Estimating the Usefulness of Humphrey Perimetry Gaze Tracking for Evaluating Structure–Function Relationship in Glaucoma

Yukako Ishiyama, Hiroshi Murata, Hiroyo Hirasawa, and Ryo Asaoka

Department of Ophthalmology, the University of Tokyo, Tokyo, Japan

Correspondence: Ryo Asaoka, Department of Ophthalmology, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8655, Japan; rasaoka-kyu@umin.ac.jp

Submitted: August 18, 2015
Accepted: October 21, 2015


PURPOSE. We have previously reported that fixation loss (FL) rates, false-positive (FP) rates, and gaze tracking (GT) parameters (average tracking failure frequency per stimulus [TFF], average blinking frequency [BF], average frequency of eye movements between 1° and 2° [move1–2], between 3° and 5° [move3–5], and equal to or more than 6° [move≥6]) are related to the over- or underestimation of visual field (VF) results. The purpose of the current study was to validate these results by investigating the effect of implementing the GT parameters on the relationship between VF results and optical coherence tomography (OCT) measurements.

METHODS. Two hundred forty-four eyes of 155 open-angle glaucoma patients were included. Vision fixation during VF tests with the Humphrey Field Analyzer (24-2 SITA standard) was evaluated using the gaze fixation chart at the bottom of the VF printout. Mean total deviation (mTD) values were calculated, and their relationship with OCT-determined circumpapillary retinal nerve fiber layer (cp-RNFL), OCT-determined macular ganglion cell complex (GCC) thickness, and axial length was investigated using the corrected Akaike Information Criterion (AICc) in linear mixed modeling.

RESULTS. In the best model, average total cpRNFL thickness, average total GCC thickness, axial length, FL, FP, move1–2, move3–5, move≥6, TFF, and BF were selected as significant predictors (mTD = 2.1 + 0.097 X average total cpRNFL thickness + 0.089 X average total GCC thickness – 0.94 X axial length + 2.7 X FL + 7.2 X FP – 7.0 X move3–5 – 1.8 X move≥6 – 4.2 X TFF – 1.7 X BF).

CONCLUSIONS. Both GT parameters and classic VF reliability indices had significant influence on the structure-function relationship analysis in glaucoma.

Keywords: gaze tracking, glaucoma, visual field, optical coherence tomography

The Humphrey Field Analyzer (HFA; Carl Zeiss Meditec, Dublin, CA, USA) is used worldwide to monitor visual field (VF) damage in glaucoma patients. Several methods are used to estimate the reliability of HFA results. Fixation loss (FL) indicates test reliability and vision fixation by presenting test stimuli to the area of the blind spot. The rate of false-positive (FP) answers is measured to identify “trigger-happy” patients, and the rate of false-negative (FN) responses is measured to estimate inattention during examinations.1–3 The FP is estimated by maximum likelihood estimation based on the number of positive answers during a “listen time,” which starts shortly after the end of the response window and ends 180 ms after the onset of the next stimulus.4 A FN occurs when a patient fails to respond to a more intense stimulus than the ones to which the patient has previously responded. Some past studies have reported on the usefulness of these indices,5,6 but more recent studies have highlighted their limitations.7,8

Gaze tracking (GT) is a method to monitor eye movements: It measures the status of fixation when targets are presented in VF tests.9 It has been reported that GT is useful for evaluating the quality of fixation, particularly when VF defects surround the blind spot.10 In addition, we recently reported that some GT parameters are useful predictors of test-retest reproducibility.11 We also found that the FN rate was significantly associated with reproducibility but FL and FP were not. However, this finding does not detract from the importance of these indices as markers of the reliability of VF measurements. In particular, FL and FP may be good indices of the accuracy of VF measurements (whether “true” sensitivity is over- or underestimated); our recent study12 suggested that high FL and FP were predictive of overestimation of VF sensitivity in longitudinal series of VFs.

Recent developments in optical coherence tomography (OCT) have enabled clinicians to evaluate glaucoma from an anatomical point of view, using the circumpapillary retinal nerve fiber layer (cp-RNFL) thickness and the macular ganglion cell complex (GCC) thickness measurements.13–21 In the current study, glaucomatous damage was measured using VFs and OCT, and the usefulness of considering GT results, in addition to classic reliability indices, on the structure-function relationship was investigated.

METHODS

The study was approved by the Research Ethics Committee of the Graduate School of Medicine and Faculty of Medicine at the University of Tokyo. Written consent was given by patients for their information to be stored in the hospital database and used...
for research. This study was performed according to the tenets of the Declaration of Helsinki.

Subjects

Two hundred forty-four eyes of 155 open-angle glaucoma patients (75 males and 80 females) were included in the study. All patients were prospectively recruited at the glaucoma clinic in the University of Tokyo Hospital. Each patient underwent VF testing using the HFA (24-2 or 30-2 Swedish Interactive Threshold Algorithm, SITA, standard program) within 3 months from the OCT measurement. Visual fields with FL, FP or FN > 50% were excluded, following our previous report.12

Axial length was measured using the IOL Master (Carl Zeiss Meditec).

All patients enrolled in the study fulfilled the following criteria: (1) Glaucoma was the only disease causing VF damage; (2) patients were followed for at least 6 months at the University of Tokyo Hospital and had experienced at least two VF measurements prior to this study; (3) all patients had glaucomatous VF defects in at least one eye defined as three or more contiguous total deviation points at P < 0.05, or two or more contiguous points at P < 0.01, or a 10-dB difference across the nasal horizontal midline at two or more adjacent points, or MD worse than −5 dB; (4) absence of other systemic or ocular disorders, including cataract except for clinically insignificant senile cataract, shallow peripheral anterior chamber that could affect the optic nerve head (ONH) and VF and history of intraocular surgeries or refractive surgeries except for uneventful intraocular lens implantation; and (5) aged 20 years or older. All of the visual acuities of the eyes examined were equal to or better than 6/12.

Gaze Tracking Measurements

The GT system in the HFA monitors patients’ gaze position at each stimulus presentation (see Fig.).9 Gaze tracking data were exported from HFA printouts as JPEG images using the Beeline (Tokyo, Japan) data filing system. An upward bar in the GT chart represents fixation disparity, and the length of the bar indicates the magnitude of disparity. A short downward bar represents tracking failure, while a long downward bar indicates eyelid closure. GT parameters were calculated as follows: average frequency of eye movement per stimulus between 1° and 2° (move1-2), 3° and 5° (denoted move1,2), between 3° and 5° (denoted move2,3), and equal to or more than 6° (denoted move ≥ 6), following our previous report.11

Optical Coherence Tomography Measurements

Optical coherence tomography measurements were carried out using the 3D OCT-2000 (Topcon Corp., Tokyo, Japan) raster-scan protocol; data were obtained in a 7.0 × 7.0-mm square area, centered on the point of fixation, in 1.3 seconds. Registration of fundus photographs and OCT images was automatically confirmed using an OCT projection image (generated from 3D-OCT data by summing different retinal depth levels) and localization of major retinal vessels. The cp-RNFL thickness was obtained using tomographic images of the parapapillary fundus with the three-dimensional (3D) disc scan and 3D macula scan (128 horizontal scan lines composed of 512 A-scans for an image area of 6 × 6 mm). In the current study, GCC thickness was measured as the thickness from the internal limiting membrane to the inner nuclear layer. All OCT measurements were performed after pupil dilation with 1% tropicamide. The optic disc center was determined in fundus photographs as the barycenter of the closed spline curve fitted to the manually determined 7 points; the point was then extrapolated in all OCT images thereafter. Images that were clearly influenced by blinks (incomplete images) or saccades were manually excluded, and those with quality factors of less than 60% were excluded.

Statistical Analysis

The means of the 52 total deviation (TD) values in each of the 24-2 VF tests were calculated (mTD). When the VF was measured using the 30-2 test pattern, only the 52 test points overlapping with the 24-2 test pattern were used for the

<table>
<thead>
<tr>
<th>Table 1. Parameters Used in Model Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Analyzed Parameters</strong></td>
</tr>
<tr>
<td>Age, y</td>
</tr>
<tr>
<td>Axial length, mm</td>
</tr>
<tr>
<td>OCT parameters</td>
</tr>
<tr>
<td>Average of total cpRNFL thickness, μm</td>
</tr>
<tr>
<td>Average of total GCC thickness, μm</td>
</tr>
<tr>
<td>Traditional reliability parameters</td>
</tr>
<tr>
<td>FL, %</td>
</tr>
<tr>
<td>FP, %</td>
</tr>
<tr>
<td>GT parameters</td>
</tr>
<tr>
<td>Average frequency of eye movement per stimulus between 3° and 5° (move1,2)</td>
</tr>
<tr>
<td>Average frequency of eye movement per stimulus of 6° or more (move ≥ 6)</td>
</tr>
<tr>
<td>Average tracking failure frequency per stimulus (TFF)</td>
</tr>
<tr>
<td>Average blinking frequency per stimulus (BF)</td>
</tr>
</tbody>
</table>
The relationship between mTD and average total cpRNFL thickness, average total GCC thickness, axial length, age, move<sub>3-5</sub>, move<sub>6</sub>, TFF, and BF, as well as FL and FP (see Table 1), was analyzed using a linear mixed model. The optimal linear model was then selected among all possible combinations of predictors (210 patterns) based on the second-order bias corrected Akaike Information Criterion (AICc) index, following our previous report.11 The AIC is a well-known statistical measure used in model selection, and the AICc is a corrected type of the AIC, which provides an accurate estimation especially when the sample size is small.23 The degrees of freedom in a multivariate regression model decreases with a large number of variables. It is therefore recommended that one use model selection methods to improve the model fit by removing redundant variables.24,25

All analyses were performed using the statistical programming language R (R version 3.1.3; The Foundation for Statistical Computing, Vienna, Austria).

### RESULTS

Characteristics of the study subjects are summarized in Table 2. The mean (± standard deviation, SD) age of the patients was 52.4 ± 13.5 years, ranging from 20 to 85 years. The mTD values were −5.0 ± 5.2 (mean ± SD [range, −24.1 to 2.0]) dB. Total cpRNFL and GCC thicknesses were 91.9 ± 16.8 (mean ± SD [range, 49.0–150.9]) and 98.4 ± 23.9 (mean ± SD [range, 49.8–91.9]) μm.

Fixation loss rates, FP rates, and FN rates were 5.1 ± 7.2 [0.00–38.5]%), 3.0 ± 3.7 [0.00–21.0]% and 4.1 ± 5.4 [0.00–46.0]%, respectively (Table 3).

Move<sub>3-5</sub>, move<sub>6</sub>, TFF, and BF results were 0.10 ± 0.11 [0.00–0.55] per stimulus, 0.051 ± 0.12 [0.00–0.94] per stimulus, 0.055 ± 0.074 [0.00–0.91] per stimulus, and 0.020 ± 0.044 [0.00–0.49] per stimulus, respectively (Table 4).

As shown in Table 5, AICc was 1404.2 for the linear relationship between mTD and average total cpRNFL thickness, average total GCC thickness, and axial length. The AICc decreased when adding FL and FP (1396.2). The smallest AICc (1373.1) was observed when average total cpRNFL thickness, average total GCC thickness, axial length, FL, FP, move<sub>3-5</sub>, move<sub>6</sub>, TFF, and BF were included (mTD = 2.1 + 0.097 × average total cpRNFL thickness + 0.089 × average total GCC thickness − 0.94 × axial length + 2.7 × FL + 7.2 × FP − 7.0 × move<sub>3-5</sub> − 1.8 × move<sub>6</sub> − 4.2 × TFF − 1.7 × BF). This optimal formula is identical to mTD − 0.097 × average total cpRNFL thickness − 0.089 × average total GCC thickness + 0.94 × axial length = 2.1 + 2.7 × FL + 7.2 × FP − 7.0 × move<sub>3-5</sub> − 1.8 × move<sub>6</sub> − 4.2 × TFF − 1.7 × BF, in which the parameters for structure and function are located on the left-hand side and reliability indices are located on the right-hand side.

Age was not selected as a variable in the optimal model.

### DISCUSSION

In the current study, GT results were extracted from VF tests and analyzed quantitatively and objectively. The relationship between VF damage (mTD) and structural damage (thickness measurements estimated with OCT) was investigated in conjunction with axial length, GT, and other reliability parameters. As a result, average total cpRNFL thickness, average total GCC thickness, axial length, FL, FP, move<sub>3-5</sub>, move<sub>6</sub>, TFF, and BF were selected as significant predictors of mTD.

In this study, cpRNFL and GCC thicknesses were both significantly related with mTD, in agreement with numerous previous studies.15–21 It remains controversial which one of the two OCT parameters, cpRNFL or GCC, is more useful for diagnosing glaucoma, despite many previous studies.15–17,19 probably because they depend on the particular characteristics of the eye, such as the size of the optic disc20 and refractive status.21 In the current study, both cpRNFL and GCC thicknesses were selected in the best linear model, with very similar coefficient values (0.097 and 0.089, respectively).

Both FL and FP were selected as significant predictors of mTD in the optimal model. This finding is supported by the result from our previous study in which the influence of FL, FP, and FN on longitudinal VF results was investigated, and, as a result, both FL and FP were selected as important predictors.12 In the current study, in the best model, both move<sub>3-5</sub> and move<sub>6</sub> were selected with negative coefficients. This is also in agreement with our previous study12; furthermore, the coefficients of these parameters were similar (FL 2.7 and FP 7.2 in the current study and FL 0.9 and FP 9.2 in the previous study).12 despite differences in study designs. Clinicians should be careful when assessing VFs with frequent large eye movements, as observed in GT records. Also, Jansonius20 has reported that the ability to detect progression is largely influenced by the variability of VF results in the time course. At the moment, variability of VFs is assessed using only FL, FP, and FN, but our results suggest it is advantageous to interpret GT parameters when assessing VF results. This is also supported by our other previous report in which the relationship between test–retest reproducibility of VFs and some GT parameters, TFF and move<sub>3-5</sub>, were confirmed.11

The FN rate was not included as a possible predictor of VF damage since it is closely related to the damage of VF.27 Nonetheless, this does not deny the usefulness of the FN rate to assess test reliability. Indeed, we have shown that this index is useful for estimating test–retest reproducibility11, hence, FN
results should certainly not be ignored when interpreting VFs in the clinical setting. The average frequency of eye movement per stimulus between 1° and 2° was not included as a variable in the optimal model. This was also not a useful parameter in our previous study.12 These results are not surprising when we consider that VF test points are located at 6° intervals in the 24-2 and 30-2 VF test patterns. Furthermore, a previous study has reported that eye movements of less than 3° are commonly observed in VF tests, even in well-trained healthy observers.28,29 Surprisingly, move3-5 had a larger coefficient than move3-6 in the optimum formula. The reason for this is not clear, but we hypothesize that move3-5 mainly reflects misfixations during the VF measurement whereas move3-6 could be a result of a lack of concentration during the VF test. As misfixations can be related to VF damage (the more damaged, the poorer fixated), move3-5 could have a larger coefficient than move3-6. Furthermore, the test grid interval in 24-2 VFs may be important since an eye movement of 6° corresponds to an adjacent test point.

In our previous study,12 it was suggested that VFs with high TFF and BF values tended to be associated with the underestimation of MD. In agreement with this finding, these parameters were also selected in the best model in the current study. Further research is needed, however, to disentangle the underlying causes of TFF, which is assumed to include long intervals in the 24-2 VFs.24,25

One possible caveat regarding the current investigation is that 10-2 VFs were not included. The imaging area used to derive GCC measurements corresponds only to the central area of the VF, which is approximately 10°; thus, a closer structure-function relationship would be expected with 10-2 VFs.

In the current study, GT data were exported as JPEG images from the Beeline data filing system, and various GT parameters were simply calculated by reading the JPEG image. Thus, GT parameters could be obtained on a personal computer; it would be advantageous to develop a clinical support tool to allow clinicians to access GT parameters.

In conclusion, we have investigated the relationship between VF damage and OCT-measured structural damage, with the inclusion of reliability indices and GT results. As a consequence, GT parameters were found to be significant predictors of VF test results.

**Acknowledgments**

Supported in part by 26462679 (RA) and 60645000 (HH) from the Ministry of Education, Culture, Sports, Science and Technology of Japan and Japan Science and Technology Agency (JST) CREST (RA, HM).

Disclosure: Y. Ishiyama, None; H. Murata, None; H. Hirasawa, None; R. Asaoka, None

**References**


