Published in final edited form as:

J Glaucoma. 2008 ; 17(3): 175–182. doi:10.1097/IJG.0b013e31815678ac.

Effect of Recording Duration on the Diagnostic Performance of Multifocal Visual-evoked Potentials in High-risk Ocular Hypertension and Early Glaucoma

Brad Fortune, OD, PhD^{*}, Xian Zhang, PhD[†], Donald C. Hood, PhD[†], Shaban Demirel, OD, PhD^{*}, Emily Patterson, MD^{*}, Annisa Jamil, MD^{*}, Steven L. Mansberger, MD, MPH^{*}, George A. Cioffi, MD^{*}, and Chris A. Johnson, PhD^{*}

^{*} Discoveries in Sight, Devers Eye Institute, Legacy Health System, Portland, OR

[†] Department of Psychology, Columbia University, New York, NY

Abstract

Purpose—To evaluate the effect on diagnostic performance of reducing multifocal visual-evoked potential (mfVEP) recording duration from 16 to 8 minutes per eye.

Methods—Both eyes of 185 individuals with high-risk ocular hypertension or early glaucoma were studied. Two 8-minute mfVEP recordings were obtained for each eye in an ABBA order using VERIS. The first recording for each eye was compared against single run (1-Run) mfVEP normative data; the average of both recordings for each eye was compared against 2-Run normative data. Visual fields (VFs) were obtained by standard automated perimetry (SAP) within 22.3±27.0 days of the mfVEP. Stereo disc photographs and Heidelberg Retina Tomograph images were obtained together, within 24.8±50.4 days of the mfVEP and 33.1±62.9 days of SAP. Masked experts graded disc photographs as either glaucomatous optic neuropathy or normal. The overall Moorfields Regression Analysis result from the Heidelberg Retina Tomograph was used as a separate diagnostic classification. Thus, 4 diagnostic standards were applied in total, 2 based on optic disc structure alone and 2 others based on disc structure and SAP.

Results—Agreement between the 1-Run and 2-Run mfVEP was 90%. Diagnostic performance of the 1-Run mfVEP was similar to that of the 2-Run mfVEP for all 4 diagnostic standards. Sensitivity was slightly higher for the 2-Run mfVEP, whereas specificity was slightly higher for the 1-Run mfVEP.

Conclusions—If higher sensitivity is sought, the 2-Run mfVEP will provide better discrimination between groups of eyes with relatively high signal-to-noise ratio (eg, early glaucoma or high-risk suspects). But if higher specificity is a more important goal, the 1-Run mfVEP provides adequate sensitivity and requires only half the test time. Considered alongside prior studies, the present results suggest that the 1-Run mfVEP is an efficient way to confirm (or refute) the extent of VF loss in patients with moderate or advanced glaucoma, particularly in those with unreliable VFs, including malingering or other “functional” forms of VF loss.

Keywords

electrophysiology; glaucoma; optic disc; multifocal visual-evoked potential; perimetry

Accumulated evidence indicates that the multifocal visual-evoked potential (mfVEP)^{1,2} is an effective technique for objective assessment of the visual field (VF) in glaucoma and other optic nerve diseases.^{2–18} Although most of these studies have demonstrated close agreement between mfVEP and standard automated perimetry (SAP), some have shown that mfVEP abnormalities appear in regions (eg, hemifields) of the VF that have normal sensitivity on SAP,^{2,11,13,19} and in glaucoma suspect eyes with completely normal SAP VFs.^{4,5,9,14,18,19} Longitudinal follow-up should reveal whether such mfVEP defects represent greater sensitivity of the mfVEP to detect early functional loss, or alternatively, a higher false alarm rate. Meanwhile, cross-sectional studies have found that the diagnostic performance of mfVEP and SAP are similar when an independent diagnostic standard is applied (eg, one based on optic disc structure).^{14,18}

The mfVEP test duration, however, can be considerably longer than SAP, particularly when the Swedish Interactive Thresholding Algorithm (SITA) is used as the SAP test strategy²⁰ and two 8-minute mfVEP recordings are acquired for each eye, as in many previous studies (eg, Ref. ¹⁸). Therefore, the purpose of the present study was to evaluate the potential effect of reducing the recording duration, from 16 to 8 minutes per eye, on the diagnostic performance of the mfVEP in eyes with high-risk ocular hypertension or early glaucoma.

METHODS

Subjects

One hundred eighty-five individuals with high-risk ocular hypertension or early glaucoma participated in this study, including 107 females and 78 males, ranging in age from 37 to 87 years (mean±SD: 60.9±11.0 y). This same cohort was the subject of another recent study.¹⁸ All subjects were fully informed with regard to the potential risks and benefits of the study and then provided voluntary written consent to participate. All procedures followed the tenets of the Declaration of Helsinki, were in accordance with Health Insurance Portability and Accountability Act regulations, and were preapproved by the Legacy Health System Institutional Review Board for the protection of human subjects.

Participants were recruited prospectively from the Devers Eye Institute, or other ophthalmic practices in the Portland, OR metropolitan area. All subjects had been tested on at least one previous occasion with full-threshold or SITA-standard white-on-white SAP. At the time of recruitment, all subjects were considered to have either high-risk ocular hypertension or early glaucoma. Thus, all subjects had a history of untreated intraocular pressure ≥22mm Hg in both eyes and at least one of the following additional risk factors: vertical cup-to-disc ratio ≥0.6 in at least 1 eye and/or an interocular cup-to-disc ratio asymmetry ≥0.2 between the 2 eyes; positive family history of glaucoma; personal history of migraine, Raynaud syndrome, or vasospasm; African American ancestry; or age >70 years. To be included, all subjects also met the following criteria for both eyes: best corrected visual acuity ≥20/40 and spectacle refraction < ±5.00-D sphere and < ±2.00-D cylinder. Subjects were excluded if they had any other previous or current ocular or neurologic disease, previous ocular surgery (except uncomplicated cataract surgery), or diabetes requiring medication. Note that SAP VF status (normal or abnormal) was not a criterion for study entry, as long as the mean deviation (MD) was better than – 6 dB at the time of recruitment.

SAP VFs were obtained within 22.3 (±27.0) days of the mfVEP. Stereo disc photographs and HRT images were obtained during the same visit, which was within 24.8±50.4 days of the mfVEP and 33.1±62.9 days of the SAP VF. (Note the SDs are large because of a few outliers; the disc images were obtained within 3mo of the functional tests for 96% of subjects.)

SAP

SAP VF testing was performed using the Humphrey Field Analyzer II (model 750, Carl Zeiss Meditec, Dublin, CA). The 24-2 VF test pattern was used for 119 (64%) of the subjects. The other 66 (36%) subjects were tested using the 30-2 pattern, with only the points corresponding to those in the 24-2 being analyzed. Nearly all tests used the SITA-standard threshold strategy (348 of 370 eyes, 94%); the other 22 eyes were tested using the Full Threshold strategy. SAP VF data were imported into Excel (Microsoft, Redmond, WA) for further analysis. The normative database for these analyses came from 348 eyes of 348 normal control subjects.²¹ As previously described,²¹ the data for each VF test were transformed into 45-year-old equivalent sensitivities before calculation of VF indices [MD, PSD, and Glaucoma Hemifield Test (GHT)], and then compared with the results for normal 5%, 2%, 1%, and 0.5% probability levels for each of the global parameters, local threshold values, and Total Deviation (TD) values.

SAP VFs were considered abnormal if the pattern standard deviation (PSD) value was beyond the normal 95% limit ($P \leq 0.05$), or if the GHT was “outside normal limits,” or if there was a 3-point cluster of abnormal locations ($P \leq 0.05$) on the TD plot contained within either the upper or lower hemifield, with at least one of these points being abnormal at the $P \leq 0.01$ level. The final for classifying SAP is similar to the Hodapp-Parrish-Anderson criteria.²²

Stereo Optic Disc Photography and Grading

Optic disc photographs were obtained in all patients using a simultaneous stereoscopic camera (3-Dx, NIDEK Co, Ltd, Gamagori, Japan) after maximal pupil dilation. Two fellowship-trained glaucoma specialists (E.P. and A.J.), masked to all other patient information, independently graded each photograph (viewed with a Stereo Viewer II, Asahi, Pentax, Tokyo, Japan) as either “normal” or “glaucomatous optic neuropathy” (GON) on the basis of the following characteristics: adequate clarity and stereopsis; neuroretinal rim thinning (generalized or localized); excavation; retinal nerve fiber layer defect; violation of “inferior-superior-nasal-temporal (ISNT)” rule; and cup-disc ratio. The presence and location of disc hemorrhages were noted, but in isolation, did not necessarily warrant a grade of GON. A third masked expert (G.A.C.) adjudicated disagreements between these 2 graders.

Confocal Scanning Laser Ophthalmoscopy

Confocal scanning laser ophthalmoscopy images were obtained using the Heidelberg Retina Tomograph (HRT, Heidelberg Engineering, Dossenheim, Germany, versions 2.01/3.04). Six 10×10 -degree scans centered on the optic disc were acquired and the 3 judged to have the best quality were combined to create a mean topography for each eye. Experienced technicians outlined the margin of the optic disc while viewing stereo disc photographs. The overall Moorfields Regression Analysis (MRA) classification²³ was recorded for each eye. All eyes classified as “borderline” by the MRA were assigned to the normal category (ie, “within normal limits”). This maintains high specificity.²⁴ For example, in a separate group of 100 control subjects,²⁵ only 2 of 200 normal eyes (1%) were “outside normal limits” whereas 18 others were classified as “borderline” by the MRA.

Multifocal VEP

Stimulus—Multifocal VEPs were obtained as previously described, 18 using one of the standard stimulus options available within the VERIS software package (“Dart Board 60 With Pattern,” VERIS 4, Electro-Diagnostic Imaging, San Mateo, CA). The stimulus had a total diameter of 44.5 degrees and consisted of 60 individual sectors. Each sector contained

16 checks, 8 white (200 cd/m²) and 8 black (<1 cd/m²) and thus had a Michelson contrast of ~99%. In all cases, the display was viewed through natural pupils with optimal refractive correction in place.

During a recording, the contrast polarity of each stimulus sector was temporally modulated (pattern reversals) according to a pseudo-random m-sequence. The m-sequence chosen for this study had $2^{15}-1$ steps in total and thus required ~8 minutes to complete. Each 8-minute recording was divided into 16 segments of equal duration and subjects were able to rest between segments as necessary. Segments contaminated by eye movements, loss of fixation, and/or noise are discarded and rerecorded. Each subject completed 2 recordings per eye in an ABBA-order with the first eye randomly chosen.

Recording Procedures—Three channels were recorded using gold disc electrodes (Grass Model F-E5GH, Astro-Med, Inc, West Warwick, RI). For the vertical midline channel, electrodes were placed 4-cm above the inion (active) and at the inion (reference). For the left and right “oblique” channels, the active electrodes were placed 1-cm above the inion and 4 cm to the left or right, respectively, of the vertical midline; each of these active sites was referenced to the electrode at the inion. The left earlobe served as ground for all 3 channels. In addition to the 3 channels recorded, data for 3 other channels were derived from the differences between various electrode pairs, as previously described.^{2,26}

In preparation for recording, the skin at each electrode site was scrubbed with Nuprep (D.O. Weaver & Co, Aurora, CO) on a cotton-tipped wooden swab. Electrodes were fixed in position with EC2 conductive cream (Astro-Med, Inc, Warwick, RI) and secured with Coban wrap (3 M, St Paul, MN). Electrode impedance was maintained below 5 kΩ in all cases and was usually below 2 kΩ. Signals were amplified by 10^5 (Grass Model 12), band-pass filtered from 3 to 100 Hz (1/2 amplitude) and sampled at 1200 Hz (ie, in 0.83-ms bins).

Signal-to-Noise Ratio—For the “2-Run” condition, both recordings for each eye were averaged after the mfVEP responses were exported from VERIS; only the first recording for each eye was exported for the “1-Run” condition. No spatial smoothing or artifact rejection was applied before exporting the data. All other mfVEP data analyses were performed using programs that were written in commercial software (Matlab; MathWorks, Inc, Natick, MA) as follows. First, mfVEP records from each of the 6 channels were low-pass filtered using a Fourier transform technique with a sharp cutoff at 35 Hz. Then, the signal-to-noise ratio (SNR) for each local response was derived as previously described.^{2,27} The root-mean-square (RMS) amplitude of a signal window (45 to 150 ms) from each local response was divided by the RMS amplitude for a noise window (325 to 430 ms); the latter was taken as the *average* value of all 60 records for a given eye.^{2,27}

The Best Array—All analyses were performed on the “best” of the responses from the 6 channels available for each eye.^{2,26} The best arrays were composed of the 60 best SNR responses, but were derived differently for the monocular and interocular tests. In the monocular test, for each eye at each location, the response with the largest SNR among the 6 channels was selected for inclusion in the best array. For the interocular test, the response with the largest SNR was selected from the 12 responses (2 eyes, 6 channels each eye) at each location. The corresponding response from the other eye from that same channel completed the pair of responses at that location in the interocular best array.

Abnormalities Defined by Cluster Criteria—The monocular SNR for each mfVEP response (ie, the SNR at each location in each eye of each subject) was converted to a z-score relative to a previously published normal distribution derived from 100 control subjects with similar demographics to the current study population (100 individuals, age

range: 22 to 92 y, average $\pm$ SD: 49.0 $\pm$ 13.6 y).²⁵ Similarly, the interocular ratio (of root-mean-squared amplitude) at each location was converted to a z-score and assigned a probability value on the basis of the corresponding local normal distribution. The z-scores for 2-Run mfVEPs were based on 2-Run normative values obtained from controls; accordingly, the 1-Run z-scores were based on 1-Run normative data from controls (ie, the first run for each eye from the ABBA group of 2).

The mfVEP for each individual (ie, the monocular mfVEP for the right and left eyes, and the interocular ratio) was determined to be either normal or abnormal based on various “cluster” criteria.^{18,25,28} The definition of a cluster was 2 or 3 contiguous locations within a single hemifield. Each of the various cluster criteria was specified as a unique combination of cluster size and probability level (z-score) of the individual points within the cluster.

Comparison of Diagnostic Performance: 1-Run Versus 2-Run mfVEP

The diagnostic performance of the 1-Run mfVEP was compared with that of the 2-Run mfVEP using sensitivity and specificity results for a range of mfVEP cluster criteria. Sensitivity of the mfVEP was defined as the percentage of eyes with a positive diagnosis that also had an abnormal mfVEP. Specificity was defined as the percentage of eyes with a negative diagnosis (ie, normal) that also had a normal mfVEP. These comparisons were performed using 4 different diagnostic standards. The first 2 diagnostic standards were based exclusively on optic disc structure: (1) the masked expert grade of the stereo disc photo; and (2) the HRT MRA classification. Thus, for these 2 conditions, performance of the mfVEP was judged for detection of GON only. The other 2 diagnostic standards included SAP in their classification: (3) stereo disc photo grade and SAP; and (4) HRT MRA and SAP. Thus, these latter 2 diagnostic standards enabled comparison of the 1-Run mfVEP with the 2-Run mfVEP for detection of manifest glaucoma (ie, glaucomatous optic disc and SAP VF). The 95% confidence interval (CI) for all proportions (sensitivity and specificity estimates) was calculated to provide an estimate of precision based on sample size.²⁹ Note that some of the results for the 2-Run mfVEP have been published previously,¹⁸ and are reiterated here for ease of comparison with the 1-Run mfVEP results.

RESULTS

The average SAP MD was +0.3 $\pm$ 2.1 dB (range: +3.9 to – 10.1 dB) and the average PSD was 2.3 $\pm$ 1.9 dB (range: 1.0 to 16.1 dB). There were 89 eyes (24%) with an abnormal SAP (as defined above). There were 185 eyes (50%) classified as GON by masked expert grades of the stereo disc photo. There were 93 eyes (25%) classified as “outside normal limits” by the HRT MRA. There were 59 eyes (16%) in which both the stereo disc photo and the SAP VF were classified as abnormal. There were 44 eyes (12%) in which both the HRT MRA and the SAP VF were classified as abnormal.

Figure 1 presents an example of the results for 1 subject. The top row (Fig. 1A) shows the SAP 24-2 grayscale plots for the right and left eye. The second row (Fig. 1B) shows the SAP TD plots for the right and left eye. VF loss in the left eye of this patient was among the most severe in this study with an MD of – 7.6 dB. The 2-Run mfVEP results are shown in the third row (Fig. 1C) and the 1-Run mfVEP results are shown in the bottom row (Fig. 1D). The left panel shows the mfVEP response array with right eye responses in blue and left eye responses in red. The inset shows 1 pair of responses at higher magnification; note the 1-Run responses are slightly noisier. The second panel shows the monocular probability plot for the right eye mfVEP and the third panel shows the monocular probability plot for the left eye mfVEP. The right-most panel shows the interocular probability plot. In all probability plots, a red colored square indicates an abnormal left eye response and a blue colored square

indicates an abnormal right eye response. A saturated color indicates $P < 0.01$ and a desaturated (pastel) color indicates $P < 0.05$.

In this case, there is no right eye abnormality (cluster) for either 2-Run or 1-Run mfVEP: there are no abnormal points on the monocular plot and only 1 abnormal point for the right eye on either of the interocular plots (inferotemporal on the 2-Run and paracentral on the 1-Run interocular plots), which does not meet any of the cluster criteria. Interestingly, the enlarged blind spot apparent on the right eye SAP does not appear as a mfVEP defect, although the pair of responses shown in the insets shows a relatively large interocular asymmetry, evidently not large enough to trigger an interocular abnormality though. The position of the blind spot overlaps 4 stimulus sectors in the mfVEP, so this SAP defect becomes “spread” over several mfVEP sectors and loses significance.

In contrast, the left eye shows abnormalities on both monocular and interocular plots for both 2-Run and 1-Run conditions. There are more abnormal points (a larger and deeper cluster) in the 2-Run plots, and better agreement with SAP, as compared with the 1-Run plots. However, the 1-Run mfVEP diagnosis was still abnormal for the left eye.

Table 1 lists the number of eyes with an abnormal mfVEP for each of the cluster criteria evaluated (top row, 2-Run mfVEP; second row, 1-Run mfVEP). Approximately, 30% fewer eyes were classified as abnormal by the 1-Run mfVEP as compared with the 2-Run mfVEP. This is most likely due to the decrease in the lower limit of the normal SNR range for the 1-Run mfVEP,²⁵ creating a “floor effect” (discussed later). The third row of Table 1, however, shows that the classification agreement between 1-Run and 2-Run mfVEPs was close to 90%. This indicates that relatively few of the eyes classified as normal by the 2-Run mfVEP switched to an abnormal classification when considering only the 1-Run mfVEP; confirming that specificity remains relatively similar for the 1-Run mfVEP, as initially suggested by prior studies in control eyes.²⁸ The bottom 2 rows of Table 1 shows that the agreement between mfVEP and SAP are essentially identical (close to 80% for all cluster criteria) for 1-Run and 2-Run mfVEPs.

Figure 2 plots the sensitivity of the mfVEP to detect GON versus false alarm rate (1-specificity), when the masked expert grade of the stereo disc photo was used as the diagnostic standard. Results for the 1-Run mfVEP (open circles) are similar to the results for the 2-Run mfVEP (filled diamonds). Each data point shows the result for a particular cluster criterion. The diagnostic performance of the 1-Run mfVEP is similar to that of the 2-Run mfVEP for discrimination of eyes with GON. That is, the 2 families of cluster criteria plot along the same detectability (d') locus. The 1-Run results (open circles) tend to have slightly lower sensitivity and false alarm rates as compared with the 2-Run outcome for a given cluster definition. For example, using the 3-point cluster “244” ($P < 0.01, 0.05, 0.05$), sensitivity to detect GON (per disc photo grade) was slightly lower for the 1-Run mfVEP at 26% (20% to 32%, 95% CI), as compared with 33% (26% to 40%, 95% CI) for the 2-Run mfVEP. The specificity of the 1-Run mfVEP (85%, 80% to 90%, 95% CI) was slightly higher than the 2-Run mfVEP (79%, 73% to 85%, 95% CI) for this cluster type.

Similarly, Figure 3 suggests that the 1-Run and 2-Run mfVEP perform equally well when the definition of GON is based on the HRT MRA classification of “outside normal limits.” Again, for any given cluster type, the sensitivity to detect GON is slightly higher for the 2-Run mfVEP, whereas the specificity is slightly lower. For the 3-point cluster 244 ($P < 0.01, 0.05, 0.05$), sensitivity to detect GON (per HRT MRA) was 39% (29% to 49%, 95% CI) for the 1-Run mfVEP and 50% (40% to 61%, 95% CI) for the 2-Run mfVEP; specificity was 86% (82% to 90%, 95% CI) for the 1-Run mfVEP and 81% (76% to 86%, 95% CI) the 2-Run mfVEP.

Figure 4 compares diagnostic performance of the 1-Run versus the 2-Run mfVEP for detection of early manifest glaucoma. In this condition, early glaucoma was defined as both a glaucomatous optic disc (by expert grade of stereo photo) and the presence of a SAP VF defect. There were 59 eyes in which the SAP VF and the stereo optic disc photo were both classified as glaucomatous, and 153 eyes in which the SAP VF and the stereo optic disc photo were both classified as normal. Figure 4 shows that the diagnostic performance of the 1-Run and the 2-Run mfVEP are similar when the standard is based on a common “clinical” definition. For the 3-point cluster 244 ($P < 0.01, 0.05, 0.05$), sensitivity to detect early glaucoma (SAP and disc photo positive) was 52% (39% to 64%, 95% CI) for the 1-Run mfVEP and 65% (53% to 77%, 95% CI) for the 2-Run mfVEP; specificity was 90% (85% to 94%, 95% CI) for the 1-Run mfVEP and 84% (78% to 90%, 95% CI) for the 2-Run mfVEP.

Figure 5 also compares the diagnostic performance of the 1-Run versus the 2-Run mfVEP for detection of early manifest glaucoma; however, in this circumstance glaucoma was defined as the presence of both a SAP VF defect and an HRT MRA classification “outside normal limits.” There were 44 eyes in which the SAP VF and the HRT MRA were both classified as glaucomatous, and 228 eyes in which the SAP VF and the HRT MRA were both classified as normal. Figure 5 shows again, that diagnostic performance of the 1-Run and the 2-Run mfVEP are similar. For the 3-point cluster 244 ($P < 0.01, 0.05, 0.05$), sensitivity to detect early glaucoma (SAP and HRT MRA positive) was 60% (46% to 75%, 95% CI) for the 1-Run mfVEP and 76% (63% to 89%, 95% CI) for the 2-Run mfVEP; specificity was 89% (85% to 93%, 95% CI) for the 1-Run mfVEP and 85% (81% to 90%, 95% CI) for the 2-Run mfVEP.

DISCUSSION

Before considering the extent to which the results suggest that the 1-Run mfVEP provides an adequate alternative to the 2-Run mfVEP, there are 2 caveats that deserve mention. First, we refer here only to glaucoma. For example, the cluster definitions used to define mfVEP abnormalities were all limited to contiguity within a hemifield, which may not be appropriate for other types of functional defects observed in diseases such as optic neuritis. Second, the current comparison was based on patients with high-risk ocular hypertension or early glaucoma, so the relative similarity of the 1-Run versus the 2-Run performance may not necessarily apply to other stages of glaucoma. For example, the sensitivity of the 1-Run and 2-Run mfVEP tests should converge toward 100% as the stage of disease advances (because severe loss of function is easier to detect), thus providing an advantage to the 1-Run mfVEP (slightly higher specificity and shorter test duration).

However, given these caveats, the major findings of this study support the conclusion that the 1-Run mfVEP can be used instead of the 2-Run mfVEP. The key results were the following. First, agreement between 1-Run and 2-Run mfVEP classification was high, close to 90% for nearly all cluster types (definitions of abnormality). Second, agreement between the SAP and mfVEP classifications was close to ~80%, regardless of whether the 1-Run or the 2-Run mfVEP outcome was used. Third, the diagnostic performance of the 1-Run mfVEP was nearly indistinguishable from that of the 2-Run mfVEP from a statistical standpoint, in eyes with high-risk ocular hypertension or early glaucoma, regardless of whether the diagnostic standard was exclusively based on optic disc structure (GON vs. normal disc structure) or based on a more complete “clinical” definition, that is, the presence of both GON and a SAP VF defect.

The fourth fundamental result was that the sensitivity of the 2-Run mfVEP was slightly greater than the 1-Run mfVEP for all 4 classification tasks: detection of GON per disc photo grade standard, detection of GON per HRT MRA standard, detection of early manifest

glaucoma per disc photo plus SAP standard, and detection of early manifest glaucoma per HRT MRA plus SAP standard. This result is consistent with the fact that the 2-Run mfVEP classified a higher number of eyes as abnormal for any particular cluster criterion (Table 1). The greater number of abnormalities determined by the 2-Run mfVEP is attributable to the improvement in SNR, especially to the increase in the lower limit of the normal SNR range at each location, a predictable consequence of increased signal averaging.²⁵

However, some of the increase in the number of abnormalities also contributed to a slightly higher false alarm rate for the 2-Run mfVEP. In general though, there was slightly greater enhancement of sensitivity, as compared with the trade-off in specificity for the 2-Run mfVEP. The apparent slight improvement for specificity of 1-Run over 2-Run mfVEP is partly explained by a floor effect. In particular, the lower limit of the normal range for SNR decreases due to fewer signals averaged in the case of the 1-Run mfVEP and this limit approaches noise in certain locations (eg, upper hemifield). Consequently, a patient's response with a poor SNR is less likely to be flagged as an abnormality in certain locations where the lower limit of the normal range is closer to 1.0 (noise level).

Therefore, the results of this study suggest that if a higher sensitivity is sought, the 2-Run mfVEP will provide better discrimination between groups of eyes with relatively high SNR (eg, early glaucoma or high-risk suspects). But if higher specificity is a more important goal, then the 1-Run mfVEP provides adequate sensitivity and requires only half the test time. Considered alongside prior studies, the present results suggest that the 1-Run mfVEP should be useful to confirm (or refute) the extent of VF loss in patients with moderately advanced or advanced glaucoma, particularly those patients whose SAP VF may have questionable reliability. This can also be extended to the more general clinical task of ruling out malingering or other "functional" forms of VF loss: the 1-Run mfVEP test provides relatively higher specificity and a shorter test duration for achieving this goal.

Acknowledgments

Funding/Support: M.J. Murdock Charitable Trust, Vancouver, WA; Legacy Good Samaritan Foundation, Portland, OR; NIH R01-EY03424 (C.A.J.) and R01-EY02115 (D.C.H.).

References

1. Baseler HA, Sutter EE, Klein SA, et al. The topography of visual evoked response properties across the visual field. *Electroencephalogr Clin Neurophysiol* 1994;90:65–81. [PubMed: 7509275]
2. Hood DC, Greenstein VC. Multifocal VEP and ganglion cell damage: applications and limitations for the study of glaucoma. *Prog Retin Eye Res* 2003;22:201–251. [PubMed: 12604058]
3. Klistorner AI, Graham SL, Grigg JR, et al. Multifocal topographic visual evoked potential: improving objective detection of local visual field defects. *Invest Ophthalmol Vis Sci* 1998;39:937–950. [PubMed: 9579473]
4. Klistorner A, Graham SL. Objective perimetry in glaucoma. *Ophthalmology* 2000;107:2283–2299. [PubMed: 11097611]
5. Graham SL, Klistorner AI, Grigg JR, et al. Objective VEP perimetry in glaucoma: asymmetry analysis to identify early deficits. *J Glaucoma* 2000;9:10–19. [PubMed: 10708226]
6. Hood DC, Zhang X, Greenstein VC, et al. An interocular comparison of the multifocal VEP: a possible technique for detecting local damage to the optic nerve. *Invest Ophthalmol Vis Sci* 2000;41:1580–1587. [PubMed: 10798679]
7. Hood DC, Zhang X. Multifocal ERG and VEP responses and visual fields: comparing disease-related changes. *Doc Ophthalmol* 2000;100:115–137.
8. Hasegawa S, Abe H. Mapping of glaucomatous visual field defects by multifocal VEPs. *Invest Ophthalmol Vis Sci* 2001;42:3341–3348. [PubMed: 11726643]

9. Goldberg I, Graham SL, Klistorner AI. Multifocal objective perimetry in the detection of glaucomatous field loss. *Am J Ophthalmol* 2002;133:29–39. [PubMed: 11755837]
10. Fortune, B.; Goh, K.; Demirel, S., et al. Detection of glaucomatous visual field loss using the multifocal visual evoked potential. In: Henson, DB.; Wall, M., editors. *Perimetry Update 2002/2003; Proceedings of the XVth International Perimetric Society Meeting*. The Hague: Kugler; 2003. p. 251-260.
11. Thienprasiddhi P, Greenstein VC, Chen CS, et al. Multifocal visual evoked potential responses in glaucoma patients with unilateral hemifield defects. *Am J Ophthalmol* 2003;136:34–40. [PubMed: 12834667]
12. Bjerre A, Grigg JR, Parry NR, et al. Test-retest variability of multifocal visual evoked potential and SITA standard perimetry in glaucoma. *Invest Ophthalmol Vis Sci* 2004;45:4035–4040. [PubMed: 15505053]
13. Hood DC, Thienprasiddhi P, Greenstein VC, et al. Detecting early to mild glaucomatous damage: a comparison of the multifocal VEP and automated perimetry. *Invest Ophthalmol Vis Sci* 2004;45:492–498. [PubMed: 14744890]
14. Graham SL, Klistorner AI, Goldberg I. Clinical application of objective perimetry using multifocal visual evoked potentials in glaucoma practice. *Arch Ophthalmol* 2005;123:729–739. [PubMed: 15955974]
15. Hood DC, Odel JG, Zhang X. Tracking the recovery of local optic nerve function after optic neuritis: a multifocal VEP study. *Invest Ophthalmol Vis Sci* 2000;41:4032–4038. [PubMed: 11053309]
16. Klistorner AI, Graham SL, Grigg J, et al. Objective perimetry using the multifocal visual evoked potential in central visual pathway lesions. *Br J Ophthalmol* 2005;89:739–744. [PubMed: 15923511]
17. Danesh-Meyer HV, Carroll SC, Gaskin BJ, et al. Correlation of the multifocal visual evoked potential and standard automated perimetry in compressive optic neuropathies. *Invest Ophthalmol Vis Sci* 2006;47:1458–1463. [PubMed: 16565379]
18. Fortune B, Demirel S, Zhang X, et al. Comparing multifocal VEP and standard automated perimetry in high-risk ocular hypertension and early glaucoma. *Invest Ophthalmol Vis Sci* 2007;48:1173–1180. [PubMed: 17325161]
19. Thienprasiddhi P, Greenstein VC, Chu DH, et al. Detecting early functional damage in glaucoma suspect and ocular hypertensive patients with the multifocal VEP technique. *J Glaucoma* 2006;15:321–327. [PubMed: 16865010]
20. Bengtsson B, Heijl A. Evaluation of a new perimetric threshold strategy, SITA, in patients with manifest and suspect glaucoma. *Acta Ophthalmol Scand* 1998;76:268–272. [PubMed: 9686835]
21. Johnson CA, Sample PA, Cio GA, et al. Structure and function evaluation (SAFE): I. criteria for glaucomatous visual field loss using standard automated perimetry (SAP) and short wavelength automated perimetry (SWAP). *Am J Ophthalmol* 2002;134:177–185. [PubMed: 12140023]
22. Hodapp, E.; Parrish, RK., II; Anderson, DR. *Clinical Decisions in Glaucoma*. St Louis: Mosby; 1993. p. 47
23. Wollstein G, Garway-Heath DF, Hitchings RA. Identification of early glaucoma cases with the scanning laser ophthalmoscope. *Ophthalmology* 1998;105:1557–1563. [PubMed: 9709774]
24. Ford BA, Artes PH, McCormick TA, et al. Comparison of data analysis tools for detection of glaucoma with the Heidelberg Retina Tomograph. *Ophthalmology* 2003;110:1145–1150. [PubMed: 12799239]
25. Fortune B, Zhang X, Hood DC, et al. Normative ranges and specificity of the multifocal VEP. *Doc Ophthalmol* 2004;109:87–100. [PubMed: 15675203]
26. Hood DC, Zhang X, Hong JE, et al. Quantifying the benefits of additional channels of multifocal VEP recording. *Doc Ophthalmol* 2002;104:303–320. [PubMed: 12076018]
27. Zhang X, Hood DC, Chen CS, et al. A signal-to-noise analysis of multifocal VEP responses: an objective definition for poor records. *Doc Ophthalmol* 2002;104:287–302. [PubMed: 12076017]
28. Fortune B, Demirel S, Zhang X, et al. Repeatability of normal multifocal VEP: implications for detecting progression. *J Glaucoma* 2006;15:131–141. [PubMed: 16633227]

29. Harper R, Reeves B. Compliance with methodological standards when evaluating ophthalmic diagnostic tests. *Invest Ophthalmol Vis Sci* 1999;40:1650–1657. [PubMed: 10393031]

FIGURE 1.

Example of results for 1 subject. The top row (A) shows the SAP 24-2 grayscale plots for the right and left eye. The second row (B) shows the SAP TD plots for the right and left eye. VF loss in the left eye of this patient was among the most severe in this study with an MD of -7.6 dB. The 2-Run mfVEP results are shown in the third row (C) and the 1-Run mfVEP results are shown in the bottom row (D). In 1C and 1D, the left panel shows the mfVEP response array with right eye responses in blue and left eye responses in red. The inset shows 1 pair of responses at higher magnification. The second panel shows the monocular probability plot for the right eye mfVEP and the third panel shows the monocular probability plot for the left eye mfVEP. The right-most panel shows the interocular probability plot. In all probability plots, a red colored square indicates an abnormal left eye response and a blue colored square indicates an abnormal right eye response. A saturated color indicates $P < 0.01$ and a desaturated (pastel) color indicates $P < 0.05$. Gray and open squares indicate that the SNR was too small to make a valid interocular comparison.

FIGURE 2.

Sensitivity of the mfVEP to detect GON versus false alarm rate when the masked expert grade of the stereo disc photo was used as the diagnostic standard. Abnormalities on the mfVEP were determined using various cluster criteria, each associated with an individual data point. Results for the 1-Run mfVEP (open circles) are similar to the results for the 2-Run mfVEP (filled diamonds).

FIGURE 3.

Sensitivity of the mfVEP to detect GON versus false alarm rate when the HRT MRA was used as the diagnostic standard. Details are same as in Figure 2 caption.

FIGURE 4.

Sensitivity of the mfVEP to detect early manifest glaucoma versus false alarm rate when the diagnostic standard was based on GON by stereo photo grade plus a SAP VF defect. Details are same as in Figure 2 caption.

FIGURE 5.

Sensitivity of the mfVEP to detect early manifest glaucoma versus false alarm rate when the diagnostic standard was based on GON by HRT MRA “outside normal limits” plus a SAP VF defect. Details are same as in Figure 2 caption.

Number of Eyes With an Abnormal mfVEP (top row, 2-Run mfVEP; Second Row, 1-Run mfVEP) for Each of the Cluster Types Listed in the top row

TABLE 1

Cluster Criterion	22	33	44	444	122	222	223	233	224	244
N abnormal eyes (2-Run)	87	107	162	112	47	48	65	75	77	100
N abnormal eyes (1-Run)	51	77	145	90	33	34	39	48	47	75
Agreement 1-Run vs. 2-Run	89%	89%	81%	89%	92%	92%	91%	91%	90%	90%
Agreement with SAP (2-Run)	78%	78%	68%	78%	81%	81%	81%	81%	79%	78%
Agreement with SAP (1-Run)	78%	79%	70%	79%	79%	79%	79%	79%	79%	78%

There were a total of 370 eyes (185 subjects). The numbers in the top row indicate the probability level of the points within the cluster (1= $P<0.005$; 2= $P<0.01$; 3= $P<0.02$; 4= $P<0.05$; 2 numbers indicate that the cluster contains 2 contiguous points within a hemifield, 3 numbers indicate 3 contiguous points within a hemifield. Agreement between 1-Run and 2-Run mfVEP classification (middle row). Agreement between SAP and mfVEP classifications (fourth row, 2-Run mfVEP; bottom row, 1-Run mfVEP).