular oedema considered to be due to a cause other than CRVO (e.g. diabetic macular oedema, Irvine-Gass syndrome).
2. An ocular condition is present that, in the opinion of the investigator, might affect macular oedema or alter visual acuity during the course of the study (e.g. vitreomacular traction).
3. Any diabetic retinopathy or diabetic macular oedema at baseline clinical examination of the study eye.
4. Moderate or severe non proliferative diabetic retinopathy (NPDR) or quiescent, treated or active proliferative diabetic retinopathy (PDR) or macular oedema in the non-study eye. Note: Mild NPDR only is permissible in the non-study eye.
5. History of treatment for MO due to CRVO in the past 90 days with intravitreal or peribulbar corticosteroids or in the last 60 days with anti-VEGF drugs or > 6 prior anti-VEGF treatments in the previous 12 months.
6. Active iris or angle neovascularisation, neovascular glaucoma, untreated NVD, NVE and vitreous haemorrhage or treatment for these conditions in the last 1 month.
7. Uncontrolled glaucoma [>30mmHg], either untreated or on anti-glaucoma medication at screening.
8. Any active periocular or intraocular infection or inflammation (e.g. conjunctivitis, keratitis, scleritis, uveitis, endophthalmitis).

Systemic exclusion criteria:
9. Uncontrolled blood pressure defined as a systolic value > 170mmHg and diastolic value > 110mmHg.
10. Myocardial infarction, stroke, transient ischaemic attack, acute congestive cardiac failure or any acute coronary event < 3 months before randomisation.
11. Women of child bearing potential unless using effective methods of contraception throughout the study and for 6 months after their last injection for the trial. Effective contraception is defined as one of the following:
    a. Barrier method: condoms or occlusive cap with spermicides.

b. True abstinence: When it is in line with the preferred and usual lifestyle of the subject. Periodic abstinence (e.g. calendar, ovulation, symptothermal, post-ovulation methods) and withdrawal are not acceptable methods of contraception.
c. Have had tubal ligation or bilateral oophorectomy (with or without hysterectomy).
d. Male partner sterilisation. The vasectomised male partner should be the only partner for the female participant.
e. Use of established oral, injected or implanted hormonal methods of contraception and intrauterine device.

12. Pregnant or lactating women.
13. Males who do not agree to an effective form of contraception for the duration of the study and for 6 months after their last injection for the trial.
14. Hypersensitivity to the active ingredients aflibercept, bevacizumab or ranibizumab or any of the excipients of these drugs.
15. Hypersensitivity to Chinese Hamster Ovary (CHO) cell products or other recombinant human or humanised antibodies.
16. A condition that, in the opinion of the investigator, would preclude participation in the study.
17. Participation in an investigational trial involving an investigational medicinal product within 90 days of randomisation.

4.0 Objectives, principal research question and associated hypotheses

4.1 Principal Trial objective

The objective of the LEAVO trial is to determine if Bevacizumab or Afilbercept are as effective as Ranibizumab in reducing visual loss from MO due to CRVO, whether they have an equivalent side effect profile and whether either could be considered as a recommended NHS treatment based on non-inferior clinical effectiveness and superior cost-effectiveness.

4.2 Principal Research Question

The principle research question is as follows: Is visual acuity following Aflibercept or Bevacizumab non-inferior to Ranibizumab in eyes with MO due to CRVO at 100 weeks?

4.3 Hypotheses

The hypotheses refer to the populations of relevant patients rather than study subjects.

The *Working hypothesis*: The so-called "working hypothesis" is the hypothesis which motivates the trial, which the trial results may or may not support. It is that the change in best corrected visual acuity is non-inferior in patients treated with either Aflibercept or Bevacizumab compared to patients treated with Ranibizumab.

The Statistical *Null Hypothesis* 1: Bevacizumab is inferior to Ranibizumab in eyes with MO due to CRVO at 100 weeks.
The Statistical *Null Hypothesis* 2: Aflibercept is inferior to Ranibizumab in eyes with MO due to CRVO at 100 weeks.

Statistical *Alternative hypothesis* 1: Bevacizumab is noninferior to Ranibizumab in eyes with MO due to CRVO at 100 weeks.
Statistical *Alternative hypothesis* 2: Aflibercept is noninferior to Ranibizumab in eyes with MO due to CRVO at 100 weeks.

4.4 Study objectives

4.4.1 Primary objective:

1. To determine whether bevacizumab is non-inferior to ranibizumab in treating visual loss due to MO secondary to central retinal vein occlusion at 100 weeks
2. To determine whether aflibercept is non-inferior to ranibizumab in treating visual loss due to MO secondary to central retinal vein occlusion at 100 weeks.

4.4.2 Secondary Objectives

1. To determine the difference between arms in mean change in best corrected visual acuity at 52 weeks.
2. To determine the difference between arms in the proportion of participants with ≥ 15 ETDRS letter improvement (appreciable visual gain), ≥ 10 letter improvement, <15 letter loss and ≥ 30 ETDRS letter loss (severe visual loss) at 52 and 100 weeks.
3. To determine the difference between arms in the proportion of participants with ≥73 ETDRS letters or better than 6/12 Snellen equivalent (ie approximate driving visual acuity), ≤58 ETDRS letters (≤6/24) and ≤ 19 letters (≤3/60)(CVI partial and severe visual impairment) at 52 and 100 weeks.
4. To determine the difference between arms in the mean change in OCT CST and macular volume at 52 and 100 weeks.
5. To determine the difference between arms in the proportion of participants with OCT CST < 320µm (Spectralis or refer to protocol appendix 1) at 52 and 100 weeks (key guide to subsequent NHS clinical practice).
6. To determine the differences between arms in the mean number of injections performed per study eye per participant at 100 weeks.
7. To determine any differences in the relative effectiveness of the investigational treatments and comparator on quality of life and resource utilization, reported as Incremental Cost Effectiveness Ratios (ICERs) at 52 weeks.
8. To determine any differences in the relative effectiveness of the investigational treatments and comparator on quality of life and resource utilisation, reported as Incremental Cost Effectiveness Ratios (ICERs) at 100 weeks.
9. To detect any differences in the prevalence of local and systemic side effects at 100 weeks.
10. To determine differences between arms at 100 weeks in the proportion i. of persistent non-responders (see protocol Section 8.14.7), ii. of participants that develop a change in retinal non-perfusion compared to screening iii. of participants that develop anterior and posterior segment neovascularisation.
11. To determine differences between arms in mean change in best corrected visual acuity at 100 weeks due to i) baseline visual acuity stratified as ≤38 letters, 39-58 letters, 59-78 letters, ii) duration of disease stratified as: <3 months, 3-6 months and > 6 months, iii) treatment stratified as naïve vs previous treatment iv) quantity of retinal ischaemia (<10, ≥10 and < 30, and ≥ 30 DA of non-perfusion).
12. To determine differences between arms in changes in area of non-perfusion at 100 weeks and OCT anatomical features over time and at 100 weeks.

5.0 Trial design

This is a two year multicentre (approximately 40 centres), double-blind pragmatic individually randomised controlled trial that will test the non-inferiority visual acuity from treatment with Bevacizumab and Aflibercept to Ranibizumab at 100 weeks in 459 adult participants with MO due to CRVO of less than 12 months duration.

5.1 Treatment arms

The trial is randomised with three arms and with equal allocation of participants in a 1:1:1 ratio to the three arms.

Arm A: Treatment: An intravitreal injection of Aflibercept (Eylea, Bayer) (2.0mg/5µl) will be administered at baseline, 4, 8 and 12 weeks. After this the retreatment criteria is ascertained (see Protocol section 8.14).

Arm B: Treatment: An intravitreal injection of Bevacizumab (Avastin, Roche) (1.25mg in 50ul) will be administered at baseline, 4, 8 and 12 weeks. After this the retreatment criteria is ascertained (see Protocol section 8.14).

Arm C: Control: An intravitreal injection of Ranibizumab (Lucentis, Novartis) [0.5mg/50ul] will be administered at baseline, 4, 8 and 12 weeks. After this the retreatment criteria is ascertained (see Protocol section 8.14).

Different labels can be used for the control arm, such as placebo, comparator, standard care, or control.

5.2 Type of RCT

This is a phase III, parallel groups' trial.

5.3 Frequency and duration of follow-up

Participants in all 3 study arms will be seen at weeks 0, 4, 8, 12, 16, 20 and 24. After this participants will potentially be seen every 4 weeks until week 96 if retreatment criteria are met. If retreatment criteria are not met at three successive visits from week 24 onwards the visit interval is increased to 8 weekly until week 96. They will also be finally seen at 100-weeks.

6.0 Trial measures

6.1 Primary outcome

The primary outcome is Best Corrected Visual Acuity (BCVA) in the study eye measured in ETDRS letter score at 4 metres at 100 weeks. Measurements of BCVA at milestone visits are included in the analysis of the primary outcome. Any BCVA measurement will be excluded from the analysis if it is both more than 3 standard deviations below the mean at that timepoint (including all measurements) AND taken within 3 months of occurrence of a vitreous haemorrhage or another cause unrelated to maculopathy secondary to CRVO (such as neovascular glaucoma).

6.2 Secondary outcomes

The secondary efficacy outcome measures are listed as follows according to the type of variable they will be formally analysed at 52 weeks and 100 weeks, but also measured at other time points.

6.2.1 Continuous outcome variables:

i. Visual Acuity and Clinical Outcomes

1. Change from baseline in ETDRS letter score measured at 4 metres at 52 weeks.
2. Change from baseline in mean OCT central subfield thickness (CST) at 52 and 100 weeks.
3. Change from baseline in macular volume at 52 and 100 weeks.
4. Number of injections performed in the study eye at 100 weeks
5. Change in retinal non-perfusion as assessed by mean disc area of non-perfusion at 100 weeks and by the ischaemic index at 100 weeks.

ii. Patient reported outcomes

1. National Eye Institute visual function questionnaire (VFQ25) composite score, distance, and near subscales at 52 and 100 weeks.
2. Quality of life (EQ-5D with and without vision bolt-on) at 52 and 100 weeks.

iii. Economic reported outcomes (this is detailed in the health economics analysis plan)

1. Quality of life scales (VFQ25 composite score, distance and near subscales, and EQ5D with and without vision bolt-on) at 0, 12, 24, 52, 76 and 100 weeks.
2. Resource utilization (Client Service Receipt Inventories) at 0, 12, 24, 52, 76 and 100 weeks.

6.2.2 Categorical outcome variables:

i. Visual Acuity and Clinical Outcomes

1. Participants with ≥15 ETDRS letter improvement (appreciable visual gain), ≥ 10 letter improvement, <15 letter loss and ≥ 30 ETDRS letter loss (severe visual loss) at 52 and 100 weeks.
2. Participants with ≥73 ETDRS letters or better than 6/12 Snellen equivalent (i.e. approximate driving visual acuity), ≤58 ETDRS letter (≤6/24) and ≤19 letters (≤3/60) (CVI partial and severe visual impairment) at 52 and 100 weeks.
3. Participants with OCT CST <320µm (Spectralis or refer to protocol appendix 1) at 52 and 100 weeks (key guide to subsequent NHS clinical practice).
4. Participants with the anatomical OCT features: diffuse intraretinal oedema, intraretinal cystic change, subretinal fluid, vitreomacular interface abnormaility (either VMT or ERM) over time and at 100 weeks.

ii. Safety and tolerability

1. Prevalence of local and systemic side effects at 100 weeks.
2. Participants that are persistent non-responders (section 8.14.7 of the protocol) and that develop anterior and posterior segment neovascularisation at 100 weeks.

6.3 Timing of measures

A full schedule on the timing of measures is provided in below:

^Mandatory Visits: Loading (wk 4 & 8) & Milestones (baseline, wks 12, 24, 52, 76, 100)	Screening	Baseline	Week 4	Week 8	Week 12	Week 16	Week 20	Week 24		Week 52		Week 76		**Week 100	Unsch. Visit.	**Withdrawal Visit
Variable treatment visits									4-8 weekly		4-8 weekly		4-8 weekly			
Weeks		0	4	8	12	16	20	24	28-48	52	56-72	76	80-96	100	1-99	13-97
Visit window (days)	-10 to 0	0	0 to +14	0 to +14	0 to +14	-14 to +14	-14 to +14	-14 to +14	-14 to +14	-14 to +14	-14 to +14	-14 to +14	-14 to +14	-14 to +14		
Informed Consent	X															
Inclusion/Exclusion Criteria review	X	X[3]														
Randomisation[1]		X														
Urine Pregnancy test in women of child bearing age.	X															
Patient demographics, medical and ophthalmic history	X															
Adverse events	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Concomitant medication	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Blood Pressure	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Best corrected ETDRS visual acuity in both eyes (refraction visit =X1)	X1	X	X	X	X1	X	X	X1	X	X1	X	X1	X	X1	X / X1[5]	X1
Standard Ophthalmic Examination	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
SD-OCT in both eyes	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X
7-field or wide-angle CFP[2]	X									X				X	+/-[5]	X
7-field or wide angle FFA[2]	X													x	+/-[5]	X
VFQ-25 and EQ-5D with and without vision 'bolt-on'		X			X			X		X		X		X	+/-[5]	X
CSRI		X			X			X		X		X		X	+/-[5]	X
Treatment Allocation Guess Form[4]														X		X
Administer IMP*		X	X	X	X	X2	X2	X2	X2	X2	X2	X2	X2		X2	X

X1 – Same day refracted best corrected visual acuity
X2 - PRN treatment.
Study Treatment Visit: non shaded square.
Study Milestone Visit: shaded square
^Milestone visits and mandated loading visit dates should be agreed with participant prior to performing randomisation
*Intravitreal injections including immediate post injection checks are performed as per each trial sites local policy and may include a check of ON perfusion or VA or IOP or a combination of these.
** Participants should be reminded to use an effective form of contraception for 6 months after their last trial injection. Females of child bearing potential should be reminded to notify the local study team if they fall pregnant during this time.
[1]Randomisation should only occur once all other assessments at baseline (week 0) have occurred
[2]Further colour fundus photographs and fluorescein angiography may be performed as per investigator discretion. Colour fundus photographs should be done if a patient converts from non-ischaemic to ischaemic CRVO.
[3]To include review of screening assessment test results and confirmation of eligibility.
[4]To be completed by participant, masked investigator, site optometrists.
[5]To be performed (as required) if unscheduled visit is a milestone visit.

6.4 Participant duration in the study

Each study subject will participate in the trial from the day that the they give informed consent to their last final visit at 100 weeks.

6.5 Final assessment

The final study assessment is when the last study subject achieves their 100 weeks assessment.

7.0 Sample Size

7.1 Determination of the primary outcome effect size

Bevacizumab and aflibercept are defined to be substantially inferior to ranibizumab, if in each case, the mean of the primary outcome (change in best corrected ETDRS visual acuity letter score) is worse by a margin of five letters.

The two null hypotheses, that bevacizumab is substantially inferior to ranibizumab, and that aflibercept is substantially inferior to ranibizumab, will each be rejected if the estimated 95% confidence interval for the difference in treatment means lies wholly above the five letter margin in each case.

The choice of a five-letter margin is 32% higher than the available estimated 12-month placebo-controlled effect of 6.6 letters (15) for ranibizumab, the standard (comparator) treatment for LEAVO. This margin choice is therefore consistent with maintaining assay sensitivity sufficiently to be able to declare non-inferiority.

7.2 Determination of the primary outcome variability

For a similar trial on CRVO (15) the standard deviation reported in Ranibizumab arms at 12 months was 14.3. In the absence of 24-month data we have assumed a comparable SD of 14.3 at 24 months.

7.3 Clustering of outcomes from eyes within subjects effects

Only one eye per subject can be selected for the study. In 95% of cases, one eye will be affected by CRVO. As explained in the protocol, bilateral RVO is rare, but if it happens and both eyes are eligible, the eye included is the 'worst seeing eye'. However, participants will be given the choice if both eyes are found to be eligible. All observations are in this way able to be assumed to be independent in the sample size calculation and statistical analysis.

7.4 Power to detect effects

There is 80% power to detect non-inferiority using a two-sided 95% confidence interval from an analysis of covariance test with adjustment for baseline visual acuity and randomisation stratifiers.

7.5 Determination of the sample size based on the primary outcome

The sample size was set to be 459 participants, 153 per arm (1 eye per subject). The target of 390 subjects followed up with primary outcome required in the analysis involves a 15% allowance for dropout and protocol exclusions.

Sample size calculations were performed using nQuery Advisor 4.0 software.

7.6 Detectable effects sizes expressed in general standardised form

For a continuous secondary outcome, with 153 subjects per arm followed up we can detect effects of size 0.45 Standard Deviation's difference between means with 80% power using a two-sided t-test at the 5% significance level. For binary outcomes, we have at least 90% power to detect a difference in proportions of 0.2 using a chi-squared test at the 5% significance level.

8.0 Randomisation and Subgroups

8.1 Arms

Each participant will be equally randomised to one of three arms: bevacizumab, aflibercept or ranibizumab.

8.2 Method of allocation

The method of minimisation incorporating a random element will be used. There will be three stratifying factors: visual acuity (stratified by baseline BCVA letter score (≤38 [approximate Snellen equivalent <6/60], 39–58 [approximate Snellen equivalent between 6/48 and 6/24 exclusive], ≥59 [approximate Snellen equivalent ≥ 6/18]) and onset of symptoms to presentation at hospital and commencement of therapy (< 3 months, 3-6 months and > 6 months) and prior treatment or not.

8.3 Relative timing of randomisation

Randomisation will be via a bespoke web based randomisation system hosted at the King's CTU on a secure server. Once a participant enters the study and their data is entered into the

eCRF, they will be allocated a unique study PIN. This, along with their date of birth and initials will be used to identify the participant and their data throughout the study.

8.4 Subgroup variables

Three subgroup variables will be considered: i) baseline visual acuity (low, moderate, high: ≤38 letters, 39-58 letters, 59-78 letters), ii) disease duration (<3 months, ≥3 months) and iii) quantity of retinal ischaemia (non-ischaemic vs ischaemic vs very ischaemic CRVO (<10, ≥10 and ≥ 30 DA of non-perfusion) .

These are based on the fact that visual gain in the low vision group may be higher than that achieved by the high vision group and this effect may be different between arms. Patients with ischaemic CRVO may not have similar visual acuity gains to those with no ischaemia and this effect may be different between treatment arms. The shorter the duration of disease, the better the visual acuity outcomes but this may vary between treatment arms.

9.0 Blinding

The trial will be double masked. Study participants, clinicians and members of the research team who will undertake key measurements (visual acuity, morphology) will be masked to group allocation. The clinician administering the drug injected into the vitreous will not be masked. This will ensure that the study has a high level of both treatment integrity (delivery of the treatment as intended) and treatment differentiation (treatment conditions differed from one another in the intended manner).

The trial statistician will have access to the accumulating outcome data that is required for reporting to the DMC.Both the trial statisticians will attend both the open and closed DMC meetings.

10.0 Data and Distributions

10.1 Data decisions made

The data manager will make limited decisions about data variables and values so that issues such as missing data can be comprehensively handled by the trial statistician. Decisions which impact on the analysis will be recorded in an appendix of this statistical analysis plan.

10.2 Outcomes requiring derivation

List of outcomes with source of derivation code:

1) **VFQ-25** (16): a validated tool for vision related quality of life. It consists of a base set of 25 vision targeted questions representing 11 vision-related sub-scales, plus an additional single-item general health rating question. The overall composite score is computed as the simple average of the vision-targeted sub-scale scores, excluding the general health rating question. The overall score can range from 0 (worst possible score) to 100 (best).

   EQ-5D (17-19) with and without vision bolt-on: The EQ-5D is a generic instrument for describing and valuing health. It is based on a descriptive system that defines health in terms of 5 dimensions (Mobility, Self-care, Usual activities,

Pain/Discomfort, Anxiety/Depression. Each dimension has 5 response categories (EQ-5D-5L) corresponding to eg. “no problems",”slight problems”, “moderate problems”, “severe problems”, and “unable to/extreme problems”. A preference-based score ranges from states worse than dead (<0) to 1 (full health), anchoring dead at 0. In addition, the EQ-5D includes a visual analogue scale (EQ-VAS), which records the respondent’s self-rated health on a vertical scale where the endpoints are labelled ‘Best imaginable health state’ (marked as 100) and ‘Worst imaginable health state’ (marked as 0).

The EQ-5D with bolt-on is similar to the EQ-5D-5L but another dimension was added (vision) in order to overcome perceived inadequacies in a particular population. The corresponding scoring system for the EQ-5D vision 'bolt-on' on has not been finalised yet. Further details are developed in the HEDMAP.

10.2.1 Procedure for deriving variables

If there is existing syntax code to derive a variable within the King's Clinical Trials Unit then this will be used. Otherwise new code will be developed by the trial statistician and verified by the senior statistician.

10.2.2 Missing items in scale and subscales

The number (%) of patients with complete data for each scale will be reported. If scales provide missing value guidance then this will be used.

10.3 Use of data transformation

It is not anticipated that any continuous outcomes will need to be considered for transformation, because the sample size is reasonably large for group comparisons in the main trial analyses. Assumptions of normality and constant variance required by the models will be examined using residual and other diagnostic plots. If it is relevant, and necessary, where sample size is reduced, a log transformation will be considered, because this retains a sensible interpretation for inferences; in relative terms between arms. If an absolute interpretation is needed, then data transformation may not be undertaken, but a nonparametric Bootstrap method for obtaining confidence intervals may be considered (20).

10.4 Defining Outliers

Outliers are observations that have extreme values relative to other observations observed under the same conditions. An outlier will be defined here as a data-point being at least four standard deviations from the mean of its distribution of values observed across other patients. This definition will apply to the transformed scale for those outcomes that have been log transformed.

A “bivariate outlier“ for checking will be defined here as a pair of successive serial data-points of the same measure for a participant whose difference is at least four standard deviations from the mean of all patients’ such differences. Simple plots of successive pairs of serial measures will be used through the 24-month period to assist in identifying outliers for data checking.

10.5 Handling outliers

Outliers will be identified for further investigation by looking at the distributions of the data through histograms, scatter plots or box-plots. Univariate tests for the compatibility of the distribution with a normal distribution will not be undertaken since they can be too sensitive to departures that are often not relevant for the comparison of means (Central Limit Theorem).

Once an outlier is found, a blinded member of the team with sufficient clinical experience will be involved in the decisions as to whether a data value is impossible versus implausible versus plausible. If the outlier is impossible, then it will be set to missing, and a list of these occurrences will be appended to this SAP. If an outlier is clinically plausible, the outlier will remain. If an outlier is clinically implausible (but possible), it will not be ignored or deleted but will be retained for ITT analysis.

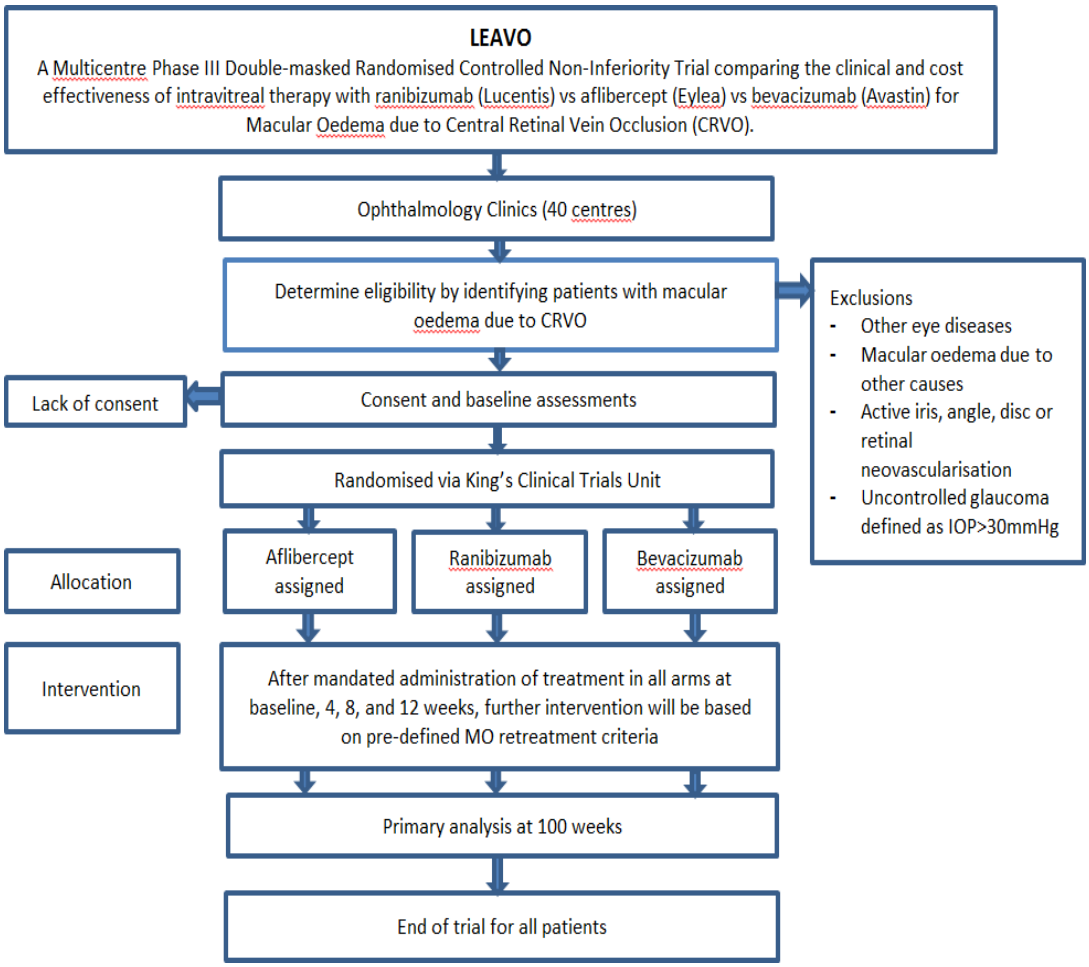
If outliers remain in the distribution of a variable, then data transformations or nonparametric methods of analysis may be considered.

Sensitivity analysis will be undertaken to check whether the outlier is influential by obtaining results with and then without inclusion of the outlier. If the conclusions are changed, then this will be noted.

11.0 Descriptive analysis

11.1 Flow diagram

The flow diagram of the study is the one below. This will include the number randomised, who comprise the intention to treat and per protocol population, and the numbers followed-up to be in the analyses of the primary outcome as well as the main reasons for missing data by stages of the trial.



11.2 Baseline comparability of randomised groups

Baseline descriptions of participants by treatment and overall will be summarised (into Table 1 of the report). No significance testing will be carried out as any differences found may be chance-generated and not for hypothesised reasons.

Continuous variables such as OCT central subfield thickness and VFQ-25 will be summarised using means and standard deviations (SD) and/or medians and interquartile range (IQR) for variables presenting a skewed distribution. Categorical variables such as proportion of patients gaining ≥15 BCVA or participants with OCT CST < 320µm will be described using numbers and percentages.

11.3 Comparison of rates of adherence and follow-up

High compliance and low attrition rates are anticipated for this study according to previous clinical trial experience (91.6% of subjects completed the active treatment arms in the 12 month CRUISE (CRVO) study and withdrawals were mainly due to physician and patients decisions (see Protocol). A cumulative drop-out of approximately 15% by year 2 was predicted and reflected in the sample size calculations. Nevertheless, compliance rates and attrition rates will be compared and reported by arm using Fisher's exact test.

12.0 Analysis covariates

12.1 Stratifiers

It is important to consider which, if any, covariates are to be adjusted for in the analyses. The ICH E9 guideline (1) recommends that consideration be given to accounting for randomisation stratifiers by adjusting for them as covariates in linear model. This tends to improve the precision of estimated treatment effects. Therefore, for continuous outcomes, the analysis will include adjustment for the randomisation stratifiers of screening BCVA letter score (3 levels) and disease duration (2 levels).

12.2 Baseline

The corresponding baseline measure for a continuous outcome is also often predictive of the outcome at follow-up. Therefore "baseline", if collected, will be an additional covariate when modelling continuous outcomes. This will be the case for visual acuity and macular volume. The continuous baseline will have precedence for inclusion in the model over the corresponding categorical randomisation stratifier, where applicable.

13.0 Primary outcome analysis

13.1 Statistical Model

The following description of the statistical analysis applies to each of the two investigational treatments, bevacizumab and afilbercept and the standard treatment, ranibizumab.
The primary efficacy measure is the change from baseline in refracted best corrected visual acuity (BCVA) in the study eye, using the ETDRS letter score at 100 weeks. As the analysis approach for continuous outcomes below makes advantage of covariate-adjustment for the baseline of the outcome, the primary endpoint can equivalently be regarded to be each participant's 100-week measurement. This is convenient because then those with a 100-week outcome, but whose baseline measurement is missing, are not regarded to be missing the endpoint. The primary outcome may therefore be referred to below as the 100-week visual acuity, rather than the change in this from baseline to 100 weeks.

The primary outcome will be analysed using a linear mixed effects (LME) model incorporating the 5 post-baseline measurements of the refracted BCVA outcome " (12, 24, 52, 76 and 100 weeks). This mixed model will have, by definition a mix of random and fixed effect terms. The random effect in the model will be *participant*, represented as a random intercept at each follow-up timepoint, with allowance for within-participant correlation in the

adjusted post-baseline outcomes. The fixed effects in the model will be the main effect terms for arm, the two stratifiers: visual acuityand disease duration, “time”, the baseline of the outcome and its missing indicator required for the missing indicator method (21). The other fixed effects to be included in the model will be the interactions between “*time*” and each of the other fixed effects in the model. This model allows the treatment effect to be formally tested at 52 weeks, at the primary timepoint of 100 weeks, and estimated at 24 and 76 weeks.

13.2 Intention to Treat Strategy

Outcome data will be valid and included if the BCVA measure is refracted. All randomised subjects who provide at least one post-baseline valid measurement will be included.

13.3 Per Protocol analysis

For the analysis of the primary outcome, the mixed effects model will be re-fitted in a reduced per protocol (PP) population already described in section 3.3.2.. Only valid (refracted) measurements will be included, and so the per protocol analysis will be a subset of the outcome measurements in the 52 and 100-week ITT analysis LME model.

13.4 Concluding non-inferiority

Non-inferiority will only be concluded if this is declared by both the ITT analysis and the PP analysis at 100 weeks. Non-inferiority will also be assessed secondarily in ITT and PP populations at 52 weeks from the same models. Non-inferiority will be declared if the estimated 95% confidence interval for the difference in means lies wholly above the margin of -5 letters in both ITT and PP analysis models primarily at 100 weeks and secondarily at 52 weeks.

13.5 Superiority

If non-inferiority is concluded, superiority will be assessed from the ITT LME model by reporting the *p*-value from the two-sided test of the hypothesis of a zero difference in population means using a 5% significance level without need for correction for multiple testing.

In addition, if both investigative treatments were considered non-inferior to the standard treatment at 100 weeks then superiority of the investigative treatments will be assessed to each other.

13.6 Subgroup analysis

The threesubgroup variables will be assessed by extending the primary outcome model to have an interaction between arm and each categorical subgroup variable. Subgroup variables with more than two categories that are ordinal will be entered as linear in the interaction. The treatment effects will also be presented within each subgroup category with a 95% confidence interval.

13.7 Sensitivity to missing data

An expert missing-data group concluded that rather than statisticians reacting to missing data at the end of a trial, there should be comprehensive, proactive planning for handling missing data at the stage of designing trials (22). The group recommended there should be consideration of missing data mechanisms (e.g Missing At Random), and, if the missing data may be informative that appropriate sensitivity analyses should be undertaken to investigate the robustness of the inferences to the different assumptions made by the main analysis. It has also been recommended that analyses allowing for non-response and low intervention uptake (or compliance) are best specified in advance and included in the analysis plan (23). As it is expected that compliance will be high from the fear of loss of sight, and as non-inferiority is concluded only when declared in both a compliant PP population and a less compliant ITT population, the focus is on handling of missing data.

A sensitivity analysis will be undertaken to assess the possibility of alternative plausible values of treatment effect arising from potential mishandling of missing data in the primary analysis model.

The LME model for the primary outcome analysis described above is the first of a two-part approach called the Intention to Treat Strategy (13) in which a second analysis examines the sensitivity of the results to missing data in the full randomised, Intention to Treat, population. This meets the ideal of ITT. The approach to missing data taken for Leavo follows the recently published implementation paper of the ITT strategy (24). This is then also applied again to the PP population so that the non-inferiority conclusion can be re-assessed under the sensitivity analysis.

For the sensitivity analysis, we pre-specify a range for best visual acuity from -20 letters to +20 letters over which the mean of the "unobserved outcome data" might *depart* (or be different) from the mean of the "observed outcome data" (24). In other words, this range can be thought of as how much a typical subject with missing data may on average have had a different estimated treatment effect compared to the corresponding subject with the outcome data observed (given the same baseline covariates and follow-up data in the LME model). The range (-20 to +20) is chosen to represent both negative and positive *departures* that could potentially arise as the "net effect" of alternative reasons which may be unknown; such as dropout due to no anticipated further improvement, or dropout due to no improvement so far together with no anticipated achievable improvement.

This range of 40 letters (from -20 to +20) is generously wide for exploring sensitivity of the main results to departures from the MAR assumption, because 20 letters (as the maximum *departure* in either direction) is larger than the detectable between-arm treatment effect of 3 lines (15 letters) seen in superiority trials (difference in means) which is a sizeable shift in the mean of the distribution for dropouts compared to completers.

At the end of the trial, the fractions of individuals with missing data for visual acuity at 100 weeks will be available in each arm f_i (for intervention) and f_c (for control). The parameter representing excess visual acuity in those missing compared to those observed, δ, *will take values by passing across the range* -20 to +20. Three scenarios will be undertaken within the sensitivity analysis (23, 24). These reflect whether departures from the MAR assumption apply within the intervention arms only (aflibercept and bevacizumab), within the control arm only (ranibizumab), or within both arms equally and in the same direction (thereby

potentially cancelling out across the sensitivity range, if the dropout rate were to be the same in both arms).

Scenario 1: the treatment effect from the LME model will be increased by $f_i\delta$
Scenario 2: the treatment effect from the LME model will be increased by $-f_c\delta$
Scenario 3: the treatment effect from the LME model will be increased by $(f_i-f_c)\delta$

13.8 Sensitivity analysis to use of concomitant treatments

The use of concomitant treatments will be monitored by the DMC. If necessary, a sensitivity analysis will be undertaken to examine the robustness of the 100-week per protocol analysis to the use of concomitant treatments.

13.9 Interim analysis

Formal interim analysis of the primary outcome for early stopping is not planned for this study. Regular interim reports will be prepared as needed for DMEC meetings.

14.0 Secondary outcome analysis

14.1 Analysis of continuous outcomes

As for the primary outcome, the analysis of continuous secondary outcomes will be compared between arms at 100-weeks using linear mixed effect model adjusting for all randomisation stratifiers and where collected, the baseline of the outcome with the associated missing indicator. Time will be represented as categorical contrasts in main effect form and in interaction with all other fixed effects.

14.2 Analysis of binary outcomes

For the binary outcomes, such as the proportion of participants with ≥15 ETDRS letter improvement, chi-squared tests will be used. Safety outcomes will be reported as unadjusted patient proportions and rates within and between arms with 95% confidence intervals using exact methods where appropriate.

14.3 Analysis methods for secondary outcomes

All study analyses will be based on tests that are two-sided, including the two-sided 95% confidence intervals.

For the secondary outcomes mentioned in section 6.2, the following analysis will be used:

Types of variables	Outcomes:	Methods:
Continuous	Best Corrected Visual Acuity at 52 weeks	Linear mixed effects model
	Mean OCT central subfield thickness (CST) at 52 and 100 weeks	Linear mixed effects model
	Macular volume at 52 and 100 weeks	Linear mixed effects model
	VFQ25 composite score, distance and near subscales at 52 and 100 weeks	Linear mixed effects model
	EQ-5D with and without vision bolt-on at 52 and 100 weeks	Linear mixed effects model
	Number of injections by 100 weeks	Difference in means
	Change in retinal non-perfusion at week 100 as assessed by two methods in different sites: i)Disc area of non-perfusion (in approx. 27 sites) ii)Ischaemic index (in approx. 13 sites)	ANCOVA (for each assessment method) and Fisher's method of combining p-values
Categorical	Participants with ≥15 and ≥ 10 ETDRS letter improvement, <15 letter loss and ≥ 30 ETDRS letter loss (severe visual loss) at 52 and 100 weeks	Chi-squared tests
	Participants with ≥73 ETDRS letters or better, ≤ 58 ETDRS letter and ≤19 letters at 52 and 100 weeks	Chi-squared tests
	Participants with OCT CST <320µm at 52 and 100 weeks	Chi-squared tests
	Persistent non-responders participants at 52 and 100 weeks	Chi-squared tests
	Participants that develop ocular neovascularisation at 52 and 100 weeks	Chi-squared tests
	Participants with OCT anatomical features: diffuse intraretinal oedema, intraretinal cystic change, subretinal fluid, vitreomacular interface abnormaility (either VMT or ERM) at 52 and 100	Chi-squared tests
	Prevalence of local and systemic side effects	Fisher's exact test

15.0 Handling multiple comparisons

Significance tests will be used sparingly and restricted where possible to addressing stated hypotheses. Secondary outcomes, as well as the primary outcome, will be summarised using an effect size with a 95% confidence interval. Interpretation for those secondary outcomes that do not directly address the stated study hypotheses will be more cautious.

16.0 Software

Data management:
An online data collection system for clinical trials (MACRO; InferMed Ltd) will be used. This is hosted on a dedicated server at KCL and managed by the MH&N CTU. The MH&N CTU Data Manager will extract data periodically as needed and provide these in comma sepa (.csv) format.

Statistical analysis:
The principal software package will be IBM SPSS Statistics 23 and R software will be available.

17.0 DMC monitoring

We expect the DMEC would want to monitor the non-inferiority of the investigational treatments in relation to the standard treatment and we would regularly provide information such as non-compliance and withdrawal and other information listed on appendix 1.

18.0 Acknowledgments

In translating the study protocol into this statistical analysis plan, we are grateful to explanations from the study team including Philip Hykin and Sobha Sivaprasad. Further versions of the plan will be commented on by members of the Data Monitoring and Trial Steering Committees.

19.0 Amendments to Versions

Version 1 was written by Joana Vasconcelos on 4th November 2013. Version 1.1 was re-written on 16th April 2014 to take into account the study protocol version 1.27.

Professor Toby Prevost verified the first version, leading to Version 2 on 21st April 2014.

Version 3 was produced on 22nd April 2014 after comments from the Chief Investigator, Phil Hykin and the co-lead investigator, Sobha Sivaprasad.

Version 4 was produced on 12th September 2014 which accounted for comments made by the DMC chair/statistician and protocol version 2.2.

Version 5 was produced on 30th September 2014 and updated on 26th November 2014 to take into account LEAVO protocol v3.0. This will be the final version approved by the independent TSC after their comments.

Amendments to versions will be listed here.

Version 5.1 was amended to Version 5.2 as a result of the DMC meeting held on 11th December, in open session, and the DMC recommendation to the TSC, discussed at the TSC meeting on 8th January 2016.

The DMC discussed the circumstances under which a BCVA score at 100 weeks, or other timepoints, would not reflect the underlying visual status of a participant. In particular, recent vitreous haemorrhages may cause low BCVA scores which would then return to normal for the patient, either spontaneously or through appropriate clinical management (vitrectomy). The challenge is that any such measurements could artificially induce very large negative changes in BCVA which would have enormous influence in statistical analysis – specifically by leading to very large inflations in the standard deviation for the change from baseline. This could have profound implications for the ability of this non-inferiority trial to achieve its objectives, which rely on the 95% confidence interval for the difference between randomised groups in the change from baseline falling within pre-specified bounds (the non-inferiority margin). As such values intrinsically do not reflect the underlying visual status of the patient, the DMC proposed that the TSC consider amending the primary analysis population measurements to exclude from analysis any refracted BCVA measurement which is both >3 SD below the mean at that timepoint (including all measurements) and taken within 3 months of occurrences of a vitreous haemorrhage. The TSC also considered a proposal from the PI that Visual Acuity loss due to other casues unrelated to maculopathy secondary to CRVO be included. The absolute number of measurements excluded across the timepoints of measurement of refracted visual acuity is expected to be small. The TSC requested confirmation that these occurrences (number and nature) will be transparently reported by arm, and this has been included in this SAP.

Version 5.2 was amended to Version 5.3 as a result of the DMC meeting held on 1st November 2016. The inclusion/exclusion criteria and the Per Protocol definition was updated to be in conformity with the Protocol version 4.0 as well as the wording of the secondary objectives. The randomisation stratifier ‘previous treatment’ was removed from the outcome analysis models as a covariate as well as a variable in subgroup analysis due to the very small number of patients having had previous treatment in the trial. Also the categories 3-6 months and >6months of the disease duration stratifier will be merged for the same reasons and will be analysed as such in the models and subgroup analysis. Finally, the method of randomisation had been mis-typed in on section 8.2 in the SAP as being stratified, whereas it has all along been minimisation. This wording has been corrected.

Reference list

1. International conference on harmonisation; guidance on statistical principles for clinical trials; availability--FDA. Notice. *Fed Regist*. 1998;63(179):49583-98.
2. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c332.
3. Committee for Proprietary Medicinal P. Committee for Proprietary Medicinal Products (CPMP): points to consider on adjustment for baseline covariates. *Stat Med*. 2004;23(5):701-9.
4. Guidance for Industry Non-Inferiority Clinical Trials.
5. Committee for Proprietary Medicinal P. Points to consider on switching between superiority and non-inferiority. *Br J Clin Pharmacol*. 2001;52(3):223-8.
6. Garrett AD. Therapeutic equivalence: fallacies and falsification. *Stat Med*. 2003;22(5):741-62.
7. Lasaffre E. Superiority, Equivalence and Non-Inferiority Trials. *Bull NYU Hosp Jt Dis*. 2008;66(2):150-42008.
8. Patient.co.uk: Egton Medical Information Systems Limited; [cited 2014]. Available from: http://www.patient.co.uk/.
9. MedlinePlus.
10. Investigators IS, Chakravarthy U, Harding SP, Rogers CA, Downes SM, Lotery AJ, et al. Ranibizumab versus bevacizumab to treat neovascular age-related macular degeneration: one-year findings from the IVAN randomized trial. *Ophthalmology*. 2012;119(7):1399-411.
11. Comparison of Age-related Macular Degeneration Treatments Trials Research G, Martin DF, Maguire MG, Fine SL, Ying GS, Jaffe GJ, et al. Ranibizumab and bevacizumab for treatment of neovascular age-related macular degeneration: two-year results. *Ophthalmology*. 2012;119(7):1388-98.
12. NICE FAD: Ranibizumab for treating visual impairment caused by macular oedema secondary to retinal vein occlusion 2013 [27-05-13].
13. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. *BMJ*. 2011;342:d40.
14. D'Agostino RB, Sr., Massaro JM, Sullivan LM. Non-inferiority trials: design concepts and issues - the encounters of academic consultants in statistics. *Stat Med*. 2003;22(2):169-86.
15. Campochiaro PA, Brown DM, Awh CC, Lee SY, Gray S, Saroj N, et al. Sustained benefits from ranibizumab for macular edema following central retinal vein occlusion: twelve-month outcomes of a phase III study. *Ophthalmology*. 2011;118(10):2041-9.
16. Mangione CM, Lee PP, Gutierrez PR, Spritzer K, Berry S, Hays RD. Development of the 25-item National Eye Institute Visual Function Questionnaire. *Arch Ophthalmol*. 2001;119(7):1050-8.
17. Janssen MF, Birnie E, Bonsel GJ. Quantification of the level descriptors for the standard EQ-5D three-level system and a five-level version according to two methods. *Qual Life Res*. 2008;17(3):463-73.
18. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health*. 2012;15(5):708-15.
19. Longworth L, Yang Y, Young T, Mulhern B, Hernandez Alava M, Mukuria C, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. *Health Technol Assess*. 2014;18(9):1-224.

20. Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Stat Med*. 2000;19(9):1141-64.
21. White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. *Stat Med*. 2005;24(7):993-1007.
22. Burzykowski T, Carpenter J, Coens C, Evans D, France L, Kenward M, et al. Missing data: discussion points from the PSI missing data expert group. *Pharm Stat*. 2010;9(4):288-97.
23. White IR, Kalaitzaki E, Thompson SG. Allowing for missing outcome data and incomplete uptake of randomised interventions, with application to an Internet-based alcohol trial. *Stat Med*. 2011;30(27):3192-207.
24. White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. *Clin Trials*. 2012;9(4):396-407.

Appendix I – Lists of Tables for DMC and for main trial

These are draft tables for the DMC meetings.

1) Recruitment by calendar month

Months:	Site 1	Site 2	Site 3	(…)	Site 40
Dec 2014 Jan 2015 Feb 2015 … Total					

2) Randomisation - Stratifiers

	Arm Ia/ Active N (% total)	Arm Ib/ Active N (% total)	Arm C/ Control N (% total)	Total
BCVA letter score ≤38 39-58 ≥59				
On onset of symptoms to presentation at hospital and commencement of therapy <3 months 3-6 months >6 months				
Treatment: Naïve Previous				
Overall				

3) Compliance

i) Up to and including 12 weeks of treatment

Weeks:	Arm Ia/ Active % (n/d)	Arm Ib/ Active % (n/d)	Arm C/ Control % (n/d)
4 8 12 **Total**			

n: Total number of patients compliant; d: Total no patients (not having withdrawn from the trial)

ii) Remaining weeks

Weeks:	Arm Ia/ Active % (n/d)	Arm Ib/ Active % (n/d)	Arm C/ Control % (n/d)
24 52 76 100 **Total**			

n: Total number of patients compliant; d: Total no patients (not having withdrawn from the trial)

4) Outcomes

Primary outcome - Visual acuity using the ETDRS letter scoring tool

Weeks:	Arm Ia/ Active Mean (SD)	Arm Ib/ Active Mean (SD)	Arm C/ Control Mean (SD)	Ia vs C z-value	Ia vs C 95% CI	Ib vs C z-value	Ib vs C 95% CI
24 52 76 100							

Secondary outcomes:

Outcomes:	Arm Ia/ Active	Arm Ib/ Active	Arm C/ Control	Ia vs C z-value	Ib vs C z-value
Change in mean OCT central subfield thickness (CST) at 52 and 100 weeks.	Mean (SD)	Mean (SD)	Mean (SD)		
Change in macular volume at 52 and 100 weeks	Mean (SD)	Mean (SD)	Mean (SD)		
Participants with ≥15 ETDRS letter improvement, ≥ 10 letter improvement, <15 letter loss and ≥ 30 ETDRS letter loss at 52 and 100 weeks.	% (n)	% (n)	% (n)		
Participants with ≥73 ETDRS letters or better, ≤ 58 ETDRS letter and ≤19 letters at 52 and 100 weeks.	% (n)	% (n)	% (n)		

Note: These will be extended and modified by the DMC.

The denominators will either be presented here or in a "Completeness of Data" Table where observed denominators are compared to expected denominators.

Appendix II – Record of data decisions during the blinded review.

21.1 Record of data decisions

During the blinded review data decisions will be recorded here.

21.2 Record of analysis decisions

During the blinded review data decisions will be recorded here.

Appendix III – Record of data decisions after the blind-break

Record of analysis decisions

i) There were some patients that were randomised using the wrong stratification category after looking at the actual continuous values at baseline. Therefore, the actual categorised baseline values of stratifying covariates, rather than those used in the randomisation which included errors, will be used in the outcome models, so that any baseline confounding by these is more fully adjusted for, and analyses are consistent with subgroup analyses using the same categorisations of these covariates. There was agreement for this approach as the trial employs minimisation. The DMC Chair approved this decision on 13th February 2019.

ii) There were three participants with BCVA scores missing at baseline due to not having completed the one meter test despite the four meter test being less than 20. These were:

Participant 11014 who had a score of 19 in the four meter test.
Participant 16171 who had a score of 17 in the four meter test.
Participant 10368 who had a score of 4 in the four meter test.

According to the eligibility criteria, which requires BCVA to be >=19, participant 11014 was in fact eligible but not necessarily the other two, who despite being randomised, have not met proof of eligibility beyond doubt as would be preferred for a per protocol population. Therefore, for the PP analysis P11014 will be included (using the missing indicator method as planned), P16171 and P10368 will be excluded, and a sensitivity analysis for the main Per Protocol analysis at 100 weeks will be carried out as follows:

a) Participant 16171 will be included in the PP population but not participant 10368.
b) Participant 10368 will be included in the PP population but not participant 16171
c) Both participants (16171 and 10368) will be included in the PP population.

This decision was agreed among statisticians on 27th February 2019 and approved by the Chief Investigator on 1st March 2019. This Document was signed between 1st March and 3rd May 2019 by Dr Joana Vasconcelos, Trial Statistician, Professor Toby Prevost, Lead Statistician, Mr Philip Hykin, Chief Investigator, Professor Sarah Walker, DMC Chair and Professor Susan Downes, TSC Chair.