**Author's response to reviews**

**Title:** Multifocal VEP and OCT findings in patients with primary open angle glaucoma: A cross-sectional study

**Authors:**

Marilita M Moschos (moschosmarilita@yahoo.fr)
Gerasimos Georgopoulos (glaucomanet@aim.com)
Irini P Chatziralli (eirchat@yahoo.gr)
Chryssanthi Koutsandrea (koutsandrea@med.uoa.gr)

**Version:** 3 **Date:** 26 June 2012

**Author's response to reviews:** see over
Dear Editor,

I am sending for publication to your distinguished Journal the revised version of our manuscript entitled: “Multifocal VEP and OCT findings in patients with primary open angle glaucoma”.

We believe that it merits consideration for publication, as it describes the joint use of mfVEP and OCT in monitoring glaucoma progression. Both methods of examining the optic nerve head are very important as they offer objective analysis of the anatomic and functional impairment of the optic nerve in glaucoma patients, without needing the patient’s collaboration, as standard perimetry does.

Please find the point-to-point response to reviewers’ comments within the following pages. Please also note that all changes in the manuscript are highlighted in red colour.

I declare that:

i) the content of this paper has not been published or submitted for publication elsewhere

ii) the protocol of the research has been approved by the Institutional Review Board and it conforms to the provisions of the Declaration of Helsinki

iii) all authors are in agreement with the content of the manuscript

iv) there is no financial support or relationship that may pose conflict of interest.

Thank you for your consideration.

Yours sincerely,

Marilita M. Moschos MD, PhD
(corresponding author)
Response to reviewers

Reviewer 1

“In this cross sectional study the authors evaluated the anatomical and functional changes of optic nerve in eyes with primary open angle glaucoma (POAG) by the joint use of optical coherence tomography (OCT) and multifocal visual evoked potentials (mfVEP). They conclude that the joint use of mfVEP and OCT could be useful in better monitoring glaucoma progression.

This is an interesting study although several issues should be addressed.”

We would like to thank you for your encouraging comments and for the opportunity you give us to revise our manuscript and present our work in a more accurate and scientific way. Please note that all changes throughout the manuscript are highlighted in red colour.

“In the Abstract the statistical significance of RNFL and VEP reduction is not evident. Please add the p values.”

We added the p values in the abstract, according to your suggestion.

“Was gonioscopy evaluated?”

We have performed gonioscopy and all patients have a Schaffer III-IV angle. We added this comment in methods section.

“Were limits of spherical refraction and cylinder correction among inclusion criteria?”

All patients had a spherical refraction less than ±3.0 D and a cylinder correction less than ±1.5 D. We added this comment in methods section.

“Definition of visual field defect on SAP should be given.”

All patients presented a typical glaucomatous visual field defect, either nasal step or Bjerrum/altitudinal scotoma. We added this comment in methods.

“Association between VEP amplitude and RNFL thickness should be measured (i.e. with logistic regression). If possible, association between VEP amplitude and macular thickness and optic disc topography measurements (rim volume, rim area) should be evaluated.”

In order to perform a logistic regression analysis we need a dichotomous outcome that could be predicted by one or more variables. We believe that there are some main limitations in performing logistic regression analysis in our dataset: (a) although it seems feasible to discuss about correlation between VEP amplitude and RNFL thickness, it is not feasible to look for a causal relationship between these two variables. How to clarify the direction of such a causal relationship between VEP amplitude and RNFL thickness? Is VEP amplitude responsible for the changes in mean RNFL thickness, or in the opposite site, is RNFL thickness responsible for the changes in mean VEP amplitude? (b) In VEP amplitude there are three areas (rings 1, 2 and 3) whereas in RNFL thickness there are four areas (superior, temporal, inferior
and nasal). How can we compare the three areas of one variable with the four areas of the second variable? In addition, when conducting multiple comparisons, we should not forget that we run the risk of inflating the error in our analysis. However, Pearson correlation coefficient was used in order to examine possible association between VEP amplitude and RNFL. All analyses were performed separately in the group of POAG eyes and in the group of control eyes. Firstly, we looked for an association between average RNFL and VEP amplitude in Ring 1. According to the findings, there was strong evidence for a positive correlation between the two variables in the group of POAG eyes (r=0.45, p=0.016), whereas in control eyes RNFL and VEP amplitude in Ring 1 were uncorrelated (r=-0.01, p=0.985). Subsequently, we tested for correlation between average RNFL and average VEP amplitude (average of Rings 1, 2 and 3). We found statistical significant evidence for a positive correlation between average RNFL and average VEP amplitude (r=0.43, p=0.021) in the POAG group, and no evidence for correlation in the control group (r=-0.03, p=0.932).

We added the results of logistic regression analysis in results section.

“Although the study is about POAG no intraocular pressure values are showed. Please add mean IOP values in the manuscript and in Table 1.”

All patients were under topical medication for the control of intraocular pressure and IOP values were normal. We provided the specific IOP values in Table 1 and describe in methods section as well.

“Reference for definition of POAG should be provided.”

We provided the respective reference.


“ROC curves for the two diagnostic tests should be provided.”

According to the statistician analysis, who works as a lecturer in the Epidemiology Department of Athens University, the data of the current study is not suitable to perform ROC curve analysis due to its relative small sample size. Generally it has been suggested that meaningful qualitative conclusions can be drawn from ROC experiments performed with a total of about 100 observations (Metz, 1978). A minimum of 50 cases may be required in each of the two groups, so that 1 case represents not more than 2% of the observations. Based on this evidence we preferred to limit our analysis and assess the differences between cases and controls without the use of ROC curve analysis. Nevertheless according to your suggestion we performed ROC curve analyses separately for (a) retinal response density (mf-VEP), (b) implicit (mf-VEP) and (c) RNFL thickness (OCT). As you can see (eg for retinal response density) the sensitivity and the specificity for Ring1 and Ring2 were estimated at 100%. We doubt about those extremely high estimates and we believe that are due to the relative small sample size of the study. We present the findings of ROC curve analysis as requested but we are not in favor of presenting them in the manuscript:
ROC curve analysis for retinal response density (mf-VEP) comparing the three rings (Figure 1)
Ring1: Sensitivity=100%, Specificity=100%, Criterion≤84, AUC=1.00 (0.92-1.00), p<0.0001
Ring2: Sensitivity=100%, Specificity=100%, Criterion≤16, AUC=1.00 (0.92-1.00), p<0.0001
Ring3: Sensitivity=100%, Specificity=78.6%, Criterion≤6, AUC=0.96 (0.85-0.99), p<0.0001

ROC curve analysis for implicit (mf-VEP) comparing the three rings (Figure 2)
Ring1: Sensitivity=31.0%, Specificity=100%, Criterion≤100.5, AUC=0.47 (0.31-0.63), p=0.70
Ring2: Sensitivity=44.8%, Specificity=100%, Criterion>119.4, AUC=0.52 (0.36-0.67), p=0.87
Ring3: Sensitivity=51.7%, Specificity=100%, Criterion>114.2, AUC=0.60 (0.44-0.74)

ROC curve analysis for RNFL thickness (OCT) comparing the four areas (Figure 3)
Superior: Sensitivity=82.8%, Specificity=100%, Criterion≤100, AUC=0.92 (0.79-0.98), p<0.0001
Temporal: Sensitivity=75.9%, Specificity=85.7%, Criterion≤66, AUC=0.85 (0.71-0.94), p<0.0001
Inferior: Sensitivity=89.7%, Specificity=100%, Criterion≤117, AUC=0.96 (0.86-0.99), p<0.0001
Nasal: Sensitivity=86.2%, Specificity=92.9%, Criterion≤69, AUC=0.88 (0.74-0.96), p<0.0001

Figure 1: ROC curve analysis for retinal response density (mf-VEP)
Figure 2: ROC curve analysis for implicit (mf-VEP)

Figure 3: ROC curve analysis for RNFL thickness (OCT)
“Figure 1 and Figure 3 are not strictly necessary and could be deleted.”
We deleted Figures 1 and 3, according to your recommendation.
Reviewer 2
We would like to thank you for your valuable comments and for the opportunity you give us to revise our manuscript and present our work in a more elaborate and scientific way. Please find all changes within the manuscript highlighted in red colour.

“Abstract
1. In several places reduce the number of significant figure reported from 4 to 3, as in “34.17±17.62” to “34.2±17.6” unless you are claiming to have measured these quantities to 1 part in 10,000. Similarly in Tables 2, 3, 4 etc.
2. Replace “This means that the RNFL thickness was lower than normal in all the peripapillary areas. Nevertheless the decrease was higher but not statistically significant in the inferior 3. and superior area.”
4. With something like “There was a non-significant trend towards thinner superior and inferior peripapillary RNFL”, at least I thought that was what you meant but Table 4 says otherwise, please clarify.”
We made the respective changes in abstract section, according to your request.

“Introduction
5. Change “objectively detect the optic disc” to “measure the optic disc” or “quantify the optic disc”
We made the respective change.

Methods
“6. When you say the patient’s VF defects were repeatable, on how many repeats?-scans), what perimeter and what test? How was the CD ratio determined? What was the BCVA cut-off? (ah ha I see you have it in the Results, move it here for clarity).”
We performed VF test twice for each patient, with an interval of 30 minutes for the patient to rest. We used Humphrey 24-2 testing algorithm, 54 individual points are tested, and the threshold value calculated for that point is compared to a database of normally-sighted individuals of similar age. Based on this comparison, the value for the threshold value at this location is classified as being normal, or abnormal at a 5% , 2%, 1% or 0.5% probability. Each individual location has a calculated deviation from the expected threshold value for a person of the same age and ethnicity. The CD ratio was determined funduscopically by two independent examiners. We added this comment in methods. We added a comment about BCVA in methods as well.

“7. The company is Carl Zeiss Meditec in Dublin CA and the model is the Cirrus 3000 HD-OCT
The company is Carl Zeiss Meditec (Dublin CA) and the model is Stratus OCT 3000.
8. on page 4 replace the second “emitting” with something like “producing”
We made the requested change within the manuscript.

9. replace the “.” In “1.024” with a “,” and all other instances if any
We made all requested changes within the manuscript, according to your suggestion.

10. you could just say “on a unit circle” the first time and remove them for the other arcs. Why did you give the data on 512 a-scans/b when you didn’t do volume scans but peripapillary scans? How did you align the 3 repeated scans?
We made the recommended change.
Most of studies measure RNFL thickness, measured by μm. This is the reason that we preferred not to show our results by volume. We took the average of the three repeated scans and used the mean value.

“11. When describing angles use the degree symbol not the numeral 0. Also why not be consistent and use the symbol for the arcs instead of the word degree? Also why not just say that “the borders of the rings fell at 0.5, 3.0, 7.0, 12.0, 18.0 and 25 deg retinal eccentricity”? “deg being the ISI unit.”
We changed the manuscript, as requested.

12. Change “Subjects were viewed with” to “Subjects wore”.
We changed the phrase in the manuscript.

13. What was the colour of the fixation cross? Also please give the colour temperature of the white of the display.
The fixation cross was red. We added this comment in the manuscript. The colour temperature of the white of the display was 6500k.

“14. Units such as cm should be separated by a space from the numerical values, i.e. 2 cm not 2cm. Correct everywhere.”
We made the correction throughout the whole manuscript.

Results
“15. The results in Table 2 are fine but we need to know if they are clinically useful so some attempt at ROC analysis is needed, providing say the sensitivity at some reasonable specificity (10%?). After all you are claiming to look at the conjoint clinical value of doing both tests. Likewise for the RNFL results of table 4.”
According to the statistician analysis, who works as a lecturer in the Epidemiology Department of Athens University, the data of the current study is not suitable to perform ROC curve analysis due to its relative small sample size. Generally it has been suggested that meaningful qualitative conclusions can be drawn from ROC experiments performed with a total of about 100 observations (Metz, 1978). A
minimum of 50 cases may be required in each of the two groups, so that 1 case represents not more than 2% of the observations. Based on this evidence we preferred to limit our analysis and assess the differences between cases and controls without the use of ROC curve analysis. Nevertheless according to your suggestion we performed ROC curve analyses separately for (a) retinal response density (mf-VEP), (b) implicit (mf-VEP) and (c) RNFL thickness (OCT). As you can see (eg for retinal response density) the sensitivity and the specificity for Ring1 and Ring2 were estimated at 100%. We doubt about those extremely high estimates and we believe that are due to the relative small sample size of the study. We present the findings of ROC curve analysis as requested but we are not in favor of presenting them in the manuscript:

**ROC curve analysis for retinal response density (mf-VEP) comparing the three rings (Figure 1)**

Ring1: Sensitivity=100%, Specificity=100%, Criterion≤84, AUC=1.00 (0.92-1.00), p<0.0001  
Ring2: Sensitivity=100%, Specificity=100%, Criterion≤16, AUC=1.00 (0.92-1.00), p<0.0001  
Ring3: Sensitivity=100%, Specificity=78.6%, Criterion≤6, AUC=0.96 (0.85-0.99), p<0.0001  

**ROC curve analysis for implicit (mf-VEP) comparing the three rings (Figure 2)**

Ring1: Sensitivity=31.0%, Specificity=100%, Criterion≤100.5, AUC=0.47 (0.31-0.63), p=0.70  
Ring2: Sensitivity=44.8%, Specificity=100%, Criterion>119.4, AUC=0.52 (0.36-0.67), p=0.87  
Ring3: Sensitivity=51.7%, Specificity=100%, Criterion>114.2, AUC=0.60 (0.44-0.74)  

**ROC curve analysis for RNFL thickness (OCT) comparing the four areas (Figure 3)**

Superior: Sensitivity=82.8%, Specificity=100%, Criterion≤100, AUC=0.92 (0.79-0.98), p<0.0001  
Temporal: Sensitivity=75.9%, Specificity=85.7%, Criterion≤66, AUC=0.85 (0.71-0.94), p<0.0001  
Inferior: Sensitivity=89.7%, Specificity=100%, Criterion≤117, AUC=0.96 (0.86-0.99), p<0.0001  
Nasal: Sensitivity=86.2%, Specificity=92.9%, Criterion≤69, AUC=0.88 (0.74-0.96), p<0.0001  

**Figure 1:** ROC curve analysis for retinal response density (mf-VEP)
Figure 2: ROC curve analysis for implicit (mf-VEP)

Figure 3: ROC curve analysis for RNFL thickness (OCT)
“16. You should use something like Garway-Heath’s structure function map to and rather than pooling your 3 rings, pool mfVEP regions that correspond to the sectors of the optic disc. Then you could compare your 4 mfVEP measures with your 4 optic disc measures, and do ROCs for both. You could then do a combined analysis also to see if measuring both things increases sensitivity and specificity or not. If both instruments measure highly correlated things there will be no improvement. That seems unlikely but if you really want to quantify the value of doing both tests you need to do something like this.”

The idea of using Garway-Heath’s structure function map is actually useful, since it makes it possible to compare the three areas of VEP amplitude (rings 1, 2 and 3) with the four areas of RNFL thickness (superior, temporal, inferior and nasal). It helps to avoid the conduction of multiple comparisons and therefore it prevents us from the risk of inflating the error in our analysis. Nevertheless, the next step would be to perform ROC curves analysis and, as stated previously, the execution of this type of analysis has the limitation of the relatively small sample size of our dataset. Therefore although, at first sight, useful the Garway-Heath’s structure function map, it does not help us to overcome the problem of ROC curves analysis.

“17. Finally you might not want to arbitrarily throw away your outer ring data, try the analysis with and without the data in a let the data tell you what is best.”

Given the relatively small sample size of our dataset, we believe that we do not have the opportunity to throw away the outer ring data. Such an idea would minimize our dataset, rendering any kind of further analysis impossible.
18. Please also tell us what proportion of patients showed defects within the central 12 deg, i.e. the visual field locations where your mfPOP data comes from. Admittedly a coarse 24-2 test doesn’t give you much data to go on but you should report this, as the denser mfVEP sampling may be an asset.

Unfortunately, when the data were collected in 2009-2010, the mfPOP was not widely used, and the logistic was not included in the VERIS mfERG recording. After your valuable comment we asked the representator of VERIS if an upgrade of the VERIS could include mfPOP recording. The answer was negative. We are currently trying to find a representator to bring this machine in Greece, in order to have mfPOP in future studies. The department of Ophthalmology of Athens University did not have such a machine to make these measurements when the patients were recruited. In will be definitely included in a future study.

19. It would be nice if in a few cases that showed no defect in the central 12 to get those patients back and do a 10-2 test. That would enhance the value of your paper for relatively little effort (or generate another paper?)

We would like to thank you for your valuable comment. It could be the purpose of a future study, together with mfPOP data.

“Discussion
20. Please give the citations nearer in the text for “Klistorner et al and Rodarte et al”, and its “et al.” al. being and abbreviation.”

We made the respective changes, according to your recommendation.