Normal Visual Field Test Results Following Glaucomatous Visual Field End Points in the Ocular Hypertension Treatment Study

John L. Keltner, MD; Chris A. Johnson, PhD; Richard A. Levine, PhD; Juanjuan Fan, PhD; Kimberly E. Cello, BS; Michael A. Kass, MD; Mae O. Gordon, PhD; for the Ocular Hypertension Treatment Study Group

Objective: To compare the occurrence of normal visual field (VF) test results following 2 vs 3 consecutive, abnormal, reliable test results in the Ocular Hypertension Treatment Study.

Methods: The Ocular Hypertension Treatment Study is a prospective, multicenter follow-up study as part of a longitudinal randomized clinical trial. Sixty-four (68 eyes) of 1636 participants developed a VF primary open-angle glaucoma (POAG) end point. We compared the proportion of normal VF test results after a VF POAG end point among eyes whose VF abnormality was confirmed by 2 (n=9 eyes) vs 3 (n=59 eyes) consecutive, abnormal, reliable VF test results.

Results: The proportion of VF test results that were normal subsequent to a VF POAG end point in eyes whose abnormality was confirmed by 2 consecutive, abnormal, reliable test results was significantly higher (73 [66%] of 110) compared with eyes whose abnormality was confirmed by 3 consecutive, abnormal, reliable test results (46 [12%] of 381) (P=.01).

Conclusions: A VF POAG end point confirmed by 3 consecutive, abnormal, reliable VF test results appears to have greater specificity and stability than either 1 or 2 consecutive, abnormal, reliable VF test results. However, some eyes whose VF POAG end point was confirmed by 3 consecutive, abnormal, reliable test results nonetheless had 1 or more normal test results on follow-up.

Participants underwent at least 2 Humphrey Visual Field Analyzer (Carl-Zeiss Meditec, Dublin, Calif) 30-2 full-threshold visual field examinations in both eyes to determine visual field eligibility. A third examination was performed if one of the 2 prior test results was abnormal, questionable, or unreliable. For final visual field eligibility, the field test results had to be normal and reliable in both eyes on 2 examinations, as determined by the Visual Field Reading Center (VFRC) (University of California, Davis, Sacramento). The optic nerve heads also had to be normal in both eyes on clinical examination and review of stereoscopic optic disc photographs, as determined by the Optic Disc Reading Center, Miami, Fla. Follow-up visual field examination results were obtained at 6-month intervals. Neither the participants nor the health care professionals were masked to the randomization assignments during follow-up.

According to OHTS criteria, reliable visual field test results were defined as abnormal if the Glaucoma Hemifield Test results were outside of normal limits and/or the corrected pattern standard deviation was P<.05%. Initially, the OHTS protocol for confirming a visual field abnormality required 2 consecutive, abnormal, and reliable visual field test results, with the abnormality on the same index and in the same location (involving similar points as the previous visual field test results). The second visual field test was performed within 1 day to 8 weeks of the first abnormal visual field test results. A visual field abnormality was not considered confirmed if it was judged to be artifactual by the VFRC readers (ie, trial lens rim artifacts that disappeared on retest or superior depression that disappeared with taping of the eyelid).

The protocol for confirming visual field abnormality was changed on June 1, 1997, to require 3 consecutive, abnormal, and reliable visual field test results with the defect in the same location and on the same index. Thus, a participant with abnormal visual field test results was tested at the next regularly scheduled follow-up visit in 6 months. If the VFRC considered the second visual field test results abnormal, it requested a third visual field test to be completed in 1 day to 8 weeks. If the third visual field test results were considered abnormal by the VFRC director (J.L.K.) and codirector (C.A.J.) to have confirmed the visual field abnormality, the VFRC then prepared a narrative description of the abnormality and sent all visual field test results to the OHTS Coordinating Center for review by the OHTS End Point Committee. The End Point Committee members made an independent determination as to whether the visual field abnormality was attributable to POAG after carefully reviewing all the optic disc photographs, visual field test results, and medical and ocular history from baseline of both eyes of the participant. Disagreement between committee members was resolved by consensus. Committee members were masked as to the POAG classification of the affected and fellow eye, randomization status, and previous intraocular surgery.

In this report, we included POAG visual field data through November 8, 2001. Sixty-eight eyes reached a POAG visual field end point as of November 8, 2001, with an average follow-up of 35 months from POAG end point to last follow-up visit. Nine eyes (from 9 patients) reached a POAG visual field end point based on the confirmation criteria of 2 consecutive, abnormal, and reliable visual field test results with an average follow-up of 59 months from diagnosis of POAG to last follow-up visit. These eyes had 110 follow-up visual field tests after reaching a POAG end point, 73 (66%) of which had normal results. Fifty-nine eyes (from 56 patients) reached a POAG visual field end point based on the confirmation criteria of 3 consecutive, abnormal, and reliable visual field test results with an average follow-up of 31 months from diagnosis of POAG to last follow-up visit. These eyes had 381 follow-up visual field tests after reaching a POAG end point, 46 (12%) of which had normal results. The correlation between eyes from the same patient was significant at the P<.05 level (95% confidence interval on the random effect [0.47-2.38] does not cover zero). As shown in the Figure, the eyes reaching a visual field POAG end point based on the confirmation criteria of 3 consecutive, abnormal, and reliable visual field test results had a significantly smaller percentage of normal follow-up visual field test results than those reaching end point based on the confirmation criteria of 2 consecutive, abnormal, and reliable visual field test results (P=.01).

As a confirmation of our results, we considered a secondary analysis that constrained the follow-up period, forcing similar follow-up intervals between the 2 groups: eyes reaching visual field POAG end point based on the confirmation criteria of 3 consecutive, abnormal, and reliable visual field test results and eyes reaching visual field POAG end point based on the confirmation criteria of 2 consecutive, abnormal, and reliable visual field test results per eye, we defined the outcome variable as the proportion of visual field tests subsequent to a visual field POAG end point with normal results. We applied the arc-sin–square root transformation to the outcome variable, a transformation commonly used for comparing proportions in regression models. Results for proportion of visual field test results that were normal subsequent to a visual field POAG end point were obtained from a fit to a mixed model of the percentage of normal visual field test results, appropriately transformed, on the end point criteria, randomization assignment, baseline IOP (fixed effects), and patient (random effect). The random effect models the correlation between eyes from the same subject. The estimates and confidence intervals were obtained using restricted maximum likelihood in S-PLUS. A Cox proportional hazards model was additionally fit to study the effect of the strike rule on first reversion to a normal visual field test result following POAG onset. This model adjusted for the effects of randomization assignment, baseline IOP, and correlation between fellow eyes.

RESULTS

Sixty-eight eyes (from 64 patients) reached a POAG visual field end point as of November 8, 2001, with an average follow-up of 35 months from POAG end point to last follow-up visit. Nine eyes (from 9 patients) reached a POAG visual field end point based on the confirmation criteria of 2 consecutive, abnormal, and reliable visual field test results with an average follow-up of 59 months from diagnosis of POAG to last follow-up visit. These eyes had 110 follow-up visual field tests after reaching a POAG end point, 73 (66%) of which had normal results. Fifty-nine eyes (from 56 patients) reached a POAG visual field end point based on the confirmation criteria of 3 consecutive, abnormal, and reliable visual field test results with an average follow-up of 31 months from diagnosis of POAG to last follow-up visit. These eyes had 381 follow-up visual field tests after reaching a POAG end point, 46 (12%) of which had normal results. The correlation between eyes from the same patient was significant at the P<.05 level (95% confidence interval on the random effect [0.47-2.38] does not cover zero). As shown in the Figure, the eyes reaching a visual field POAG end point based on the confirmation criteria of 3 consecutive, abnormal, and reliable visual field test results had a significantly smaller percentage of normal follow-up visual field test results than those reaching end point based on the confirmation criteria of 2 consecutive, abnormal, and reliable visual field test results (P=.01). As a confirmation of our results, we considered a secondary analysis that constrained the follow-up period, forcing similar follow-up intervals between the 2 groups: eyes reaching visual field POAG end point based on the confirmation criteria of 3 consecutive, abnormal, and reliable visual field test results and eyes reaching visual field POAG end point based on the confirmation criteria of 2 consecutive, abnormal, and reliable visual field test results per eye, we defined the outcome variable as the proportion of visual field tests subsequent to a visual field POAG end point with normal results. We applied the arc-sin–square root transformation to the outcome variable, a transformation commonly used for comparing proportions in regression models. Results for proportion of visual field test results that were normal subsequent to a visual field POAG end point were obtained from a fit to a mixed model of the percentage of normal visual field test results, appropriately transformed, on the end point criteria, randomization assignment, baseline IOP (fixed effects), and patient (random effect). The random effect models the correlation between eyes from the same subject. The estimates and confidence intervals were obtained using restricted maximum likelihood in S-PLUS. A Cox proportional hazards model was additionally fit to study the effect of the strike rule on first reversion to a normal visual field test result following POAG onset. This model adjusted for the effects of randomization assignment, baseline IOP, and correlation between fellow eyes.
consecutive, abnormal, and reliable visual field test results. We limited the follow-up period for eyes under the 2 consecutive, abnormal, reliable visual field test results criteria to at most the median of the follow-up time of eyes under reaching POAG end point based on the confirmation criteria of 3 consecutive, abnormal, and reliable visual field test results.

Consequently, in this analysis, both groups had a median follow-up of 1135 days and a median of 6 follow-up visual field tests after POAG end point (thus, an equal follow-up interval and an equal number of follow-up visual field tests in the 2 groups). We fit a mixed model analogous to that put forth herein (adjusting for correlation between fellow eyes, treatment assignment, and IOP) and drew identical conclusions. Namely, the eyes reaching a visual field POAG end point based on the confirmation criteria of 3 consecutive, abnormal, and reliable visual field test results had a significantly smaller percentage of normal follow-up visual field test results (12%) than those reaching end point based on the confirmation criteria of 2 consecutive, abnormal, and reliable visual field test results (65%) (P = .03).

The frequency distribution of the percentage of subsequent normal follow-up visual field test results following an OHTS visual field POAG end point is presented in the Table. Of the 9 eyes reaching an end point based on 2 consecutive, abnormal, and reliable visual field test results, more than 76% of the subsequent follow-up field test results were normal for 4 of the eyes. Of the 59 eyes reaching end point based on 3 consecutive, abnormal, and reliable visual field test results, less than 25% of the subsequent follow-up field test results were normal for 49 of the eyes.

The odds of a reversion to normal visual field test results following POAG onset are significantly larger for eyes reaching a visual field POAG end point based on the confirmation of 2 consecutive, abnormal, and normal visual field test results compared with the confirmation of 3 consecutive, abnormal, and reliable visual field test results (P < .001; odds ratio, 5.70 [95% confidence interval, 2.94-11.05]). Furthermore, on adjusting for the confirmation rule (2 vs 3 abnormal test results), treatment is not significantly related to reversion to normal visual field test results (P = .10). Since all eyes received treatment after reaching a visual field POAG end point, this latter finding indicates that the extended treatment received by the medication group does not have any significant beneficial effect on reversion to normal visual field test results.

### Table. Frequency Distribution of Normal Visual Field Test Results Following an OHTS Visual Field POAG End Point

<table>
<thead>
<tr>
<th>Subsequent Normal Visual Field Test Results After POAG End Point, %</th>
<th>No. of Eyes at End Point Based on 2 Consecutive, Abnormal, and Reliable Visual Field Test Results (n = 9 eyes)</th>
<th>No. of Eyes at End Point Based on 3 Consecutive, Abnormal, and Reliable Visual Field Test Results (n = 59 eyes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-10</td>
<td>0</td>
<td>42</td>
</tr>
<tr>
<td>11-25</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>26-50</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>51-75</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>76-100</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>No visual field tests performed</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Abbreviations: OHTS, Ocular Hypertension Treatment Study; POAG, primary open-angle glaucoma.

In this report, we compare the proportion of visual field test results that are normal subsequent to a visual field POAG end point confirmed by 2 vs 3 consecutive, abnormal, and reliable visual field test results with the defect in the same location. We found that 66% (73/110) of the subsequent visual field test results were normal in the group of 9 eyes that reached a visual field POAG end point based on 2 consecutive, abnormal, and reliable visual field test results. However, we found that only 12% (46/381) of the follow-up field test results were normal from the 59 eyes that reached a visual field POAG end point based on 3 consecutive, abnormal, and reliable visual field test results. Thus, an OHTS visual field POAG end point confirmed by 3 consecutive, abnormal, and reliable visual field test results appears to have greater specificity and stability than either 1 or 2 consecutive, abnormal, and reliable visual field test results.

The OHTS protocol for determining the onset of POAG was designed to be highly specific to minimize clinical uncertainty as to who had developed POAG. The OHTS protocol included several provisions that protected the specificity of the diagnosis of POAG. The criteria for visual field abnormality were based on STATPAC. The protocol required that visual field abnormality be reproduced in the same location on separate, reliable tests completed at different visits. In June 1997, the visual field end point criterion was increased from 2 to 3 consecutive, abnormal, and reliable visual field test results when it was found that 85.9% of the eyes with initial visual field abnormalities in the OHTS had normal visual field test...
results on repeat testing. In addition, 3 consecutive, abnormal, and reliable visual field test results maintained a high specificity for determining POAG visual field end points. To protect against artifactual results, each abnormal visual field test result was reviewed by masked, trained readers at the VFRC. Visual field abnormalities that met the confirmation criteria were reviewed by the End Point Committee to ensure that only those abnormalities specifically due to POAG were defined as POAG end points. The masked End Point Committee reviewed case report forms, visual field test results, and optic disc photographs for both eyes from baseline to render an informed judgment as to whether the abnormality could be attributed to POAG. Only 125 (57%) of the 218 confirmed visual field abnormalities or optic disc deteriorations that were reviewed by the End Point Committee were attributed to POAG; 43% of the confirmed changes were due to causes other than POAG.7 Despite such thorough screening to protect the specificity of the determination of a visual field POAG end point, 12% of the eyes had 1 or more normal visual field test results subsequent to a visual field POAG end point, even though confirmation required 3 consecutive, abnormal, and reliable visual field test results. These results suggest that perimetric testing and/or early glaucomatous visual field loss may be inherently variable.

The follow-up interval differed between eyes in each of the 2 confirmation groups, POAG end point confirmed by 2 consecutive, abnormal, and reliable visual field test results and POAG end point confirmed by 3 consecutive, abnormal, and reliable visual field test results. In the analyses performed herein, we studied the percentage of normal visual field test results, which is the number of normal test results divided by the total number of tests taken, thus overcoming difficulties presented by differences between follow-up intervals. This conclusion is further validated by a secondary analysis that forced an identical median follow-up interval and median number of follow-up visits between the 2 confirmation groups and drew identical conclusions.

The treatment differed in timing for the observation group, that is the time of treatment application after 2 vs 3 consecutive, abnormal, and reliable visual field test results. For eyes assigned to the medication group, the confirmation rule effect is not compromised by such a timing difference because all eyes received treatment at baseline. In the observation or nonmedication group, there is potential for a treatment timing effect because the eyes converting under the 2 consecutive, abnormal, and reliable visual field test results confirmation rule, on average, start treatment sooner (after only 2 consecutive, abnormal, and reliable visual field test results as opposed to 3 consecutive, abnormal, and reliable visual field tests) and thus may have a higher likelihood of normal visual field test results following POAG end point. This timing effect cannot be accounted for using the data from this study. Our best option, as put forth in the analysis herein, is to adjust the confirmation rule effect for treatment/randomization assignment. The confirmation rule effect is thus a real effect for eyes treated at baseline and indicative of a significant confirmation rule effect among eyes randomized to the observation group.

SENSITIVITY AND VARIABILITY

Automated static perimetric threshold tests exhibit variability within a test procedure and from 1 examination to another, as reported in earlier investigations. The amount of variability is much higher in patients with glaucomatous visual field loss,7–16 especially at locations with reduced sensitivity. Previous studies have reported that increased visual field variability may be an early sign of glaucomatous damage.11 Hart and Becker demonstrated that glaucomatous visual field tests go through 3 transitional phases, which we have previously described.1 The initial phase has no defect demonstrable despite the fact that early damage is occurring. The second phase is a period in which shallow defects are often transient and are barely detectable. In the third phase, visual field defects progress at an uneven pace to become very dense.

Variability in visual field results is unlikely to be due to poor reliability or poor testing procedure. A high quality of visual field data has been maintained because of strict quality control measures used by the OHTS VFRC. The OHTS VFRC quality control system provides feedback on a regular basis to the clinical centers and to the visual field technicians about their performance and handling of the visual field data. In a prior presentation,17 the VFRC reported only 1325 (2.6%) of the 50925 regular follow-up visual field test results were beyond the 33% reliability limits for fixation losses, false-positive errors, or false-negative errors.

LONG-TERM VARIABILITY AND PROGRESSION

It is difficult to distinguish between progression of glaucomatous visual field loss and long-term variability unless several visual field test results are obtained over time. Thus, it is necessary to confirm changes to avoid “overcalling” progressive visual field loss. For example, Shulzer13 found in the Low Tension Glaucoma Study that 4 to 6 confirming visual field tests were needed to reliably determine visual field progression. Chauhan and colleagues14 defined progression as at least 4 nonedge test points beyond the P<.05 probability level on the Glaucoma Change Probability program, with complete overlap of at least 4 of these points on a confirming field. Although similar strategies for determining visual field progression were used in the Early Manifest Glaucoma Trial,18 the Glaucoma Change Probability program was based on the pattern deviation values rather than the total deviation values. In addition, 3 test points beyond the 5% level were needed for confirmation of progression on 3 successive visual field test results.

The Advanced Glaucoma Intervention Study (AGIS) visual field defect score is based on both the extent and depth of clusters of adjacent depressed test locations relative to age-matched normative data for individuals tested with the Humphrey Visual Field Analyzer 24–2 full-threshold testing program. The complex AGIS scoring system was designed for use with more advanced glaucomatous visual field loss and has been described in detail elsewhere.19,20 In the AGIS study, only 1 baseline visual field was obtained and 3 consecutive field tests
showed worsening of 4 AGIS units, indicating progression. During follow-up, in a subset in which the IOP was always lower than 18 mm Hg and there was no net visual field progression in 8 years, 14% of the subjects had an improvement of 4 AGIS units and 14% worsened at 5 years of follow-up. The Collaborative Initial Glaucoma Treatment Study (CIGTS) scoring system has also been described elsewhere and is based on the total deviation data from a Humphrey Visual Field Analyzer 24-2 full-threshold testing program but differs because it is based on probability rather than the depth of the defect. At 5 years of follow-up in CIGTS, no net visual field progression was noted in either the medicine-first or surgery-first arms. About 10% of the subjects met the criterion for progression at any point. The percentage improving by the same criterion was not reported. Comparison of AGIS and CIGTS scoring systems using a common longitudinal data set has shown that the CIGTS system identifies progression twice as frequently as the AGIS system.

An alternative to using specified end point changes on a specified number of consecutive field tests is to use a linear regression analysis. The length of follow-up required to detect progression using linear regression is influenced by a number of factors, including examination frequency, underlying rate and type of progression, the specific variable being evaluated, the degree of variability, and the order of the visual field tests within the series. Spry and Johnson and Vesti et al reported that a minimum of 7 or 8 visual field tests are required to achieve reasonable levels of sensitivity and specificity.

Confirmation of visual field change at successive examinations and correlation of visual field changes with other clinical observations (eg, optic disc change, nerve fiber layer change) seem to be the best method of detection of progression.

SPECIFICITY

It is critical to have a high specificity to evaluate the treatment effect in clinical trials. It is also important to have high specificity when diagnosing early POAG. Before a physician commits an individual to a lifetime of medical treatment, he or she should have a high degree of certainty about the diagnosis. The physician has the ability to evaluate other clinical parameters, such as optic nerve head measurements and intraocular pressures. Thus, the physician may not need 3 consecutive, abnormal visual field test results to confirm a glaucomatous visual field abnormality if an optic disc hemorrhage or progressive cupping is present. However, if the optic nerve findings are equivocal, the criterion of 3 consecutive, abnormal, and reliable visual field test results may be useful before instituting a lifetime of glaucoma treatment.

The clinical significance of a sporadic, abnormal visual field test result is not yet clear. It is possible that such a field test result is an indication of early glaucomatous damage. It is also possible that a dose-response relationship exists so that an individual with 2 abnormal visual field test results over 5 years of testing may be more likely to develop permanent glaucomatous damage than an individual with 1 abnormal field test result over 5 years of testing. Further follow-up of the OHTS participants may shed light on this question. Currently, we can conclude that confirmation of visual field abnormalities on at least 3 visual field tests helps to distinguish between the development of glaucomatous visual field loss and long-term variability.

Submitted for Publication: October 7, 2004; final revision received May 6, 2005; accepted May 16, 2005.

Correspondence: John L. Keltner, MD, Department of Ophthalmology, University of California, Davis, 4860 “Y” St, Suite 2400, Sacramento, CA 95817 (jlkeltner@ucdavis.edu).

Financial Disclosure: None.

Funding/Support: This study was supported in part by grants EY 09307 and EY 09341 from the National Eye Institute and the National Center on Minority Health and Health Disparities, National Institutes of Health, Bethesda, MD; Merck Research Laboratories, White House Station, NJ; and by an unrestricted grant from Research to Prevent Blindness, Inc, New York, NY.


**Correction**

In the Clinical Sciences article by Wollstein et al titled “Optical Coherence Tomography Longitudinal Evaluation of Retinal Nerve Fiber Layer Thickness in Glaucoma,” published in the April issue of the ARCHIVES (2005; 123:464-470), an error occurred in the key to Figure 2. The correct figure and key are reproduced below. The journal regrets the error.

**Figure 2.** Kaplan-Meier survival curve for visual field mean deviation decrease of 2 dB (VF-MD2) and optical coherence tomography mean retinal nerve fiber layer thickness (OCT-mean) for the entire study group ($P = .12$).