Evaluating the Accuracy of the Visual Field Index for the Humphrey Visual Field Analyzer in Patients with Mild to Moderate Glaucoma

RENÉE TALBOT, IVAN GOLDBERG, AND PATRICK KELLY

• PURPOSE: **To evaluate the accuracy of the visual field index (VFI) for the Humphrey Visual Field Analyzer in a population of patients with mild to moderate glaucoma.**
• DESIGN: **Retrospective cohort study.**
• METHODS: **The study included 42 patients (61 eyes) with at least 11 years of follow-up, and annual automated visual fields (VFs). Patients with mean deviations ≤−20 dB were excluded. All unreliable fields were omitted (fixation losses ≥20%, false-positive ≥15%, false-negative ≥33%). The VFs were divided into two 5-year series and the data were analyzed by the new Humphrey Visual Field Analyzer software. Projected VFIs from the first 5 years were compared with observed values obtained from the last 5 years. Unreliable fields initially excluded were reintroduced into the series (22 eyes) to create a comparison.**
• RESULTS: **Predicted VFIs were accurate with a mean overestimation of VF deterioration of 1.37% (95% CI: −0.22%, 2.96%). Of the predicted values, 95% were between −4.5% and 5.2% of the observed values when the predicted VFI was ≥90%; and between −13.8% and 20.5% when the predicted VFI was <90%. No statistical difference was found between the reliable and unreliable series (mean difference of 0.09% [95% CI: −0.41%, 0.59%]).**
• CONCLUSION: **The new software for the Humphrey Visual Field Analyzer projects an accurate value for patients when the predicted VFI is ≥90%. Clinicians should consider the limitations of the software, especially for those patients with greater initial VF loss. (Am J Ophthalmol 2013;156:1272–1276. © 2013 by Elsevier Inc. All rights reserved.)**

AS A PROGRESSIVE OPTIC NEUROPATHY, THE PATHOgenesis of glaucoma remains enigmatic. Intraocular pressure (IOP) is one of the only risk factors that can be modified to stabilize this disease. However, IOP levels alone do not dictate management decisions. Visual field status is of paramount importance.

Accepted for publication Jul 25, 2013.

From the Eye Associates, Glaucoma Unit, Sydney, Australia (I.G.); the School of Public Health, The University of Sydney, Sydney, Australia (P.K.); private practice (R.T.).

Inquiries to Renée Talbot, 209-1919 Riverside Drive, Ottawa, Ontario K1H 2A1; e-mail: renee.a.talbot@gmail.com

Developed by Bengtsonn and Heijl[1] for the Humphrey Visual Field Analyzer (HVFA) (Carl Zeiss Meditec, Dublin, California, USA), the Glaucoma Progression Analysis II has, at its core, the visual field index (VFI). This is a global measurement that emphasizes central visual field sensitivities and focal loss as a result of glaucoma. Using a linear regression technique, the software calculates the percentage of vision lost to glaucoma and predicts the stability or progression of visual field loss over the next 5 years based on the results of prior visual field examinations. We studied the accuracy of these predictions in a population diagnosed with glaucoma.

METHODS

THIS WAS A RETROSPECTIVE COHORT STUDY. PATIENTS DID not have direct contact with the investigator, and all data were handled anonymously. All patients had signed a consent form that their clinical data could be utilized for knowledge advancement and research purposes; therefore, the South Eastern Sydney Area Health Service Ethics Committee waived the need for formal approval. The study conformed to the principles of the Helsinki Declaration. At Eye Associates in Sydney, Australia, 834 charts of patients with glaucoma were selected at random. Their visual fields were reviewed and selected according to the following inclusion criteria: patients were required to have had a minimum of 11 consecutive years of follow-up, including at least a yearly automated visual field test performed with the 30-2 or 24-2 Swedish Interactive Threshold Algorithms (SITA) on the HVFA and to not have advanced glaucoma as determined by a mean deviation worse than –20 dB; 42 patients (61 eyes) were included. The majority of the patients' whose charts were reviewed were not selected because the patients lacked the required follow-up period, the visual data were not available for review, or they had advanced glaucoma.

The visual field data over the 11 years of follow-up were divided into the first 5 years and the last 5 years. Unreliable visual fields were removed; fixation losses ≥20%, false-positive rates ≥15%, and false-negative rates ≥33%. If the first or last visual field in a series was unreliable, the patient was excluded, as well as if there was more than 1 unreliable field in the first or last 5 years. The new HVFA software

0002-9394/$36.00
http://dx.doi.org/10.1016/j.ajo.2013.07.025

TABLE. Characteristics of patients with mild to moderate glaucoma, forming the reliable and unreliable data groups in the evaluation of the visual field index

Characteristics	Reliable (n = 42, 61 eyes)	Unreliable (n = 21, 22 eyes)
Mean age	61	61
Mean baseline VA*	0.05	0.12
Mean final VA	0.15	0.15
Mean baseline MD (range)	−3.87 (+1.25 to −19.78)	−4.01 (+1.01 to −19.78)
Mean final MD (range)	−4.94 (+1.50 to −19.89)	−5.13 (+0.88 to −19.89)
Laser (% of patients)	14 (33)	6 (29)
Trabeculectomy (% of patients)	10 (24)	9 (43)
Number of right eyes	33	11
Number of left eyes	28	11
Females (%)	21 (50)	10 (48)
Males (%)	21 (50)	11 (52)

MD = mean deviation in dB; VA = visual acuity.
*All visual acuities are presented in logMAR.

analyzed the data from the fields of the first half of the series so as to generate a prediction. The prediction was then compared to the actual outcome as assessed during the last 5 years. The equation for the Humphrey regression slope is proprietary, so in order to obtain an accurate VFI value (a numerical value that is not provided on the printout), the histograms were enlarged to obtain a scale of 1 mm = 1%. In a masked fashion, each histogram was measured twice for accuracy.

To determine whether the software's predictive capabilities were susceptible to greater fluctuation if unreliable fields were included in the data set, we identified a subgroup of 21 patients (50%) (22 eyes) whose complete series included an unreliable test in either the first or the second half of the series or both. The data from this group were compared with its own more reliable counterpart.

Patients who underwent laser trabeculoplasties (argon or selective), surgical filtering procedures, or both were identified to determine whether such interventions had any effect on the slope of the regression analysis.

RESULTS

ALL ANALYSES WERE CONDUCTED USING SAS 9.1 (SAS INSTITUTE, Cary, North Carolina, USA). Differences between observed and predicted VFIs were calculated for each patient. The accuracy of each predicted value was analyzed by plotting histograms of the differences (data not included), calculating a 95% confidence interval (CI) of the mean difference as well as constructing Bland-Altman plots with 95% limits of agreement. The data obtained from the subgroups were subjected to hypothesis tests and calculation of the CIs to identify any differences among the following groups: (1) reliable versus unreliable; (2) surgical versus nonsurgical; and (3) laser versus no laser.

All statistical analyses, including the limits of agreement, considered the multiple observations provided by 19 patients (45%) by using linear mixed models.[2]

The characteristics of the patients forming the reliable and unreliable data groups are described in the Table. Of the 42 patients included in the study, the majority (22 patients) had primary open-angle glaucoma; 8 patients had ocular hypertension, 6 patients had a diagnosis of chronic primary angle-closure glaucoma, 5 patients had secondary open-angle glaucoma, and 1 patient had normal tension glaucoma. There was a slight, not statistically significant difference between the predicted and observed values for the 61 eyes included in the study. Progression was overestimated, on average, by 1.37% (95% CI [−0.22%, 2.96%], $P = 0.05$), with 95% of the predicted values found to be between 12% lower and 9.3% higher than the observed values (Fig. 1). Within the group, 21 patients (22 eyes) required the removal of unreliable fields from their series prior to evaluation. Those unreliable fields were later reintroduced into their data set and analyzed using the same methods and compared with their own reliable data. There was little alteration in the accuracy of the prediction model, with a mean difference of 0.09% (95% CI [−0.41%, 0.59%], $P = 0.71$) (Fig. 2).

Within the entire group, 14 eyes underwent laser trabeculoplasties and 10 eyes underwent surgical trabeculectomies; 4 eyes underwent both laser treatment and surgery. Given these interventions, it might be expected that the predicted values would overestimate the deterioration of the visual field more extremely than in the eyes that had not undergone such treatment. This was not so for the eyes in the laser-treatment subgroup, but it was the case for the eyes after surgery. The mean difference in the accuracy of prediction was 0.12% (95% CI [−3.88%, 4.13%], $P = 0.94$) for eyes that had undergone laser trabeculoplasties compared with eyes that had not. The mean difference for the surgery subgroup, compared to the group without

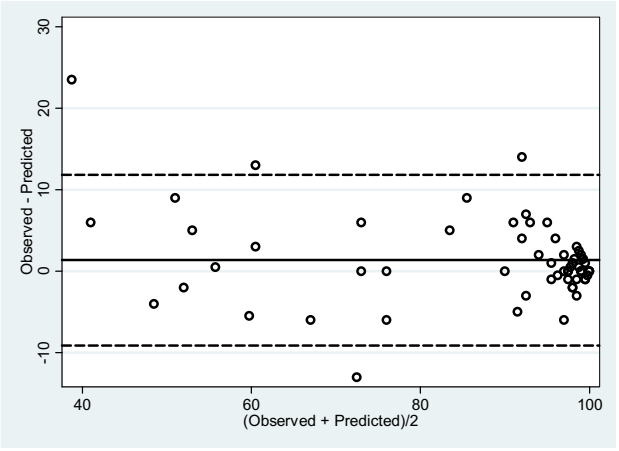

FIGURE 1. Difference between the observed and the predicted visual field index values in the reliable data group of patients with mild to moderate glaucoma. Bland-Altman plot showing the difference between the observed VFI minus the predicted VFI versus the average of the observed and predicted VFI (n = 61). The solid horizontal line is the mean of the differences (y = 1.37). The horizontal dashed lines are the 95% limits of agreement (y = −9.30 and y = 12.03).

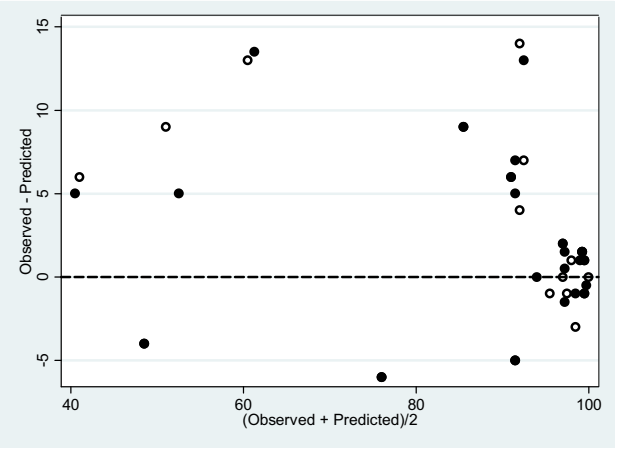

FIGURE 2. Comparison between the unreliable and the reliable data sets in the evaluation of the visual field index in patients with mild to moderate glaucoma. Bland-Altman plot for the predicted VFI using unreliable series (black dots) from 22 eyes, compared with the predicted VFI for the same 22 eyes obtained from their reliable series (white dots).

any surgical intervention, was 4.52% (95% CI [0.21%, 8.84%] $P = 0.02$). The surgery group demonstrated a higher VFI value than that predicted.

The prediction was more accurate if either the estimated slope of the prediction was greater than –1.5%/year, or the predicted VFI was ≥90%, with 95% of the predicted values expected within −7.2% and +7.2% (range, 14.4%), and −4.5% and +5.2% (range, 9.7%) of the actual observed VFI, respectively. This latter prediction is twice as accurate as previously measured with all the data sets included. However, if the slope was more negative than −1.5% per year or the predicted VFI was <90%, the accuracy declined considerably, with 95% of the predicted values for the latter case being within −13.8% and +20.5% of the actual value (range, 34.3%) (Fig. 3).

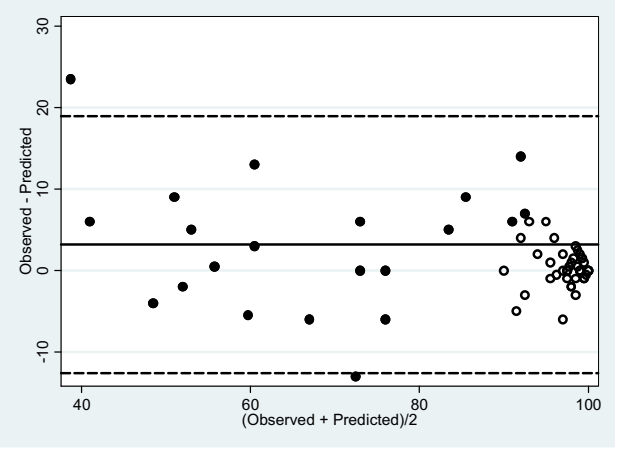

FIGURE 3. Comparing the accuracy of a predicted visual field index value of ≥90% to a value predicted to be <90% in patients with mild to moderate glaucoma. Bland-Altman plot. Black dots represent eyes where the predicted VFI scores are less than 90% (n = 21) and the white dots are those with a predicted value 90% or more (n = 40). The solid horizontal line is the mean of the differences (y = 3.36); the horizontal dashed lines are the 95% limits of agreement (y = −13.8 and y = 20.5) based on only those observations with a predicted VFI less than 90%.

DISCUSSION

BENGTSSON AND HEIJL[1] CREATED THE VFI BY TRANSLATING the below-normal threshold values identified by the pattern-deviation maps into a percentage and calculating the mean of those values. An emphasis is placed on the central visual field, reflecting its importance,[3] with gradual de-emphasis as field eccentricity increases. Pattern-deviation probability plots are used to identify the points below normal sensitivity in eyes with mean deviations better than −20 dB, whereas the total deviation plots are used for the same purpose when the mean deviation is worse than −20 dB, because in more advanced disease; pattern deviation plots underestimate the visual field loss.[4] In our study, we excluded the patients with mean deviations worse than −20 dB; therefore, only the pattern deviation probability plots were used to render predictions. In a study by Rao and associates, the effect of advanced disease on the VFI was evaluated. As the mean deviation crossed the −20dB threshold, a large range in the decline

of the VFI was seen, ranging from 3% to 33%.[5] A similar range of variability was obtained in our study (34.4%) when the predicted VFI was <90%. However, the accuracy of the prediction was excellent when the VFI was ≥90%.

To determine whether unreliable fields disrupted the accuracy of the software, we formed an internal control group; it proved not to be the case. This was important because it would be a considerable undertaking to remove fields that did not meet strict criteria prior to analysis by the regression model. Furthermore, the removal of unreliable fields would prolong the time needed to detect and to quantify progression reliably.[6] Only 1, with a maximum of 2 visual fields, was unreliable in the comparison group. The effects of consistent unreliable fields on the software's ability to predict an accurate value have not been determined.

In addition to reliability indices utilized as exclusion criteria, full-threshold fields were excluded. Although the software was designed specifically for the HVFA SITA 30-2 and 24-2 algorithms,[1] it allows full-threshold fields to be used as baseline and subsequently as follow-up examinations. They were not permitted in the analysis if they were not included as 1 of the 2 baseline fields. This selection criterion limited our candidates to those with at least yearly visual field examinations starting in 1997 because this was the transition period from full-threshold to SITA strategies. Therefore, the baseline examinations in the patients' data sets are not necessarily their first visual fields. Full-threshold examinations allow a greater range of variability at each test location compared with the SITA standard algorithm. The latter may be better able to detect pathologic changes sooner because the former would classify the same reduction in sensitivity as "normal variability."[7] Would full-threshold tests as baseline and SITA standard tests as follow-up visual fields overcall progression? When Budenz and colleagues[8] compared the field defects using SITA Fast, SITA standard, and full-threshold algorithms, visual field defects were larger but shallower with SITA standard compared with full-threshold algorithms; no difference was found between SITA fast and full-threshold for scotoma size. There was a significant difference in the depth of the scotoma between the SITA algorithms and full-threshold algorithms; the latter produced a denser field defect.[8] Care must be taken when comparing results from 2 different algorithms. It may be best to obtain new baseline measurements when changing over to the SITA algorithms from the full-threshold algorithms.[1,8]

All patients had at least 4 visual fields per series (initial 5 years and last 5 years). This is one more field than is required by the software to generate a prediction. The number of data points available is important when linear regression models are used.[3,6,9,10] When more data are available, the forecasted value will be more accurate.[6] Chauhan and colleagues[11] suggested that at least 6 visual fields over a 2-year period are needed to detect rapid progression reliably. In addition, as patient performance may differ for each test; long-term fluctuation must be monitored. Variability is correlated with the number of visual fields required to detect progression; the less the variability, the smaller the number of VFs required to detect change.[3,6,10–12] In our data, linear regression analysis slightly overestimated the visual field progression overall, a discovery Rasker and associates[9] also made when the collective data of their patients with progression and significant linear regression were analyzed. This conclusion was also reached by Bengtsson and associates when analyzing their new software.[13] This overestimation was not important clinically when our total data sets were considered as a whole. However, when patients with greater visual field loss were separated from those with less loss, the software showed less precision for the former group and allowed a larger margin of error; with a range of 34%. Similar limits of precision were found by Bengtsson and colleagues when 70% of their patients' predicted VFIs were within ± 10% of the observed values, and 3% of their patients' predictions differed by ≥50%.[13]

Glaucomatous visual loss does not always progress linearly; episodic progression has been demonstrated.[9,14] A linear regression model may not be the ideal method to predict future change in this group of patients, so perhaps a nonlinear or piecewise-linear regression analysis might suit these types of data better.[15] The glaucoma change probability analysis was shown to be better at detecting episodic progression in a study comparing it with pointwise linear regression.[6] Moreover, linear regression failed to detect progression in 25% of patients in 1 study.[9]

There is no perfect method to predict visual field progression accurately in all patients with glaucoma. It is important to remember that many variables confound the results generated by computer software, especially when human behavior is involved. Patients' adherence to medication, clinic attendance, performance in visual field testing, and patterns of progression cannot be predicted or standardized. The new HVFA software's accuracy is greatest for patients whose visual field indices are not expected to progress below 90%, and the precision declines when the predicted VFI is less. Nonetheless, for many patients, it is a helpful tool when it comes to monitoring glaucoma progression; it is proving to be a powerful visual aid and reference for patients so they can understand the states of their visual fields and any changes over time.

ALL AUTHORS HAVE COMPLETED AND SUBMITTED THE ICMJE FORM FOR DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST, and the following were reported. Funding for our statistical analysis was provided by Glaucoma Australia, which is a registered lay charity. Dr Goldberg has received payment for lectures from Alcon Laboratories and Pfizer; is a consultant for Alcon, Allergan and Forsight Laboratories; has received payment for the development of educational programs from Alcon and Allergan; and is an advisory board member of Alcon, Allergan, Merck, and Pfizer. Dr Kelly

received payment for statistical analysis. Design and conduct of study (R.T., I.G.); Collection of data (R.T.); Management of data (R.T., I.G.); Analysis of data (P.K.); Interpretation of data (R.T., P.K., I.G.); Preparation of manuscript (R.T., I.G.); and Review of manuscript (R.T., I.G.).

REFERENCES

1. Bengtsson B, Heijl A. A visual field index for calculation of glaucoma rate of progression. *Am J Ophthalmol* 2008;145(2):343–353.
2. Carstensen B, Simpson J, Gurrin LC. Statistical models for assessing agreement in method comparison studies with replicate measurements. *Int J Biostat* 2008;4(1):1–26.
3. Heijl A, Bengtsson B, Chauhan BC, et al. A comparison of visual field progression criteria of 3 major glaucoma trials in early manifest glaucoma trial patients. *Ophthalmology* 2008;115(9):1557–1565.
4. Arnalich-Montiel F, Casas-Llera P, Munoz-Negrete F, Rebolleda G. Performance of glaucoma progression analysis software in a glaucoma population. *Graefes Arch Clin Exp Ophthalmol* 2009;274(3):391–397.
5. Rao H, Senthil S, Choudbari N, et al. Behavior of visual field index in advanced glaucoma. *Invest Ophthalmol Vis Sci* 2013;54(1):307–312.
6. Nouri-Mahdavi K, Hoffman D, Ralli M, Caprioli J. Comparison of methods to predict visual field progression in glaucoma. *Arch Ophthalmol* 2007;125(9):1176–1181.
7. Bengtsson B, Heijl A. Comparing significance and magnitude of glaucomatous visual defects using SITA and full threshold strategies. *Acta Ophthalmol Scand* 1999;77(2):143–146.
8. Budenz D, Rhee P, Feuer W, et al. Comparison of glaucomatous visual field defects using standard full threshold and Swedish interactive threshold algorithms. *Arch Ophthalmol* 2002;120(9):1136–1141.
9. Rasker M, van den Enden A, Bakker D, Hoyng P. Rate of visual field loss in progressive glaucoma. *Arch Ophthalmol* 2000;118(4):481–488.
10. Casas-Llera P, Rebolleda G, Munoz-Negrete F, et al. Visual field index rate and event-based glaucoma progression analysis: comparison in a glaucoma population. *Br J Ophthalmol* 2009;93(12):1576–1579.
11. Chauhan BC, Garway-Heath DF, Goni FJ, et al. Practical recommendations for measuring rates of visual field change in glaucoma. *Br J Ophthalmol* 2008;92(4):569–573.
12. Nouri-Mahdavi K, Caprioli J, Coleman A, et al. Pointwise linear regression for evaluation of visual field outcomes and comparison with the advanced glaucoma intervention study methods. *Arch Ophthalmol* 2005;123(2):193–199.
13. Bengtsson B, Patella VM, Heijl A. Prediction of glaucomatous visual field loss by extrapolation of linear trends. *Arch Opthalmol* 2009;127(12):1610–1615.
14. Mikelberg F, Schulzer M, Drance S, Lau W. The rate of progression of scotomas in glaucoma. *Am J Ophthalmol* 1986;101(1):1–6.
15. Montgomery D, Peck E, Vining GG. Regression and model building. In: Bloomfield P, et al., eds. Introduction to Linear Regression Analysis. 3rd edition, New York: Wiley, 2001:1–6.