Diagnostic accuracy of optical coherence tomography Inter-Eye Percentage Difference (IEPD) for optic neuritis in multiple sclerosis

Danko Coric, MD [1,2], Lisanne J. Balk, PhD [1,2], Bernard M.J. Uitdehaag, MD, PhD [1], Axel Petzold, MD, PhD [1,2,3,4]

1. Department of Neurology, VU University Medical Center, Amsterdam, the Netherlands

2. Expertise-center Neuro-ophthalmology, VU University Medical Center, Amsterdam, the Netherlands

3. Moorfields Eye Hospital, City Road, London, UK

4. The National Hospital for Neurology and Neurosurgery, Queen Square, London, UK

Contact information corresponding author:

Dr. A. Petzold

VU University Medical Center Amsterdam

Mailbox 7057, 1007 MB Amsterdam, the Netherlands

a.petzold@ucl.ac.uk, phone +31-20-4445292

Word count manuscript: 3294

Keywords: multiple sclerosis, optic neuritis, optical coherence tomography, diagnostic criteria

DECLARATION OF CONFLICT OF INTEREST

ACKNOWLEDGEMENTS

ABSTRACT

Background: Multiple sclerosis (MS) associated optic neuritis (MSON) causes atrophy of the inner retinal layers, which can be quantified by optical coherence tomography (OCT). It has been suggested that the Inter-Eye Percentage Difference (IEPD) of atrophy may be of diagnostic value in MSON.

Methods: Prospective, longitudinal study in MS patients and healthy controls (HC). Spectral-domain OCT of both eyes was performed, followed by automated retinal layer segmentation for the peripapillary retinal nerve fibre layer (pRNFL) and macular ganglion cell and inner plexiform layer (mGCIPL). Receiver Operator Characteristics Curves were plotted and the area under the curve (AUC) was calculated for group comparisons of the IEPD of the pRNFL and mGCIPL ~~at baseline and 2 year follow-up~~.

Results: There were 39 patients with bilateral MSON, 62 with unilateral MSON, 106 without MSON and 63 HC. Diagnostic accuracy (AUC) of the IEPD was 0.73-0.86 for pRNFL and 0.75-0.94 for mGCIPL. The diagnostic sensitivity of the mGCIPL IEPD was 70% with a specificity of 97% for distinguishing unilateral MSON from HC at baseline. For the comparison of bilateral MSON with HC sensitivity was 86% with a specificity of 97%. ~~Validation at 2 year follow up demonstrated sensitivity/specificity levels of 71%/97% (unilateral vs. HC) and 83%/97% (bilateral vs. HC).~~

Conclusion: The IEPD of the pRNFL and more so the IEPD of the mGCIPL is an useful diagnostic measure for MSON. The IEPD is a dimensionless unit and may therefore contribute to overcome device and proprietary segmentation algorithm limitations.

INTRODUCTION

The intimate relationship between optic neuritis (ON) and multiple sclerosis (MS) has long been recognised [1]. In fact, one of the most frequent clinically isolated syndromes remains MS associated ON (MSON) [2]. The diagnosis for both conditions has been clinical [3; 4] and in the case of MS is based on finding evidence for dissemination in space (DIS) and dissemination in time (DIT) [5]. With the increasing recognition of the diagnostic sensitivity of MRI for DIS and DIT the need for clinical evidence has declined in the past two decades [6-8].

There are cases however, either in clinical or research setting, in which there is a need to establish the presence of a previous (subclinical) episode of MSON but can be challenging. Clinically recommended [9; 10], the art of using an ophthalmoscope in practice has declined with emergence of new diagnostic techniques such as optical coherence tomography (OCT). In recognition of this development it was proposed to use a 20% inter-eye percentage difference (IEPD) of the peripapillary retinal nerve fibre layer (pRNFL) as a diagnostic criterion for MSON [10]. There are several advantages to a dimensionless variable such as the IEPD, one of which is that it is less likely that differences between devices and segmentation algorithms will matter. Because of class III evidence only for recommended consensus of a 20% IEPD cut-off of the pRNFL on OCT this study aimed to define optimised cut-off ranges for the IEPD of the pRNFL and combined ganglion cell and inner plexiform layer (mGCIPL) for making a diagnosis of MSON.

METHODS

This study was approved by the Medical Ethical Committees of the VU University Medical Center in Amsterdam, the Netherlands and is in accordance with the 1964 Declaration of Helsinki. Written informed consent was obtained from all subjects before study inclusion.

Study design and patient population

Patients with MS [8] and healthy control subjects (HC) were enrolled from the VU University Medical Center Amsterdam (VUmc).

All subjects underwent clinical and OCT assessments at baseline and were retested after 24 months. The assessment of MSON was based on medical history and always preceded the OCT measurement, making sure that the physician who established the diagnosis of MSON did not have the OCT data available at the time of assessment. Disease duration was defined as the time from the first MS symptoms. The expanded disability status scale (EDSS) was obtained by a certified examiner. As potential swelling of the pRNFL during the acute stages of MSON may confound OCT measurements, patients were excluded if they had experienced symptomatic MSON three months prior to either OCT measurement (baseline or follow-up), according to a consensus protocol [10]. Patients were classified into those without MSON (nonMSON), unilateral MSON or bilateral MSON. or 'new MSON' if an episode occurred within the 2 year observation period

Optical Coherence Tomography

Spectral Domain OCT (SD-OCT, Spectralis, Heidelberg Engineering, Heidelberg, Germany) was performed in all subjects at baseline and follow-up as described previously [11]. Scans were excluded from the analyses if manual correction for algorithm segmentation failures was not possible or if they did not meet the remaining six quality control criteria (OSCAR-IB) [12]. This led to a rejection rate of 12.8%. at baseline and 17.8% at 2 year follow-up.

Statistical analyses

Data were assessed for normality using the Shapiro-Wilk test followed by graphical inspection in SAS (version 9.4). According to distribution non-parametric (Mann Whitney U-test) or parametric (independent sample T-test) tests were used for comparison of two groups or ANOVA and GLM followed by post-hoc analyses for comparison of more than two groups. For OCT data all analyses performed on eye-level (two separate eyes per subject) were done using generalized estimation equations (GEE), with an exchangeable correlation matrix and adjustments for intra-subject inter-eye correlations. Analyses for OCT and visual acuity (VA) data averaged from both eyes were done by GLM. Additional adjustments for confounding factors such as age, sex and disease type were performed as indicated. The IEPD was calculated for the pRNFL and mGCIPL for each patient. Receiver Operator Characteristics (ROC) curves were used to calculate the Area Under the Curve (AUC) to describe the level of diagnostic accuracy [13]. The diagnostic value was rated as 'no or low discriminatory power' for an AUC 0.5-0.7, as of 'moderate discriminatory power' for an AUC of 0.7-0.9 and of 'high discriminatory power' for an AUC >0.9 [14]. The ROC were plotted to determine graphically optimised IEPD cut-off values as the shortest distance from the top left corner to the ROC curve [15]. p-values of <0.05 were accepted as significant.

RESULTS

In total 296 subjects were assessed. Of these 26 had to be excluded because a diagnosis of previous MSON could not be established with certainty or because of visual symptoms due to Uhthoff's phenomenon or eye movement problems [10]. The overall prevalence of MSON was 48.8%. The highest prevalence was found in patients with a secondary progressive (SP)

disease course, followed by those with a relapsing remitting (RR) disease course. In patients with SP MS a previous episode of MSON was established in the right eye of 11/55 (20.0%), in the left eye of 9/55 (16.4%) and bilateral in 13/55 (23.6%). In patients with RR MS a previous episode of MSON was established in the right eye of 20/123 (16.3%), in the left eye of 21/123 (17.1%) and bilateral in 24/123 (19.5%). Of the patients suffering from primary progressive MS 26/29 (89.7%) had no history of MSON, but there was convincing evidence for MSON in 3 patients. The baseline data of the patients with and without MSON and HC are summarised in Table 1.

[Table 1 – around here]

The HC were younger compared to patients without MSON (p=0.0006), but there was no age difference between HC, uni- or bilateral MSON (p>0.05). Likewise there was a difference in disease duration between groups (p=0.023) which was longer in patients with bilateral MSON compared to those without MSON (p=0.008). The VA averaged from both eyes was worse in patients with bilateral MSON compared to patients with unilateral MSON (p=0.0006) or nonMSON (p=0.024). There was no significant difference between groups for the EDSS (Table 1).

Inter-eye percentage differences (IEPD)

The results of the ROC curves for the IEPD are summarised in Table 2. The AUCs and optimised cut-off values are presented for the IEPD of the pRNFL and mGCIPL. For each group comparison the optimised IEPD cut-off values were below 10%. The highest AUC

7

(0.94) was found for the mGCIPL for the comparison of patients with bilateral MSON and HC with a cut-off as small as 5% inter-eye difference.

The IEPD of either the pRNFL or mGCIPL was of no diagnostic value for comparing unilateral with bilateral MSON. Whilst the IEPD of the mGCIPL did distinguish between HC and nonMSON (AUC 0.77) the IEPD of the pRNFL did not (0.56).

[Table 2 – around here]

For further illustration the raw data for the IEPD of the pRNFL and mGCIPL are shown in Figure 1. The IEPD was highly significantly (p<0.0001) larger in patients with either unilateral or bilateral MSON compared to patients with nonMSON or HC. The IEPD for the mGCIPL but not the IEPD for the pRNFL was significantly (p<0.01) higher in patients with nonMSON compared to HC.

[Figure 1 – around here]

Diagnostic sensitivity and specificity of IEPD

The diagnostic sensitivity and specificity levels of the IEPD were calculated for those comparisons where the AUC indicated a 'high discriminative power' (AUC >0.9) [15]. This was the IEPD for the mGCIPL comparing patients with unilateral and bilateral MSON to HC (see Table 2). The ROC optimised cut-off value was compared to the consensus cut-off value of 20% [10].

Using the consensus cut-off value of 20% the diagnostic sensitivity of the IEPD for the mGCIPL for unilateral MSON compared to HC was 34% with a specificity of 100%. Using the optimised cut-off value of 6% (see Table 2) this changed to a sensitivity level of 70% and specificity level of 97%.

Using the consensus cut-off value of 20% the diagnostic sensitivity of the IEPD for the mGCIPL for bilateral MSON compared to HC was 36% with a specificity of 100%. Using the optimised cut-off value of 5% (see Table 2) this changed to a sensitivity level of 86% and specificity level of 97%.

Validation of IEPD cut-off values

For validation, sensitivity and specificity levels were re-assessed at follow-up. Two patients experienced an episode of MSON between baseline and follow-up and where therefore excluded from further analysis. The clinical data at follow-up is summarised in Table 3. The drop out of patients was considerable. Overall patients were clinically relatively stable and OCT thickness changes within the known physiological variation [16].

[Table 3 around here]

The diagnostic sensitivity and specificity values for a 6% cut-off for the IEPD of the mGCIPL for comparing unilateral MSON with HC were 71% and 97%, respectively.

The diagnostic sensitivity and specificity values for a 5% cut-off for the IEPD of the mGCIPL for comparing bilateral MSON with HC were 83% and 97%, respectively.

The raw data for the IEPD of the pRNFL and mGCIPL at follow-up are shown in Figure 2. Interestingly, the two patients who experienced a confirmed episode of MSON between baseline and follow-up had a very high IEPD for both retinal layers. But numbers of new episodes were too small to allow for a meaningful statistical analysis.

[Figure 2 – around here]

DISCUSSION

This study showed that the IEPD of the pRNFL and the IEPD of the mGCIPL were of diagnostic value for MSON. Detailed ROC analyses suggest that the published consensus on a 20% inter-eye difference for the pRNFL [10] was overly conservative. While inter-eye analysis of OCT data is not new (klistorner, Jette), previous studies have looked at the absolute difference between the two eyes and the 20% IEPD consensus for the pRNFL is based on expert opinion only. There are yet no published data on the inter-eye percentage difference for the mGCIPL. The present study suggests that a 5% IEPD of the mGCIPL is of excellent diagnostic sensitivity and specificity for bilateral MSON in the setting investigated.

These findings are relevant because one advantage of the IEPD is that the dimensionless value may help to overcome limitations to pool data from different devices and segmentation algorithms. As many new OCT devices enter the market adding to heterogeneity between center equipment [17], there is a need to start thinking about strategies to facilitate future data analyses. The IEPD could potentially be a valuable secondary outcome measure in ON treatment trials, pending independent validation.

There are statistical advantages to the IEPD, too. As correctly pointed out the statistical analyses of OCT data needs to consider inter-eye interactions [18]. Consequently, a concerted effort developed OCT reporting guidelines outlining the use of GEE of binocular OCT data [19]. The robustness of the model is a strength, but does not lend itself for modelling of associations with single outcome variables [11]. Clinical outcome measures such as the EDSS, Multiple Sclerosis Functional Composite, cognition, relapse rate and treatment response are all relevant [20]. Statistical evaluation of these data makes use of multivariate Cox models, Kaplan-Meyer curves, the many flavours of GLM, forms of multimodal modelling, principal component analysis and many more [21, 22] to which the IEPD may have advantages over device and eye dependent absolute values of OCT data.

Finally, the IEPD may overcome present exclusion criteria due to high hypermetropia or myopia as well as ethnic retinal layer thickness differences and thickness variations due to physiological variability. The observation that the IEPD for the pRNFL and mGCIPL were very high in two patients with a new episode of confirmed MSON is encouraging, but will require further validation prospectively.

A likely limitation of the IEPD will be a low diagnostic specificity for separating MSON from other forms of ON such as isolated optic neuritis, chronic relapsing isolated optic neuropathy (CRION) and neuromyelitis optica ON. None of these patient groups were investigated in the present study, which is a limitation of the study.

Another limitation of the present study is that the IEPD has not been tested in parallel on different devices. With the advent of many new OCT devices [17], testing this will be resource demanding. The IEPD is also unlikely to overcome pertinent quality control issues [12, 19].

There were two unexpected findings in this study. First, the IEPD of the mGCIPL was of high diagnostic accuracy for bilateral MSON as well as for the expected unilateral MSON. Second, the IEPD of the mGCIPL was superior to the IEPD of the pRNFL. There is a likely anatomical explanation for both results. The large variability of the optic disc appearance between patients will cause a larger degree of variability of pRNFL data. In contrast, the more homogeneous appearance of the macula lutea will make the mGCIPL data more consistent between eyes and patients. It is well known that MSON can cause any type of visual field defect, the anatomical basis for which is loss of retinal axons and their ganglion cells [2,10]. With bilateral MSON the visual field defect frequently is asymmetric between eyes. Therefore the diffuse, asymmetric damage caused is likely to affect the IEPD of the mGCIPL independent to whether the damage was due to unilateral or bilateral MSON.

CONCLUSION

This study suggest that the IEPD of the pRNFL and the IEPD of the mGCIPL may prove a useful variable for interpretation of OCT data. Advantages of the IEPD are simplicity for clinical practise, the possibility to become a cross OCT device measure, the suitability for many statistical models on established MS outcome measures and a potentially robust and useful secondary outcome measure for treatment trials. There is a need for further validation of the IEPD.

AUTHOR CONTRIBUTIONS

DC performed data collection, preformed automated and manual layer segmentation, statistical analyses and co-wrote the first draft of the manuscript. LB contributed to designing the study concept, performed data collection, revision and QC assessment of OCT scans, automated and manual layer segmentation and statistical analyses. BU contributed to designing the study concept and revised the manuscript. AP had the idea for the study, contributed to designing the study concept, revision and QC assessment of OCT scans, automated and manual layer segmentation, statistical analyses, wrote the first draft and revised the manuscript. All authors gave final approval of the version submitted.

REFERENCES

[1] Parinaud H. Troubles oculaires de la sclérose en plaques. *Progr. méd. (Paris)* 1884; 12: 641-650.

[2] Toosy AT, Mason DF, Miller DH. Optic neuritis. *Lancet Neurology* 2014; 13: 83-99.

[3] McDonald WI. Acute optic neuritis. *British Journal of Hospital Medicine* 1977; 18: 42-48.

[4] McDonald WI, Halliday AM. Diagnosis and classification of multiple sclerosis. *British Medical Bulletin* 1977; 33: 4-9.

[5] Filippi M, Rocca MA, Ciccarelli O, et al. MRI criteria for the diagnosis of multiple sclerosis: MAGNIMS consensus guidelines. *Lancet Neurology* 2016; 15: 292-303.

[6] Poser C, Paty DW, Scheinberg L, et al. New Diagnostic Criteria for Multiple Sclerosis: Guidelines for Research Protocols. *Annals of Neurology* 1983; 13: 227-231.

[7] McDonald WI, Compston A, Edan G, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Annals of Neurology* Neurol 2001; 50: 121-127.

[8] Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Annals of Neurology* 2011; 69: 292-302.

[9] McDonald WI. Diagnosis and management of optic neuritis. *Transections of the Ophthalmological Society of New Zealand* 1976; 28: 11-17.

[10] Petzold A, Wattjes MP, Costello F, et al. The investigation of acute optic neuritis: a review and proposed protocol. *Nature Reviews. Neurology* 2014; 10: 447-458.

[11] Coric D, Balk LJ, Verrijp M, et al. Cognitive impairment in patients with multiple sclerosis is associated with atrophy of the inner retinal layers. *Multiple Sclerosis Journal* 2017; 1352458517694090.

[12] Tewarie P, Balk LJ, Costello F, et al. The OSCAR-IB Consensus Criteria for Retinal OCT Quality Assessment. *PLoS One* 2012; 7: e34823.

[13] Faraggi D, Reiser B. Estimation of the area under the ROC curve. *Statistics in Medicine* 2002; 21: 3093-3106.

[14] Grzybowski M, Younger JG. Statistical methodology: III. Receiver operating characteristic (ROC) curves. *Academic Emergency Medicine: Official Journal of the Society for Academic Emergency Medicine* 1997; 4: 818-826.

[15] Fan J, Upadhye S, Worster A. Understanding receiver operating characteristic (ROC) curves. *Canadian Journal of Emergency Medicine* 2006; 8: 19-20.

[16] Balk LJ, Oberwahrenbrock T, Uitdehaag BMJ, Petzold A. Physiological variation of retinal layer thickness is not caused by hydration: A randomised trial. *Journal of the Neurological Sciences* 2014; 344: 88-93.

[17] Costello FE. Optical coherence tomography technologies: which machine do you want to own*? Journal of Neuro-ophthalmology: the Official Journal of the North American Neuro-ophthalmology Society* 2014; 34 Suppl: S3-S9.

[18] Balcer LJ, Galetta SL. OCT and NMO: are there methods to our madness? *Journal of Neuro-ophthalmology: the Official Journal of the North American Neuro-ophthalmology Society* 2013; 33: 209-212.

[19] Cruz-Herranz A, Balk LJ, Oberwahrenbrock T, et al. The APOSTEL recommendations for reporting quantitative optical coherence tomography studies. *Neurology* 2016; 86: 2303-2309.

[20] van Munster CEP, Uitdehaag BMJ. Outcome Measures in Clinical Trials for Multiple Sclerosis. *CNS Drugs* 2017; 31: 217-236.

[21] Petkau J. Statistical methods for evaluating multiple sclerosis therapies. *Seminars in Neurology* 1998; 18: 351-375.

[22] Sormani MP, Gasperini C, Romeo M, et al. Assessing response to interferon-β in a multicenter dataset of patients with MS. *Neurology* 2016; 87: 134-140.

Fig 1. *Inter-eye percentage difference (IEPD) for the peripapillary retinal nerve fibre layer (pRNFL) and macular ganglion cell and inner plexiform layer (mGCIPL) are shown per subject and per group at baseline. Levels of significance are indicated as p<0.0001 (***) and p<0.01 (*). HC = healthy controls*

Fig 2. *Inter-eye percentage difference (IEPD) for the peripapillary retinal nerve fibre layer (pRNFL) and macular ganglion cell and inner plexiform layer (mGCIPL) are shown per subject and per group at follow-up. Levels of significance are indicated as p<0.0001 (***) and p<0.001 (**), p<0.01 (*). HC = healthy controls*