Intravitreal ranibizumab versus aflibercept versus bevacizumab for macular oedema due to central retinal vein occlusion: the LEAVO non-inferiority three-arm RCT

Philip Hykin,[1,2*] A Toby Prevost,[3] Sobha Sivaprasad,[1,2] Joana C Vasconcelos,[3] Caroline Murphy,[4] Joanna Kelly,[4] Jayashree Ramu,[1] Abualbishr Alshreef,[5] Laura Flight,[5] Rebekah Pennington,[5] Barry Hounsome,[†] Ellen Lever,[4] Andrew Metry,[5] Edith Poku,[5] Yit Yang,[6] Simon P Harding,[7] Andrew Lotery,[8] Usha Chakravarthy[9] and John Brazier[5]

[1]National Institute for Health Research Moorfields Biomedical Research Centre, London, UK
[2]Institute of Ophthalmology, University College London, London, UK
[3]Imperial Clinical Trials Unit, School of Public Health, Imperial College London, London, UK
[4]King's Clinical Trials Unit at King's Health Partners, King's College London, London, UK
[5]School of Health and Related Research, University of Sheffield, Sheffield, UK
[6]The Eye Infirmary, New Cross Hospital, Wolverhampton, UK
[7]Eye and Vision Science, University of Liverpool, and St Paul's Eye Unit, Royal Liverpool University Hospitals, Liverpool, UK
[8]Faculty of Medicine, University of Southampton, Southampton, UK
[9]Department of Biomedical Sciences, Queen's University Belfast, Belfast, UK

*Corresponding author philhykin@gmail.com
†In memoriam

Declared competing interests of authors: Philip Hykin has received grants, personal fees and non-financial support from Bayer AG (Leverkusen, Germany), Novartis International AG (Basel, Switzerland) and Allergan plc (Dublin, Ireland). A Toby Prevost was a member of the National Institute for Health Research (NIHR) Public Health Research Committee (June 2014–June 2020) and is a member of the NIHR COVID Call (Recovery and Learning) Committee (June 2020–present); he has received a lecture fee from Bayer AG and a grant from Boehringer Ingelheim (Ingelheim, Germany). Sobha Sivaprasad is part of the NIHR Health Technology Assessment (HTA) Commissioning Board (31 July 2017–present) and has received personal fees and grants from Novartis International AG, Bayer AG, Allergan plc and Boehringer Ingelheim, personal fees from F. Hoffmann-La Roche AG (Basel, Switzerland), personal fees and non-financial support from Optos plc (Marlborough, MA, USA) and personal fees from Heidelberg Engineering Inc. (Franklin, MA, USA). Laura Flight has received grants from NIHR for her NIHR

Doctoral Fellowship, *Adaptive Design Clinical Trials and their Impact on the Economic Evaluation of Health Care Technologies* (DRF-2015-08-013), and another NIHR project, *Value-based Adaptive Clinical Trial Designs for Efficient Delivery of NIHR Research* (NIHR130354). Rebekah Pennington has received research grants from F. Hoffmann-La Roche AG. Yit Yang has received personal fees and non-financial support from Novartis International AG, Bayer AG and personal fees from Allergan plc and F. Hoffmann-La Roche AG. Simon P Harding has received grants to his employer from the NIHR HTA Clinical Evaluation and Trials UK and F. Hoffmann-La Roche AG. Andrew Lotery has received travel support from Bayer AG, and personal fees from Novartis International AG and Allergan plc. Usha Chakravarthy has received grants from NIHR [for the Inhibit VEGF in Age-related choroidal Neovascularisation (IVAN) trial, 07/36/01], was part of the HTA Prioritisation Committee (2010–14) and has received personal fees from Allergan plc and Novartis International AG.

Published June 2021
DOI: 10.3310/hta25380

Scientific summary

The LEAVO non-inferiority three-arm RCT

Health Technology Assessment 2021; Vol. 25: No. 38
DOI: 10.3310/hta25380

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Approximately 5200 cases of visual impairment due to central retinal vein occlusion-related macular oedema occur yearly in England and Wales and require treatment with repeated intraocular injections of antivascular endothelial growth factor agents. Treatment typically lasts for 2 years. Two agents, ranibizumab (0.5 mg/0.05 ml Lucentis®; Novartis International AG, Basel, Switzerland) and aflibercept (2 mg/0.05 ml Eylea®; Bayer AG, Leverkusen, Germany), are licensed and recommended by the National Institute for Health and Care Excellence.

An alternative low-cost option, unlicensed bevacizumab (1.25 mg/0.05 ml Avastin®; F. Hoffmann-La Roche AG, Basel, Switzerland), is utilised globally. All three antivascular endothelial growth factor agents are also used in the treatment of other retinal disorders. Despite clinical evidence that bevacizumab is non-inferior to ranibizumab and is cost-effective in neovascular age-related macular degeneration and diabetic macular oedema, it is not used in the NHS. The reasons for this include a lack of clinical evidence in certain indications, concerns over whether or not high-quality bevacizumab could be manufactured on the scale required for NHS use and the fact that it is not licensed or recommended by the National Institute for Health and Care Excellence. Therefore, in 2012, the National Institute for Health and Care Excellence Decision Support Unit recommended further comparative studies of these agents in retinal diseases, resulting, in 2014, in the development of this multicentre, Phase III, double-masked, randomised controlled non-inferiority trial comparing the clinical effectiveness and cost-effectiveness of intravitreal therapy with ranibizumab (Lucentis) versus aflibercept (Eylea) versus bevacizumab (Avastin) for macular oedema due to central retinal Vein Occlusion (LEAVO). No new antivascular endothelial growth factor agents or other treatments have superseded antivascular endothelial growth factor agents in vein occlusion-related macula oedema. Since LEAVO was initiated, the US Study of Comparative Treatments for Retinal Vein Occlusion 2 (SCORE2) trial (Scott IU, VanVeldhuisen PC, Ip MS, Blodi BA, Oden NL, Awh CC, *et al.* Effect of bevacizumab vs aflibercept on visual acuity among patients with macular edema due to central retinal vein occlusion: the SCORE2 randomized clinical trial. *JAMA* 2017;**317**:2072–87) reported the non-inferiority of bevacizumab to aflibercept with respect to visual acuity at 6 months in 362 patients with macula oedema due to central retinal vein occlusion or hemiretinal vein occlusion. A systematic review of antivascular endothelial growth factor therapy confirmed that there were no randomised controlled trials comparing all three antivascular endothelial growth factor agents in vein occlusion. LEAVO is, therefore, the first randomised controlled trial, to our knowledge, that has evaluated the comparative clinical effectiveness and cost-effectiveness of these three antivascular endothelial growth factor agents in central retinal vein occlusion-related macula oedema over the typical duration of the disease.

Objectives

The following research questions were addressed in this trial:

- Is bevacizumab non-inferior to ranibizumab in eyes with macula oedema due to central retinal vein occlusion in terms of best corrected visual acuity at 100 weeks?
- Is aflibercept non-inferior to ranibizumab in eyes with macula oedema due to central retinal vein occlusion in terms of the best corrected visual acuity at 100 weeks?
- What is the short-term and long-term cost-effectiveness of aflibercept and bevacizumab versus ranibizumab in the treatment of macula oedema due to central retinal vein occlusion?

Methods

Design

This was a multicentre, prospective, three-arm, double-masked, randomised controlled non-inferiority trial to evaluate the clinical effectiveness, cost-effectiveness and side-effect profile of three antivascular endothelial growth factor agents in the management of central retinal vein occlusion-related macula oedema over 100 weeks.

Setting

The trial was set in the ophthalmology departments of 44 UK NHS trust hospitals.

Participants

Participants were adults with visual impairment due to central retinal vein occlusion-related macula oedema of < 12 months' duration, with a visual acuity letter score in the study eye of between 19 (≈ 3/60 Snellen) and 78 (≈ 6/9 Snellen) and spectral-domain optical coherence tomography central subfield thickness of ≥ 320 µm.

Interventions

Using a web-based randomisation service, eligible patients were allocated (1 : 1 : 1) to repeated intravitreal injections of ranibizumab, aflibercept or bevacizumab by the method of minimisation, with the following factors: visual acuity (19–38, 39–58 or 59–78 Early Treatment Diabetic Retinopathy Study letters), disease duration (< 3 months, 3–6 months or > 6 months) and treatment naive (yes or no). Participants in all trial arms had a mandated injection at baseline and at 4, 8 and 12 weeks. From weeks 16 to 96, treatment was given if one or more of the predefined re-treatment criteria were met; the criteria were a decrease in visual acuity of > 5 letters between the previous and the current visit, attributed to an increase in optical coherence tomography central subfield thickness; an increase in visual acuity of > 5 letters between the previous and the current visit; an optical coherence tomography central subfield thickness of ≥ 320 µm due to intraretinal or subretinal fluid; and an optical coherence tomography central subfield thickness increase of > 50 µm from the lowest previous measurement. From week 24, the visit interval could be increased from 4 to 8 weeks if re-treatment criteria were not met at three consecutive visits. Re-treatment was withheld if visual acuity was > 83 letters; it could be suspended if there was minimal response to three consecutive injections and could be restarted if clinical deterioration occurred.

Follow-up

Participants were followed up for 100 weeks.

Clinical outcomes

The primary outcome was the change in refracted visual acuity letter score from baseline to 100 weeks in the study eye. Secondary outcomes in the study eye included a gain of ≥ 10 and ≥ 15 visual acuity letters, losses of < 15 or ≥ 30 visual acuity letters at 52 and 100 weeks, change in optical coherence tomography central subfield thickness from baseline to 52 and 100 weeks, optical coherence tomography central subfield thickness of < 320 µm at 52 and 100 weeks, and the number of injections by 100 weeks. Adverse events were recorded over the weeks.

Statistical analysis

Part of this section has been reproduced from Hykin P, Prevost AT, Vasconcelos JC, Murphy C, Kelly J, Ramu J, *et al.* Clinical effectiveness of intravitreal therapy with ranibizumab vs aflibercept vs bevacizumab for macular edema secondary to central retinal vein occlusion: a randomized clinical trial. *JAMA Ophthalmol* 2019;**137**:1256–64. This is an open access article distributed under the terms of the CC-BY 4.0 license (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium. The text includes minor additions and formatting changes to the original.

The standard deviation was anticipated to be 14.3 letters, based on the available data, and the sample size was set at 459 patients for at least 80% power to detect non-inferiority against a margin of −5 Early Treatment Diabetic Retinopathy Study letters for each intervention compared with ranibizumab using a two-sided 95% confidence interval from an analysis-of-covariance test with adjustment for baseline visual acuity. The primary outcome of refracted visual acuity was compared between the aflibercept and ranibizumab groups and between the bevacizumab and ranibizumab groups, primarily at the 100-week point, adjusting for baseline using a linear mixed-effects model and allowing for within-patient correlation of repeated measures over time using an unstructured covariance matrix. All participants with at least one milestone visit were included in the model; therefore, those without follow-up data did not contribute to the analysis. Fixed effects included the main effects and interactions with 'time' (defined as milestone visits at 12, 24, 52, 76 and 100 weeks) of treatment group, disease duration (< 3 or ≥ 3 months), the baseline of the outcome and its missing indicator required for the missing indicator method. The test for non-inferiority was one-sided at the 2.5% significance level, and presented as an estimated effect with two-sided 95% confidence intervals compared with the non-inferiority margin of −5 letters. The per-protocol population was defined as a subset of the intention-to-treat population who were eligible and received minimal sufficient treatment exposure, defined as four treatments correctly assessed and received during the first six visits. For the analysis of the primary outcome, the mixed-effects model was re-fitted in the per-protocol population. Non-inferiority was declared if the estimated 95% confidence interval for the difference in means lay wholly above the margin of −5 letters in both the intention-to-treat and the per-protocol analysis models, primarily at 100 weeks and secondarily at 52 weeks (and implicitly one-sided $p < 0.025$ for both). Analyses were completed according to the intention-to-treat strategy under a missing-at-random assumption, together with a principled sensitivity analysis in the full intention-to-treat and per-protocol populations. This assessed sensitivity to the handling of missing 100-week data using three recommended scenarios affecting either any or all groups. Secondary continuous outcomes were analysed only on the intention-to-treat basis, for superiority, and with the same model specification as for the primary outcome, except with baseline visual acuity represented by its minimisation categories, and reported as adjusted differences in means. Safety and Antiplatelet Trialists' Collaboration events were reported as proportions and compared between groups, with Wilson's 95% confidence intervals for rare events. All superiority tests were two-sided at the 5% significance level and effect sizes were interpreted cautiously with 95% confidence intervals.

Health economic analysis

The primary health economic analysis was a model-based cost–utility analysis adopting a lifetime horizon and an NHS payer perspective, using discrete event simulation modelling. The model utilised data from LEAVO, which were supplemented with evidence from external sources. Cost-effectiveness was expressed in terms of the incremental cost per quality-adjusted life-year, estimated using the Visual Function Questionnaire-Utility Index, the EuroQol-5 Dimensions and the EuroQol-5 Dimensions with vision bolt-on. A within-trial analysis was conducted as a secondary analysis. Scenario analyses considered the impact of price discounts for aflibercept and ranibizumab.

Results

Between December 2014 and 2016, eligibility was determined for 586 patients; 463 patients were randomly assigned to receive ranibizumab ($n = 155$), aflibercept ($n = 154$) or bevacizumab ($n = 154$). Participants' baseline characteristics were similar between the treatment groups. A total of 454 and 443 participants were included in the prespecified intention-to-treat and per-protocol linear mixed-effects models, and the 100-week visit was completed by 135 (87.1%) participants in the ranibizumab group, 133 (86.4%) participants in the aflibercept group and 139 (90.3%) participants in the bevacizumab group.

Clinical results

The mean gain in visual acuity letter score was ranibizumab +12.5 (standard deviation 21.1), aflibercept +15.1 (standard deviation 18.7) and bevacizumab +9.8 (standard deviation 21.4) at 100 weeks. At 100 weeks, the trial was unable to demonstrate that bevacizumab was non-inferior to ranibizumab in either the intention-to-treat (adjusted mean best corrected visual acuity difference−1.73 letters, 95% confidence interval −6.12 to 2.67 letters; $p = 0.071$) or the per-protocol population (adjusted mean best corrected visual acuity difference−1.67 letters, 95% confidence interval −6.02 to 2.68 letters; $p = 0.066$). Aflibercept was non-inferior to ranibizumab in both the intention-to-treat (adjusted mean best corrected visual acuity difference 2.23 letters, 95% confidence interval −2.17 to 6.63 letters, $p = 0.0006$) and the per-protocol populations (adjusted mean best corrected visual acuity difference was 3.49 letters, 95% confidence interval −0.91 to 7.88 letters; $p < 0.0001$), but it was not superior. At 52 weeks, aflibercept and bevacizumab were non-inferior to ranibizumab. The proportions of participants in the three groups who had a best corrected visual acuity letter gain of ≥ 15 were similar: 47% in the ranibizumab group, 52% in the aflibercept group and 45% in the bevacizumab group. There were no differences across the groups in the proportion of patients who had ≥ 10 best corrected visual acuity letter gain or < 15 best corrected visual acuity letter loss.

The adjusted difference in optical coherence tomography central subfield thickness at 100 weeks for aflibercept versus ranibizumab was −29.3 (95% confidence interval −60.9 to 2.3), whereas for bevacizumab versus ranibizumab, it was 21.9 (95% confidence interval −9.7 to 53.4). However, a significantly greater proportion of participants had an optical coherence tomography central subfield thickness of < 320 µm at 52 weeks in the aflibercept group (76%) than in the ranibizumab group (63%), a mean difference of 12.4% (95% confidence interval 1.7% to 23.1%). This was also the case at 100 weeks for aflibercept (81%) compared with ranibizumab (66%), a mean difference of 15.3% (95% confidence interval 4.9% to 25.7%), but for bevacizumab compared with ranibizumab a difference was found only at week 24 (mean difference −18.7%, 95% confidence interval −30.1% to −7.4%). The corresponding proportions at 52 weeks and 100 weeks for bevacizumab were −10.7% (95% confidence interval −22.3% to 0.9%) and −7.4% (95% confidence interval −18.9% to 4.1%).

By 100 weeks, participants in the ranibizumab group had received a mean of 11.8 injections, compared with 10.0 injections received by participants in the aflibercept group and 11.5 injections received by those in the bevacizumab group. The difference between the aflibercept and ranibizumab groups was significant at week 24 (mean difference −0.4, 95% confidence interval −0.6 to −0.2), week 52 (mean difference −1.1, 95% confidence interval −1.6 to −0.5) and week 100 (mean difference −1.9, 95% confidence interval −2.9 to −0.8). There was one case of infectious endophthalmitis in the bevacizumab group. The frequency of all ocular adverse and Antiplatelet Trialists' Collaboration-defined events occurred with an expected, and similar, frequency in the three groups.

Aflibercept became a standard of care after LEAVO was initiated, so the comparative effectiveness of aflibercept and bevacizumab became highly relevant and a post hoc analysis was conducted. This analysis showed that bevacizumab was not non-inferior to aflibercept in both the intention-to-treat (adjusted mean best corrected visual acuity difference−3.96 letters, 95% confidence interval −8.34 to 0.42 letters; $p = 0.32$) and the per-protocol populations (adjusted mean best corrected visual acuity difference −5.15 letters, 95% confidence interval −9.52 to −0.79 letters; $p = 0.47$).

Economic results

The main findings of the model-based and within-trial cost–utility analyses suggest that bevacizumab is an economically attractive alternative to the licensed products ranibizumab and aflibercept.

The model-based economic analysis found that all three antivascular endothelial growth factor agents generated similar quality-adjusted life-years. Aflibercept generated the highest costs, followed by ranibizumab and then bevacizumab. Using the Visual Function Questionnaire-Utility Index, bevacizumab generated more quality-adjusted life-years than ranibizumab and aflibercept. The mean difference in

quality-adjusted life-years between ranibizumab and bevacizumab was −0.044 (95% confidence interval −0.074 to 0.013), and the mean difference in costs was £11,873 (95% confidence interval £11,458 to £12,288), so bevacizumab was said to dominate ranibizumab, and the 95% confidence interval for the incremental net monetary benefit at £30,000 per quality-adjusted life-year was −£14,316 to −£12,067. The mean difference in quality-adjusted life-years between aflibercept and bevacizumab was −0.109 (95% confidence interval −0.161 to −0.057), and the mean difference in costs was £4800 (95% confidence interval £4445 to £5154), so bevacizumab was said to dominate aflibercept, and the 95% confidence interval for the incremental net monetary benefit at £30,000 per quality-adjusted life-year was −£21,864 to −£18,040. The mean difference in quality-adjusted life-years between aflibercept and ranibizumab was −0.065 (95% confidence interval −0.097 to −0.033), and the mean difference in costs was £4800 (95% confidence interval £4445 to £5154), so ranibizumab was said to dominate aflibercept, and the 95% confidence interval for the incremental net monetary benefit at £30,000 per quality-adjusted life-year was −£7917 to −£5603. The finding that bevacizumab was the most cost-effective intervention was robust to scenario analyses. The costs of aflibercept and ranibizumab would need to be discounted by at least 95% to be comparable to the cost of bevacizumab (at £28 per injection over a patient's lifetime).

In the within-trial base-case analysis, the difference in mean total costs was £1245 between aflibercept and ranibizumab (95% confidence interval £421 to £2070), −£6760 between bevacizumab and ranibizumab (95% confidence interval −£7546 to −£5973) and £7984 between aflibercept and bevacizumab (95% confidence interval £7209 to £8759). Bevacizumab was dominant (less costly and with no difference in benefit) compared with ranibizumab, with a probability of cost-effectiveness of 1.00 at the £20,000 per quality-adjusted life-year threshold. Aflibercept was more costly than ranibizumab, with a mean quality-adjusted life-year difference of 0.004 (95% confidence interval −0.0430 to 0.0518), an incremental cost-effectiveness ratio of £283,595 per quality-adjusted life-year gained and a probability of cost-effectiveness of 0.04 at the £20,000 per quality-adjusted life-year threshold. Aflibercept was dominated by bevacizumab (more costly, with a mean quality-adjusted life-year difference of −0.015, 95% confidence interval −0.0618 to 0.0322) with a probability of cost-effectiveness of 0.00 at both the £20,000 and the £30,000 per quality-adjusted life-year thresholds.

Conclusions

All three antivascular endothelial growth factor agents are effective therapies for macula oedema secondary to central retinal vein occlusion, with no differences from a safety perspective. Although aflibercept was demonstrated to be non-inferior to ranibizumab, the trial was unable to demonstrate that bevacizumab was non-inferior to either, meaning that we cannot rule out the possibility that bevacizumab may be worse by 5 visual acuity letters. However, patients' health-related quality-of-life assessments were similar across treatment groups, and bevacizumab was found to be the most cost-effective option. The trial results are, therefore, divergent. We believe that bevacizumab could be introduced into the NHS as a first-line agent for this condition only after a review of these results and in agreement with patients, their representatives and funders. If patients are fully informed and understand the clinical results of the trial, as our small post-trial patient questionnaire suggests, a majority may consent to bevacizumab treatment with the proviso that licensed medications be available to them as an option if their response to bevacizumab is less than expected. If adopted, bevacizumab would result in substantial savings to the NHS, and potentially to health-care systems around the world.

Trial registration

This trial is registered as ISRCTN13623634.

Funding

This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 25, No. 38. See the NIHR Journals Library website for further project information

Health Technology Assessment

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 3.370

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, the Cochrane Library and Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 11/92/03. The contractual start date was in June 2014. The draft report began editorial review in August 2019 and was accepted for publication in June 2020. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).