Reviewer's report

Title: Comparison of Eylea(R) with Lucentis(R) as First-line Therapy in Patients with Treatment-naïve Neovascular Age-related Macular Degeneration in real-life clinical practice: Retrospective case-series analysis

Version: 2  Date: 20 May 2015

Reviewer: Inês Marques

Reviewer's report:

The manuscript "Comparison of Eylea® with Lucentis® as First-line Therapy in Patients with Treatment-naïve Neovascular Age-related Macular Degeneration in real-life clinical practice: Retrospective case-series analysis" presents a retrospective comparative study. The authors aimed to evaluate the therapeutic effect of 2 anti-VEGF drugs, Aflibercept and Ranibizumab, in the treatment of treatment naïve nvAMD patients. They evaluate this response using visual acuity as well as OCT parameters. This is an interesting research topic, not widely explored in the literature, and with a pertinent question. However, in my opinion, the manuscript requires substantial improvement, thus justifying a major review. The fact of being retrospective is a major concern, with all the drawbacks of this kind of analysis, namely the selection bias in the choice of the treatment arm as well as the possible follow up missings which was not reported in the present article. Furthermore, I would like to suggest and explain some concerns that I will detail below.

Major Compulsory Revisions

Background
1. In the last sentence of the section the authors say “treatment naïve nvAMD and similar prognosis…”. In my opinion they cannot conclude that because no analysis was done in terms of equivalence between groups and baseline characteristics.

2. Additionally, the novelty and relevance of this manuscript should be emphasized. What are the current gaps in knowledge in this field? There is no current data comparing both 3+PRN dose of both Ranibizumab and Aflibercept as the first line therapy in treatment naïve patients. Would Aflibercept require less frequent injections than Ranibizumab?

Methods

It was a retrospective analysis of consecutive patients, identified by electronic medical records as having newly diagnosed nvAMD. The bias of collecting data from electronic records is not negligible, as many important conditions are not taken into account. The authors made a reference of handwritten clinical records search for exclusion criteria, which could improve its detection. The observation period was one year with monthly follow up were BCVA and OCT was
performed.

3. Included subjects: Please define the inclusion criteria further. Did you considered patient with a long evolution of loss of vision, did you considered all the types of CNV?

4. Exclusion criteria: When authors mention the exclusion criteria they put the “n” of patients that were excluded. But they started with 27 eyes, so I think it is not useful and can confuse the reader. Or they can start with the number of eyes that met the inclusion criteria and then limit the number considering the exclusion criteria.

5. Exclusion criteria: What is the meaning of: “When conditions stabilized, defined as absence of indication for intravitreal injection for at least 6 months, patients were henceforth cared for by their attending physician”. Doesn’t make sense here, it means that if after 3 injections the patient doesn’t need further treatments after 9 months he can be withdraw…?

6. Third paragraph- Treatment choice: “Whether patients were started on Ranibizumab or Aflibercept was decided at the practitioner’s discretion at first visit and under the assumption of similar efficacy as no proof of superiority for the one or the other substance in specific features of nvAMD was available. Therefore, the assignment was considered quasi-random, even more so as several practitioners were involved”. In my opinion you can not make these considerations, assuming that the distribution between study arms were random, first because the treating physician was not the same and can differ in personal choices and also because there are reports favoring Aflibercept in some kind of lesions, which can be a source of selection bias. This question should be addressed in the results section when comparing baseline characteristics between both arms.

7. Fourth paragraph- Retreatment criteria: When you talk about the retreatment criteria have you considered loss of VA as a retreatment criteria or you based only on OCT characteristics? Why did you choose 20 um as a cut off for the retinal thickness?

8. Fifth paragraph- Demographic data: This paragraph needs a major review. Age, gender and other demographic characteristics as CNV type should be presented in a Table 1 (Comparison of baseline characteristics between the Ranibizumab and Aflibercept group), with the correspondent p values. I would suggest presenting the baseline BCVA and CFT values as well as the mean change in the analyzed time points and the number of treatments in another table. The p values should be present for all the given data.

9. Primary and secondary outcomes should be better defined. The authors refer to it twice: “Main outcome measures were defined as change in BCVA, (CHANGE) in CFT measured by OCT and number of intravitreal injection needed during a treatment period of 12 months”. And in the 7th paragraph: “Outcome was measured by change in BCVA and change in CFT measured by OCT after a treatment-period of 12 months with either substance compared to the values obtained after the loading dose, therefore 4 months after the initiation of treatment, which is consistent with the beginning of the PRN regimen…”. Please
define it better in one sentence.

10. Methods, 5th paragraph- Clinical procedures, BCVA assessment: The authors should state who performed the VA assessment (ETDRS), was it a technician or the treating doctor that was aware of the study drug? This can be a major source of bias...

11. Methods, 5th paragraph- Clinical procedures, OCT protocol: The authors should state who performed those exams, if only one technician or more. Furthermore, the acquisition protocol that was used to obtain the CFT data should be detailed and stated if it was the same for all the patients studied. It is essential that it is clearly stated in this methods’ section.

12. Statistical analysis, first paragraph: The authors stated, “We summarised continuous variables with means and standard deviations”. The study sample is a small one (16±11), so it is essential to test whether the variables were normally distributed. Normality of measures should be evaluated using the Shapiro–Wilk test. If not, and because the means are affected by extreme values, the median should be presented as well as the IQR or min and max values as a descriptive of the central tendency for all the tested variables. Furthermore, the t test can only be used if the normality of the values is assured or a sensitivity analyses should be done using the equivalent non-parametric test (U Mann Whitney).

13. When the authors refer to the association between change in BCVA or CFT between both groups they used a Two multivariate random intercept models. However in the abstract they say they have used a Three multivariate random intercept models. I don’t understand why did they use a random model. With a little sample as the one we have, it is difficult to find a statistical significant difference. Furthermore, if we do a multivariate analysis accounting for baseline values this is even more difficult. I would suggest that the authors create a variable “change in BCVA” and “change in CFT” and do all the analysis in a simple model, using for that a paired t test or the non parametric equivalent (Wilcoxon rank test) when appropriate, in the different time points.

Results

In order to what have been told, the baseline characteristics of the population should ne entered in a classic “Table 1” chart with mean±SD or median±IQR were all the p values should be presented. Also, another table should accomplish for the VA and CFT values in the analysed time points in both groups, with the correspondent p values. You should state the mean difference observed in each of the treatment group and the corresponding p values obtained with the appropriate statistical test.

14. Second paragraph: The results should be presented in a clear way (when the authors say “there was no difference within one year of follow-up ((-0.97 letters (95% CI. -6.06 to 4.12); p=0.709)) it is not clear what does -0.97 letters refers to (is it to the mean difference observed in the entire sample or the difference between study groups ate one year, and if so which group had the best achievements?). I think it would be good to present the median values here;
given the wide range of values I suspect that the values doesn’t have a normal behaviour…

15. Third paragraph, Effect of loading dose:
Once again I think authors should present the median values (or prove normality of the sample). The difference between treatment groups should be tested with the appropriate statistical test. Once again, given the small sample size, it is difficult to find a statistical difference.

16. Fourth paragraph, Effect during the follow up: Here the authors present the VA results in terms of logMAR. It was the first time in the article and in my opinion the authors should maintain the number of letters as before.

17. Fourth paragraph, Effect during the follow up: Again, the time point of the presented analysis should be clarified (was that at 12 months follow up?) - please enter theses data in the appropriate table that was suggested before. The respective p values for comparison between groups should be presented, both in the 4th and the 12th months. As mentioned before it is not clear what “trend for a small increase of CFT (0.40 µm)” refers to (all the patients or difference between groups)?

Discussion

18. Main findings: The first sentence should be reformulated. The authors could not say that Aflibercept “showed no advantages” over Ranibizumab, they only could say that no advantages were shown in this retrospective cohort, which is different. The study was not powerful enough to allow this conclusion. Furthermore, “our study revealed no clinical difference between changes of the CFT”- The concept of clinical difference appears for the first time and should be clarified. It will be better to say, “no statistically significant differences were found between CFT changes along 1 year of follow up between both treatment groups.

19. The “considerable baseline differences” that the authors refer in the discussion section should be referred and better explained in the results section.

20. Implications for practice:
“As we applied the very commonly used PRN treatment regimen, our results inform a broad sector of nvAMD management.”- The authors should explain what do they mean by that in another way.

21. Implications for practice: The authors say “Our preliminary data showing equality in terms of use and visual outcome suggest that both drugs can be equally promoted as first-line therapies in treatment of nvAMD”. In my opinion it is better to say “The retrospective analysis made in this cohort of patients did not detect significant differences between… suggesting that both drugs could be used as first line therapies ”

22. “Latest evidence [15-22] was able to prove Aflibercept’s utility as a 'salvage therapy' in cases of suspected tachyphylaxis [23-25] to Ranibizumab, with gain in visual acuity and reduction of
after conversion”.
Reference 18 doesn’t study Aflibercept. This sentence is not completely correct. References 15,17,20 and 22 showed visual acuity stability; only ref 16 and 21 showed some VA improvement. Furthermore, there are other important studies as: Visual and anatomical outcomes following intravitreal aflibercept in eyes with recalcitrant neovascular age-related macular degeneration: 12-month results, Grewal DS1 et al, Eye (Lond) or Injection frequency and anatomic outcomes 1 year following conversion to aflibercept in patients with neovascular age-related macular degeneration, Messenger WBet al, Br J Ophthalmol (Sep 2014), with a big sample size, that also agree in visual acuity stability in converted cases.

Conclusion
23. “Both drugs were equivalent in terms of number of injections and effect on the visual outcome”.
This conclusion should be reformulated. For instance: “The present results suggest that both drugs are equivalent in terms of visual and anatomical outcomes. However, further well design comparative studies…”

Abstract
Considering all the comments detailed above, I think that the abstract should also be improved. As I mentioned before the groups didn’t have a “similar prognosis”.
24. The authors talk about a “three multivariate random intercept models” weather in the statistical analysis they mentioned “two multivariate random intercept models”. When authors say that there were no differences in VA and CFT between both groups they should clarify in which time point.

Minor Essential Revisions
1. Background: “Ranibizumab[1, 2] (Lucentis®, Gentech Inc., South San Francisco, CA, USA and Novartis AG, Basel Switzerland) in neovascular age-related macular degeneration (nvAMD) management ”[1, 2] should be after.
2. Methods, Third paragraph: “Therefore, patients with pre-existing cataract were still included”- this sentence is a repetition of the previous one.
3. Methods, Fifth paragraph: “Main outcome measures were defined as change in BCVA, (add CHANGE) in CFT measured by OCT and number of intravitreal injection needed during a treatment period of 12 months”.
4. Methods, 6th paragraph: “of the inner limiting membrane and Bruch (and not basal) membrane”
5. Results, 4th paragraph, Effect during the follow up: “increase of CFT (0.40 µm/d …)” should be “increase of CFT (0.40 µm)”.

Discretionary Revisions
1. Background: “To what extent the advantages of Aflibercept found for example in registration studies VIEW 1[3] and 2[4] materialize in real-life clinical practice is unclear.” - I would take out this sentence, for me it doesn't make sense here.
2. Results: I would suggest the authors to perform a subgroup analysis in terms of visual and anatomical outcomes according to the type of CNV lesion, once they have made that characterization, and to detect any difference in the treatment response.

3. Discussion, Implications for research: In my opinion this section should be approached after mentioning the main findings of this study, where you mention the “findings in the context of the literature”. It is a bit confusing and repeated. Authors should refer to their findings and compare it to the existing literature. The VIEW study result, which was the boost for this study, were already mentioned in the background and should not be repeated.

4. Discussion, last paragraph: Clinical data suggest that serous PED respond well to Aflibercept, as was also apparent in reference 16 and 20. It would be of interest to do a qualitative subgroup analysis in these patients and compare it with data from literature.

**Level of interest:** An article of limited interest

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests' below. If your reply is yes to any, please give details below.