Comparison of Visual Field Point-Wise Event-Based and Global Trend-Based Analysis for Detecting Glaucomatous Progression

Zhichao Wu¹−⁴ and Felipe A. Medeiros¹,²

¹ Duke Eye Center and Department of Ophthalmology, Duke University School of Medicine, Durham, NC, USA
² Department of Ophthalmology, University of California, San Diego, La Jolla, CA, USA
³ Centre for Eye Research Australia, Royal Victorian Eye and Ear Hospital, East Melbourne, Australia
⁴ Ophthalmology, Department of Surgery, The University of Melbourne, Melbourne, Australia

Correspondence: Felipe A. Medeiros, Duke Eye Center, Duke University, 2351 Erwin Rd, Durham, NC 27705, USA. e-mail: felipe.medeiros@duke.edu

Received: 3 April 2018
Accepted: 25 June 2018
Published: 27 August 2018

Keywords: visual field; glaucoma; progression; longitudinal

Citation: Wu Z, Medeiros FA. Comparison of visual field point-wise event-based and global trend-based analysis for detecting glaucomatous progression. Trans Vis Sci Tech. 2018;7(4):20, https://doi.org/10.1167/tvst.7.4.20

Copyright 2018 The Authors

Purpose: To compare global trend-based and point-wise event-based analysis for detecting visual field progression in eyes with glaucoma.

Methods: The study included a cohort of 367 glaucoma eyes from 265 participants seen over a mean follow-up period of 10 years to develop a computer simulation model of “real-world” visual field results. Progression was evaluated with point-wise event-based analysis using the Guided Progression Analysis (GPA) and global trend-based analysis using the mean deviation (MD) and visual field index (VFI) measures. The specificities of the methods were matched based on a simulated dataset of stable glaucoma eyes, to allow an adequate comparison of their sensitivities for detecting progression.

Results: The 5-year cumulative false-positive rate for the GPA alert of “possible progression” and “likely progression” (significant change from baseline at two and three consecutive visits, respectively) were 34.0% and 7.0%, respectively. At matched specificities, 27.7% eyes were detected as having progressed after 5 years using the GPA “likely progression” criterion, while 24.6% and 23.8% were detected as having progressed using the global trend–based analysis with MD and VFI, respectively. There was a moderate level of agreement between the GPA and global trend-based analyses.

Conclusions: Pointwise event-based and global trend–based methods had similar performances to detect glaucoma progression when rigorously matched for specificity.

Translational Relevance: Although both point-wise event- and global trend–based analyses perform similarly, they could provide complementary information that could be exploited to improve the overall detection of progression in clinical practice and clinical trials.

Introduction

Accurate detection of progression is key to the clinical management of eyes with glaucoma, as it is required for appropriate management decisions to prevent or minimize the risk of functional disability. One of the primary tools used to detect glaucomatous progression is visual field testing on standard automated perimetry (SAP), although accurate detection of progression on SAP remains a challenging task in both clinical practice and research.

Many methods have been developed to detect progressive glaucomatous visual field changes,¹ and can involve either event- or trend-based analyses of global or point-wise parameters; exhaustive reviews of these methods have been published in recent years.²,³ Methods of analyses that are typically used include point-wise event-based analysis, such as the guided progression analysis (GPA) on the Humphrey Field

¹ TVST | 2018 | Vol. 7 | No. 4 | Article 20

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.
short-term test-retest data; however, such data may adversely impact patients.6,7 The specificity of a method can be determined by examining eyes with glaucoma where systematic changes in visual sensitivity (termed “possible progression” on the GPA software).4 A similar observation can also be made from the results from another study,5 but a key limitation of both studies is that the ability for each method to correctly identify stable eyes as nonprogressing (or its specificity) was not evaluated. An increased ability to detect progression (or its sensitivity) for any method can be achieved at the expense of its false-positive detection rates by varying the thresholds used to identify progression, which can result in making improper clinical management decisions that may adversely impact patients.6,7

This problem can be addressed by matching the specificities of the methods when comparing their ability to detect progression. The specificity of a method can be determined by examining eyes with glaucoma where systematic changes in visual sensitivity are absent. This may be achieved through using short-term test-retest data; however, such data may not fully reflect the variability characteristics of longitudinal visual field tests performed clinically. Random re-ordering of longitudinal visual field data can also remove systematic changes in visual sensitivity when evaluating the specificity of trend-based analyses.10 However, this would not be suitable for event-based analysis, because randomly re-ordered tests from progressing glaucoma eyes would have a larger degree of measurement variability (due to the systematic decline in visual sensitivity) compared with eyes that were truly stable, thus prohibiting accurate estimates of specificity.

Instead, we recently developed a computer simulation model that allows “real-world” visual field results from glaucoma eyes to be reconstructed at a point-wise level over time.11 We showed that the simulated visual field results closely reflect the variability characteristics of those seen in the longitudinal clinical cohort it was modelled from. This model thus provides a robust framework for evaluating the sensitivity of the point-wise event- and global trend–based analyses at matched specificities, because longitudinal visual fields that are truly stable can be simulated. We therefore undertook this evaluation in this study to understand the clinical utility of these two approaches for detecting visual field progression.

Methods
Participants

This study included participants who were enrolled in a prospective longitudinal observational study evaluating structural and functional damage in glaucoma. The study received institutional review board approval, and was conducted in adherence with the Declaration of Helsinki and the Health Insurance Portability and Accountability Act. All participants in this study provided written informed consent after the test procedures were explained.

Participants in this study underwent a comprehensive ophthalmologic evaluation that included a review of their medical history, visual acuity measurements, visual field testing, slit-lamp biomicroscopy, ophthalmoscopic examination, gonioscopy, intraocular pressure measurement, and stereoscopic optic disc photography. This study only included eyes considered to have glaucoma, based on the masked evaluation of the optic nerve on stereophotographs.12 This study also included only glaucoma eyes with 10 or more abnormal visual field tests (defined as having a pattern standard deviation [PSD] value with \( P < 0.05 \), or glaucoma hemifield test being outside normal limits) over at least 5 years. Participants were also required to have open angles on gonioscopy, and a best-corrected visual acuity of 20/40 or better, and were excluded if they had any other ocular or systemic disease that could affect the optic nerve or the visual field.

Visual Field Testing

All visual field tests were performed on the Humphrey Field Analyzer II-i (Carl Zeiss Meditec, Inc.) using the Swedish Interactive Thresholding Algorithm Standard 24-2 strategy, with the results being considered unreliable and excluded from the analyses if it had more than 33% fixation losses or false negative errors (with the exception for false negative errors when the visual field mean deviation [MD] was less than −12 dB), or more than 15% false-positive errors. The visual field tests were reviewed for the presence of artifacts including fatigue or learning
effects, inattention, inappropriate fixation, eyelid or rim artifacts, and evidence that visual field results were influenced by a disease other than glaucoma (such as a homonymous hemianopia); tests with such artifacts were not included in the analyses. Visual fields were then repeated if found to be unreliable or contained artifacts.

Computer Simulations of Visual Field Point-Wise Sensitivity

This study used computer simulations to recreate “real-world” visual field results from glaucoma eyes in order to evaluate the point-wise event- and global trend-based analyses. Evaluating these methods using computer simulations is particularly advantageous, because it allows their specificities to be evaluated in a scenario that reflects “real-world” longitudinal follow-up, instead of evaluating it in a short-term test-retest scenario. We have described the details of this simulation model in a previous publication, but describe it briefly below.

As an overview, “real-world” visual field results were simulated by combining “true” point-wise sensitivity estimates with a “noise” component, both of which were derived from the cohort of glaucoma eyes under routine clinical care included in this study. To obtain longitudinal estimates of the “true” point-wise visual field sensitivity for each eye included in this study, a sigmoid regression model was fitted to the measured threshold sensitivities at each location over time using a method described recently. The sigmoid model assumes a nonlinear rate of visual field loss, with natural asymptotes occurring at normal levels of sensitivity and the perimetric floor. The model can be expressed as follows: 

\[ s = \gamma / \left( 1 + e^{x + \beta} \right) \]

where \( s \) denotes the measured sensitivity in decibels, \( \gamma \) indicates the estimate of the initial sensitivity, \( x \) indicates how soon the sigmoid function begins a steep decline, \( \beta \) indicates the steepness of this decline, and \( x \) indicates the time. This regression model was fitted using an iterative feasible generalized nonlinear least squares method (being equivalent to maximum likelihood estimation), except for locations where at least two out of the three initial tests had a measurement of 0 dB, which were fitted with a value of 0 dB throughout the entire duration of the follow-up. An example illustrating two locations that were fitted with this sigmoid regression model over the entire perimetric range is shown in Figure 1. The parameters of the sigmoid model could then be used to estimate “true” sensitivities at each location for an eye at any given time point; these derived sensitivity estimates were termed the “sensitivity template.”

To obtain estimates of measurement “noise,” residuals were derived by subtracting the measured values from those fitted by the sigmoid regression model, and binned according to these fitted values (rounded to the nearest 1 dB). Residuals were pooled across all locations and eyes to generate residual distributions for each fitted sensitivity bin, which were termed the “empirical probability distribution functions” (PDFs). The residuals at each location for each test of each eye were then converted into probabilities based on the empirical PDFs of its fitted sensitivity, providing a standardized estimate of the deviation of the individual’s response from the fitted sensitivity. The probabilities at each location for a test thus provided a template of patient performance, accounting for the correlation between the measured values at each location during a visit (i.e., a global visit effect, such as from varying levels of attention between visits); this was termed a “noise template.” A “noise template” could then be combined with a “sensitivity template” to simulate “real-world” visual field results, as described in detail previously.

In this study, 100 sequences of “real-world” visual field results were simulated for each eye (with each eye having a different pattern of damage and change over time), and each sequence consisted of 12 tests over a 5-year follow-up period, where 2 tests were performed at baseline and one test every 6 months afterward. To assess the specificity of each method for detecting visual field progression, 100 sequences were also simulated for each eye when using the baseline “sensitivity template” for all subsequent tests, to provide a scenario when visual field sensitivities remained stable over time.

Methods for Detecting Visual Field Progression

The simulated visual field tests were evaluated using the GPA (by submitting the simulated results to Carl Zeiss Meditec), a point-wise event-based analysis of visual field progression. Alerts at each visit were raised if three or more test locations showed a change exceeding the test-retest limits expected based on the baseline measurements at two or three consecutive visits, which corresponded to the “possible progression” and “likely progression” alerts.

The simulated visual field tests were also evaluated using global trend-based analysis of the MD and VFI.
In brief, MD is a weighted, age-corrected average of visual sensitivity in decibels. The VFI is also a weighted average of the percentage of normal age-corrected visual function; however, in an attempt to reduce the effects of confounders like cataracts, it incorporates locations in its calculations only when their pattern deviation probabilities are outside normal limits. Visual field progression was considered to have occurred if a statistically significant negative slope was detected at two consecutive visits. The sensitivity and specificity of this method was first evaluated when the level of statistical significance was set at the conventional \( P < 0.05 \). The sensitivities of these methods were then evaluated after their specificities were matched with those by the GPA by changing the \( P \) value used to define statistical significance.

Visual field progression in this study was evaluated in a way that reflects a clinical practice scenario, where the presence of progression is re-assessed at each follow-up visit after a new test has been acquired. Progression was then considered to have occurred at the first time point when the criterion from the point-wise event- or global trend–based analyses have been met.

### Statistical Analysis

The cumulative proportion of eyes identified as having progressed using each method to detect visual field progression was plotted against time to compare the ability of each method with detect progression at matched specificities. Proportional Venn diagrams and kappa (\( \kappa \)) coefficients were used to examine the level of agreement between each method. All analyses were performed using Stata Version 14 (StataCorp, College Station, TX).

### Results

**Participant Characteristics**

A total of 367 eyes from 265 participants with glaucoma were included in this study, and they were on average 62.7 ± 11.3-years old at the first visit (range, 25- to 88 years old) and were seen at 15.6 ± 5.0 visits (range, 10–39 visits) over 10.1 ± 2.5 years (range, 5–17 years). At the first visit, the median (interquartile range) MD and PSD of these eyes was \(-4.04 \text{ dB} (-7.86 \text{ to } -2.17 \text{ dB})\) and \(4.38 \text{ (2.53 to } 8.59 \text{ dB})\), respectively, and was \(-6.37 \text{ dB} (-11.88 \text{ to } -3.37 \text{ dB})\) and \(6.55 \text{ (3.40 to } 10.03 \text{ dB})\) respectively at the last visit.

**Specificity of the Point-Wise Event- and Global Trend–Based Analyses**

For the GPA, the false positive rate of detecting progression after a 5-year follow-up period in glaucoma eyes was 34.0% and 7.0% when requiring three or more visual field locations to have exceeded the test-retest limits at two and three consecutive tests, corresponding to the “possible progression” and “likely progression” flags, respectively. The cumulative false-positive rates for detecting progression over time using these two criteria are shown in Figure 2. Given the high false positive rate using the “possible progression” criterion, the subsequent GPA analyses are all performed using the “likely progression” criterion.

When using the standard criteria of progression for the global trend–based analysis (the presence of a statistically significant slope at \( P < 0.05 \) at two consecutive visits) for MD and VFI, their 5-year cumulative false-positive rate for detecting progression were 4.4% and 5.1%, respectively.
Sensitivity of Progression Detection of the Three Methods

Without matching the specificities of the three methods, the cumulative proportion of glaucoma eyes detected as having progressed after 5 years for the GPA was 27.7%, and for the global trend–based analysis with MD and VFI were 20.9% and 21.4%, respectively.

After matching the specificities of the global trend–based analyses with the GPA, the cumulative proportion of eyes detected as having progressed increased to 24.6% and 23.8% for MD and VFI, respectively, with all three methods performing relatively similarly. A plot of the cumulative proportion of glaucoma eyes that progressed over time is shown in Figure 3, showing how all three methods also performed similarly at detecting progression irrespective of follow-up duration and number of tests performed.

When the three methods were compared at matched specificities, there was a moderate level of agreement for the eyes detected as having progressed after the 5-year follow-up period between the GPA and global trend–based analysis with MD (κ = 0.47) and VFI (κ = 0.50), and substantial agreement between the two global trend–based analyses (κ = 0.74). A proportional Venn diagram illustrating these results is shown in Figure 4.

Impact of Disease Severity on the Ability to Detect Progression

Figure 5 shows the 5-year cumulative percentage of simulated sequences detected as having progressed for glaucoma eyes grouped based on their baseline MD (in 3-dB bins), with no clear patterns observed for the relative effectiveness of each methods to detect progression.

Discussion

This study showed that, when matched by specificity, the performance of point-wise event-based
false-positive rate of progression of 34.0% and 7.0% for the GPA alerts of “possible progression” and “likely progression,” respectively. This level of specificity for the GPA underscores the need for caution when considering whether progression has occurred using the former criterion. These false-positive rates are considerably higher than those in a previous study, with estimates of 18.5% and 2.6% reported when using the “possible progression” and “likely progression” alerts, respectively. However, specificity in that previous study was estimated by requiring participants to perform 12 visual field tests over a 3-month period, assuming that no systematic changes due to glaucoma progression would have occurred over such short period of time. The visual field test results could then be randomly re-ordered and test sequences showing progression based on the GPA were considered false-positives. Although in principle this approach is correct, it ignores the fact that long-term variability in glaucoma patients followed over time is likely to be significantly higher than the short-term variability seen over just a few months. In addition, participants who have such frequent testing over a short period of time are likely to be very experienced test takers (or “perimetry athletes”) and differ significantly from most patients seen in clinical practice. Given the high cumulative false-positive rates associated with the “possible progression” alert, clinicians should exhibit caution when considering progression to have occurred with this criterion.

In this study, we observed a 5-year cumulative
is particularly important because the incorrect identification of progression may lead to inappropriate initiation or intensification of treatment, or even its mere diagnosis can have a negative impact on an individual.6,7

In our study, we found slightly higher sensitivity for detecting progression with the GPA (27.7%) as compared with global trend–based analyses with MD (24.6%) and VFI (23.8%), when all three methods were specificity matched. Previous studies have also observed a similar finding, but reported much larger differences for the improved performance for detecting progression with the GPA.4,5 Those studies accounted for those observations by suggesting that the global methods require many more tests in order to reliably detect progression,9 and that the global methods would missed localized progression detected by the GPA.5 However, these findings were based on using the “possible progression” alert of the GPA (which is characterized by having high false-positive rates). When the “likely progression” criteria for GPA was used instead, trend-based analyses actually performed better than the GPA.4,5 This probably explains why the differences in sensitivity between the methods were lower in our study when we appropriately matched the specificities. In addition, our findings do not support the previous suggestion that trend-based analyses are likely to be more effective when more visual field tests over a longer duration of follow-up are available,4 because we found that all three methods performed similarly across the entire follow-up period when all three methods were specificity matched, highlighting how their effectiveness was not dependent on the number of tests included.

The moderate level of agreement between the methods investigated in our study suggest that there may be valuable information that could be exploited by combining the methods to improve the overall detection of progression. However, care must be taken when using both event- and trend-based analyses in clinical practice, as simply considering progression to have occurred when one of the analyses is flagged as statistically significant can result in increased false-positive rates by virtue of multiple testing. We have recently proposed a Bayesian modeling approach,17 which combines results of event- and trend-based analyses while retaining the desired specificity.

The similarity in performance of both methods for detecting visual field progression also has implications when seeking to determine optimal outcome measures for glaucoma clinical trials. Most landmark glaucoma trials to date have used a form of point-wise event-based analysis18–23 when defining whether visual field progression has occurred at an individual level. However, our findings suggest that global trend–based analysis may be a similarly effective outcome measure, accompanied by the advantage significantly lowering sample size requirements (Wu Z, et al. IOVS. 2017;58:ARVO E-Abstract 2465). In addition, trend-based methods capture information about the rate of change. This is especially valuable because an increasing body of evidence has revealed the importance of the velocity of visual field loss, in addition to the level of loss itself, on functional disability.24–31

In conclusion, this study demonstrated that point-wise event-based and global trend–based methods had similar performances to detect progression when rigorously matched for specificity. The moderate level of agreement between these approaches suggests that information from both methods could be exploited to improve the overall detection of visual field progression in clinical practice and clinical trials.

Acknowledgments

Supported in part by a National Institutes of Health/National Eye Institute grant EY021818 (FAM), National Health and Medical Research Council Early Career Fellowship (#1104985, ZW).

Disclosure: Zhichao Wu, None; Felipe A. Medeiros, Alcon Laboratories, Fort Worth, TX (F, R), Allergan, Irvine, CA (F, C, R), Bausch & Lomb, Garden City, NY (F), Carl Zeiss Meditec, Jena, Germany (F, C, R), Heidelberg Engineering, Heidelberg, Germany (F), Merck, White House Station, NJ (F), National Eye Institute, Bethesda, MD (F), Novartis, Basel, Switzerland (C), Reichert, Dewey, NY (F, R), Topcon, Livermore, CA (F)

References


