Glaucoma

The Relationship Between Visual Acuity and the Reproducibility of Visual Field Measurements in Glaucoma Patients

Masato Matsuura,1 Kazunori Hirasawa,1,2 Hiroshi Murata,2 and Ryo Asaoka1

1Department of Ophthalmology, University of Tokyo Graduate School of Medicine, Tokyo, Japan
2Orthoptics and Visual Science, Department of Rehabilitation, School of Allied Health Sciences, Kitasato University, Kanagawa, Japan

Correspondence: Ryo Asaoka, Department of Ophthalmology, University of Tokyo Graduate School of Medicine, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan; rasaoka-tky@umin.ac.jp.
Submitted: June 27, 2015
Accepted: July 16, 2015
Citation: Matsuura M, Hirasawa K, Murata H, Asaoka R. The relationship between visual acuity and the reproducibility of visual field measurements in glaucoma patients. Invest Ophthalmol Vis Sci. 2015;56:5630–5635. DOI:10.1167/iovs.15-17576

Purpose. The purpose of this study was to investigate the association between visual acuity (VA) and reproducibility of test-retest visual field (VF) measurements in glaucoma patients.

Methods. Subjects were comprised of 627 eyes of 627 open-angle glaucoma patients. The reproducibility of two Humphrey VFs (24-2 or 30-2 Swedish Interactive Threshold Algorithm tests) examined twice within the period of 3 months was calculated using the root mean squared error (RMSE) of each VF test point’s sensitivity. Visual acuity was measured once at the time of either of the VF measurements. Linear modeling was used to investigate the relationship between reproducibility of VF tests (RMSE) and the following variables: mean total deviation value (mTD), fixation losses (FLs), false positives (FPs), false negatives (FNs), refractive error, age, and VA.

Results. The optimal model to predict test-retest variability (RMSE) of VFs included age, VA, mTD, and FNs as dependent variables. Root mean squared error was significantly larger in eyes with logMAR VA > 0.5 than in eyes with logMAR VA ≤ 0.5.

Conclusions. Reproducibility of VF tests becomes poor with the deterioration of VA. Careful consideration is needed when a patient’s logMAR VA exceeds 0.5.

Keywords: visual field, visual acuity, reproducibility

Glaucoma can severely damage a patient’s visual field (VF) and is one of the major causes of blindness worldwide.1,2 In glaucoma, early detection of VF progression is essential, because the damage is irreversible. Measured VF sensitivity fluctuates3,4 and it has been widely reported that the detectability of progression is markedly reduced in series of VFs with poor reproducibility.5 Thus, the reliability of VF measurements is very important and is usually estimated using fixation loss (FL), false-positive (FP), and false-negative (FN) rates. While previous studies have reported on the usefulness of these indices,6,7 more recent studies have pointed out their limitations. A high rate of FL can result from mislocalization of the blind spot,8 and fixational instability can be found even in well-trained observers.9,10 False negatives, on the other hand, are inherently associated with VF deterioration.11

In general, VF reproducibility deteriorates with the progression of glaucoma.12 In addition, visual acuity (VA) is closely correlated with sensitivities in the central area of the measured VF13 which is usually spared until later stages of the disease. Consequently, VA can be closely related to the reproducibility of VFs. Indeed in many studies where VF tests are measured, VA is included in the inclusion criteria; however, peculiarly, there is no consensus regarding an appropriate cut-off level for VA. Hugely variable VA inclusion criteria have been used across previous studies, such as equal to or better than 6/24,14 20/40,15–21 (including in publications from the Early Manifest Glaucoma Trial,12,13 The Collaborative Initial Glaucoma Treatment Study,15 The Ocular Hypertension Treatment Study,16 The United Kingdom Glaucoma Treatment Study,17 and The Blue Mountains Eye Study20), 6/12,29 20/20,30 20/30,31–33 (including publications from the Tajimi study34 and the Collaborative Normal-Tension Glaucoma Study35), 20/80,36 (including publications from the Advanced Glaucoma Intervention Study37–39 and The Beijing Eye Study40), and 20/200 41 (VAs were converted to fractional VA).

The purpose of the current study is to investigate the association between VF reproducibility and VA. This is an important question since VF measurements are often the primary end point in clinical research studies, including clinical trials evaluating the effect of pharmacologic agents,29–42 surgery,43–48 and other interventions.49–51

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 and the University of Kitasato Hospital. 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

Six hundred twenty-seven eyes of 627 open-angle glaucoma patients (339 males and 288 females) were included in this
TABLE 1. Subjects’ Demographics

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean ± SD (range)</td>
<td>61.0 ± 14.0 (21 to 87)</td>
</tr>
<tr>
<td>Sex, M:F</td>
<td>339:288</td>
</tr>
<tr>
<td>Refractive error, D</td>
<td>−2.6 ± 3.8 (−21 to 8)</td>
</tr>
<tr>
<td>VA, logMAR</td>
<td>0.068 ± 0.22 (−0.18 to 1)</td>
</tr>
<tr>
<td>mTD in the initial VF, dB, mean ± SD (range)</td>
<td>−10.6 ± 7.4 (−28.6 to 2.5)</td>
</tr>
<tr>
<td>mTD in the last VF, dB, mean ± SD (range)</td>
<td>−10.4 ± 4.1 (−28.6 to 1.6)</td>
</tr>
<tr>
<td>FL, %, mean ± SD (range)</td>
<td>5.0 ± 4.2 (0 to 17.0)</td>
</tr>
<tr>
<td>FP, %, mean ± SD (range)</td>
<td>2.5 ± 2.4 (0 to 11.0)</td>
</tr>
<tr>
<td>FN, %, mean ± SD (range)</td>
<td>5.6 ± 5.5 (0 to 33.0)</td>
</tr>
</tbody>
</table>

All participants underwent 24-2 or 30-2 Humphrey Field Analyzer II (HFA) (Carl Zeiss Meditec, Inc., Dublin, CA, USA) VF tests twice (the same mode was used in both measurements) within 3 months at the glaucoma clinic in the University of Tokyo Hospital or University of Kitasato Hospital. Only one eye in a patient was included in the current study, and if both eyes satisfied the inclusion criteria, one eye was chosen at random.

All participants in the study fulfilled the following criteria: glaucoma was the only disease causing VF damage; patients were followed for at least 6 months at either the University of Tokyo Hospital or the University of Kitasato Hospital and experienced at least two VF measurements prior to the study; and all patients had glaucomatous VF defects in at least one eye defined as three or more contiguous total deviation points at P < 0.05, two or more contiguous points at P < 0.01, a 10-dB difference across the nasal horizontal midline at two or more adjacent points, or mean deviation (MD) worse than −5 dB.52

Visual Acuity

All participants underwent measurements of subjective refraction and corneal dioptic power with an autokeratorefractometer without cycloplegia. Visual acuity was measured using a Landolt ring chart at a distance of 5 m from the illuminated target; each ring was presented one by one, and the visual target was presented between three and five times in various directions, and VA was decided only when three answers were correct. If the VA was not 20/20 or better, refractive correction was carried out beginning with the results of autokeratorefractometry, and the corrective lenses were adjusted manually. The refractive error was measured in −0.25 diopter (D) steps, and the cylindrical power was measured and recorded in negative form. The refractive error was determined according to the results of corrective lenses that provided the best-corrected VA in eyes.34 Visual acuity was calculated as logMAR VA and study eyes were divided into five subgroups using logMAR VA (group 1, ≤0; group 2, >0 and ≤0.15; group 3, >0.15 and ≤0.35; group 4, >0.35 and ≤0.5; group 5, >0.5).

Visual Field

Visual field tests were undertaken using the HFA II, with a Goldmann size III stimulus under standard perimetric conditions (background, 10 candela [cd]/m²) and the Swedish Interactive Threshold Algorithm (SITA) standard strategy. Visual fields with fixation loss (FL) > 20% or false positive (FP) > 15% were excluded, following the criteria used by the HFA software; false negative (FN) was not used as an exclusion criterion.

Statistical Analysis

Reproducibility of the measured VF was assessed by taking the root mean squared error (RMSE) of the measured threshold at each test point:

\[
\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{52} (\text{threshold of the ith test point [first VF]} - \text{threshold of the ith test point [second VF]})^2}{52}}.
\]

For this calculation, only the 52 test points corresponding to the 24-2 VF test pattern were used when the VF was measured using the 30-2 program. Mean of the total deviation (mTD) values of the 52 test points were calculated for each VF, and the mean of the mTD values in the first and second VF were calculated. Then the relationship between the reproducibility (RMSE) and the variables VA, FP, FN, FL, refractive error, mTD, and age was analyzed using linear modeling. The optimal linear model was then selected among all possible combinations of predictors based on the second-order bias corrected Akaike Information Criterion (AICc) index. The AIC is a well-known statistical index used in model selection, and AICc is a corrected version of the AIC, which provides an accurate estimation, especially when the sample size is small.53 The degree of freedom in a multivariate regression model decreases with a large number of variables. This leads to an unbiased estimate of the risk of the linear regression fit, and it is recommended to use the model selection to improve the model fit by adequately removing redundant variables.54,55 False positive, FL, FN, and mTD were calculated based on their mean values in the two VF tests. Values were compared using the Wilcoxon test (paired), and Benjamini’s method was used to correct P values for the problem of multiple testing.56 The relationship between variables was analyzed using linear regression.

All analyses were performed using the statistical programming language R (R version 2.15.1; The Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Demographics of the study subjects are summarized in Table 1. The mean (±SD) age of patients was 61.0 ± 14.0 years, ranging from 21 to 87 years. The mTD value in the initial VF was −10.6 ± 7.4 [−28.6 to 2.50] (mean ± SD [range]) and −10.4 ± 4.1 [−28.1 to 1.6] in the second VF. The pattern standard deviation (PSD) in the initial VF and second VF was 9.7 ± 4.2 [17.5–1.5] and 9.6 ± 4.1 [17.6–1.3], respectively. There was no significant difference between mTD and PSD in paired VF tests (paired t-test, P = 0.21 and 0.51, respectively).

The distribution of RMSE in each VA group is shown in Figure; the mean root mean squared error (RMSE) was 5.0 ± 2.0 [1.1–15.7] in group 1 (logMAR VA ≤ 0), 5.3 ± 2.0 [1.7–12.9] in group 2 (logMAR VA > 0 and ≤ 0.15), 5.1 ± 1.9 [1.8–13.9] in group 3 (logMAR VA > 0.15 and ≤ 0.35), 5.4 ± 2.1 [1.7–11.8] in group 4 (logMAR VA > 0.35 and ≤ 0.5), and 6.4 ± 2.1 [2.4–11.6] in group 5 (logMAR VA > 0.5). The RMSE in group 5 was significantly larger than in group 1 (Wilcoxon test, with Benjamini’s correction method for multiple comparisons, P < 0.001); however, there was no significant difference between RMSE in group 1 versus groups 2, 3, and 4 (Wilcoxon test, with Benjamini’s correction method for multiple comparisons, P = 0.15, P = 0.65, and P = 0.25, respectively). There was a significant relationship between RMSE and logMAR (RMSE = 5.08 ± 1.38 × logMAR VA; R² = 0.021; P = 0.0002).

Average FL, FP, and FN rates are shown in Table 1. There was no significant relationship between VA and these VF...
reliability indices: \( P \) values were 0.25, 0.23, and 0.16, respectively (simple linear regression). Nor was any significant relationship observed between VA and these indices when using multivariable regression analysis (\( P = 0.20 \)). Conversely, mTD was significantly related to VA (simple linear regression, mTD = \(-9.9 - 9.4 \times \text{logMAR VA} \); \( R^2 = 0.075, P = 0.001 \)).

Demographic data in each VA group are shown in Table 2. Significant differences were observed between groups 2, 3, 4, and 5 against group 1 for age and mTD in the initial and last VFs, group 5 against group 1 for refractive error, and groups 2 and 5 against group 1 for FP rate (Wilcoxon test, with Benjamini's correction method for multiple comparisons, \( P < 0.05 \)).

Age, VA, mTD, and FN were selected as predictors in the optimal model (RMSE = 4.46 + 4.38 \times \text{age} + 13.28 \times \text{logMAR VA} + 0.84 \times \text{mTD} - 0.01 \times \text{FN}).

**DISCUSSION**

In the current study, the relationship between the reproducibility of VF measurements and VA was investigated. As a result, it was observed that the reproducibility of the VF appreciably becomes worse as VA decreases, as shown in the significant relationship between RMSE and logMAR VA; in particular, reproducibility significantly decreases when logMAR VA is > 0.5 (Table 2). Furthermore, VA was not significantly related to well-known VF reliability indices (FL, FP, and FN). We recommend that a cut-off value of logMAR VA > 0.5 should be used as an inclusion criterion when using VF outcomes in research, such as studies evaluating structure-function, progression prediction, and progression detection and also when defining the outcome of any observation/intervention, including assessing the effect of pharmacologic agents and surgical treatments. Jansonius previously analyzed the relationship between the variability of the VF and the time needed to detect progression. As a result, it was clearly shown that the detection of progression is delayed in VFs with large variability.\(^5\)\(^7\) Thus, careful consideration should be given when interpreting VF progression in eyes with logMAR VA > 0.5.

In this study, the optimal model for explaining the reproducibility of VFs included age, VA, mTD, and FN as predictors, whereas FP, FL, and refractive error were not

---

**TABLE 2. Demographic Data in Each VA Group**

<table>
<thead>
<tr>
<th>Demographics</th>
<th>( &gt;0.5 )</th>
<th>( &gt;0.35 ) and ( \leq 0.5 )</th>
<th>( &gt;0.15 ) and ( \leq 0.35 )</th>
<th>( \leq 0.15 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean ± SD (range)</td>
<td>64.6 ± 14.6 (22 to 88)</td>
<td>64.7 ± 12.4 (23 to 80)</td>
<td>65.3 ± 12.6 (36 to 87)</td>
<td>65.3 ± 12.6 (36 to 87)</td>
</tr>
<tr>
<td>Sex, n/MF</td>
<td>202/170</td>
<td>193/180</td>
<td>200/180</td>
<td>193/180</td>
</tr>
<tr>
<td>mTD in the initial VF, dB</td>
<td>-3.8 ± 7.3 (-10 to 10)</td>
<td>-3.7 ± 5.3 (-10 to 10)</td>
<td>-3.7 ± 5.3 (-10 to 10)</td>
<td>-3.7 ± 5.3 (-10 to 10)</td>
</tr>
<tr>
<td>mTD in the last VF, dB</td>
<td>-3.8 ± 7.3 (-10 to 10)</td>
<td>-3.7 ± 5.3 (-10 to 10)</td>
<td>-3.7 ± 5.3 (-10 to 10)</td>
<td>-3.7 ± 5.3 (-10 to 10)</td>
</tr>
<tr>
<td>FL, %, mean ± SD (range)</td>
<td>5.7 (0 to 30)</td>
<td>4.7 (0 to 30)</td>
<td>4.7 (0 to 30)</td>
<td>4.7 (0 to 30)</td>
</tr>
<tr>
<td>FP, %, mean ± SD (range)</td>
<td>2.3 (0 to 10)</td>
<td>1.9 (0 to 10)</td>
<td>1.9 (0 to 10)</td>
<td>1.9 (0 to 10)</td>
</tr>
<tr>
<td>FN, %, mean ± SD (range)</td>
<td>5.8 (0 to 30)</td>
<td>5.3 (0 to 30)</td>
<td>5.3 (0 to 30)</td>
<td>5.3 (0 to 30)</td>
</tr>
</tbody>
</table>

*P* < 0.05 against logMAR VA \( \leq 0.5 \) group (Wilcoxon test with Benjamini's correction method for multiple comparisons).
selected. It is widely acknowledged that the reproducibility of VFs varies according to the level of VF damage; reproducibility decreases with early to moderate VF damage, although it increases again because of the floor effect. In agreement with this, we found that mTD had a negative coefficient in our optimal regression model. Refractive error was not selected as an important predictor, in agreement with a previous study.

We previously analyzed the relationship between VF reproducibility and VF reliability indices, and as a result, it was shown that FN rate is a good predictor of VF reproducibility even when the level of VF damage is considered, but FL and FP were not found to be useful parameters in assessing VF reproducibility. In agreement with this, FN was selected in the best model in the current study, but FL and FP were not selected. Fixation loss is recorded when a stimulus projected onto the eye’s blind spot is perceived, and it indicates the test reliability and vision fixation. However, FL can also result from mislocalization of the blind spot, and moreover, it has been reported that fixational instability can be found even in well-trained observers. Importantly, FL does not directly measure eye fixation at the time a stimulus is presented (and the actual threshold measurement is taken at each test point). Indeed, we previously reported on the usefulness of the gaze tracking record as a reliability indicator, and this is a direct measure of eye fixation during stimulus presentation. In the SITA algorithm, any response prior to the minimum response time (~180 ms), adjusted according to the patient’s individual mean response time, is considered an FP error. Thus, all FP responses after the minimum response time are ignored in the FP calculation: this will reduce the performance of the indicator to measure the “trigger happy” attitude of some patients.

In glaucoma, central visual function is usually preserved until late stage disease. We previously reported that VA is closely related with foveal sensitivity. A close relationship between VA and foveal sensitivity was also observed in the current study (R² = 0.45, P < 0.001; data not shown), and RMSE was significantly related with foveal sensitivity (R² = 0.024, P < 0.001; data not shown). Visual acuity is most closely related to VF sensitivity in the innermost 17 test points (including the fovea) of the 10-2 HFA VF; this region overlaps with the 4 most central test points in the 24-2/30-2 HFA VF (approximately 3° from the fovea). Interestingly, it has been reported that a small fixational instability, such as within 3°, can be found even in well-trained observers, and we recently reported that the average frequency of eye movement between 1° and 2° was 62% in reliable VFs (defined as FL < 20% and FP < 15%). Another report suggested that eyes move 2.9° on average in reliable 10-2 HFA tests.

False-negative rate is known to increase with the progression of glaucoma, which, as already pointed out, is associated with lower VF reproducibility. Bengtsson and Heijl reported that only FN rate is associated with VF reproducibility among the traditional reliability indices of FL, FP, and FN. In agreement with this, our result suggests that FN is a useful parameter to estimate the reproducibility of VF even when the status of glaucomatous damage (mTD) is considered. We reported a similar result recently.

In the current study, VFs with FL > 20% or FP > 15% were excluded. This is because the purpose of the study was to investigate the influence of VA on the reproducibility of VFs in reliable VF, as required in clinical trials and indeed in general clinical assessment. As a result, it was suggested that careful consideration is needed when logMAR VA exceeds 0.5, even when these reliability criteria are met. Thus, it was not investigated how VA is related to poor reliability indices. A future study would be needed to shed light on this issue.

In conclusion, VA is a useful parameter to assess the reproducibility of VFs independently from traditional reliability indices. In particular, careful consideration is needed when logMAR VA exceeds 0.5, because the reproducibility becomes significantly poorer at this level.

Acknowledgments

Supported by the Technology of Japan and Japan Science and Technology Agency (JST) CREST and Ministry of Education, Culture, Sports, Science and Technology of Japan Grant 26462679.

Disclosure: M. Matsura, None; K. Hirasawa, None; H. Murata, None; R. Asaoka, None

References


