# Visual field progression in glaucoma: What is the specificity of the Guided Progression Analysis?

**Short title** Specificity of the Guided Progression Analysis

**Authors**
Paul H. Artes PhD,[1]
Neil O'Leary PhD,[1,2]
Marcelo T Nicolela, MD[1]
Balwantray C. Chauhan PhD,[1]
David P. Crabb, PhD[3]

**Length** 2,200 words, 3 figs (+ 3 in supplementary material), & 2 tables

**Section Code** Glaucoma

**Prev. Pres.** Association for Research in Vision and Ophthalmology meeting 2011

**Keywords** glaucoma, visual field, progression, permutation, Guided Progression Analysis, Glaucoma Change Probability

**Affiliation**
(1) Ophthalmology and Visual Sciences, Faculty of Medicine
Dalhousie University, Halifax, NS, Canada

(2) Biostatistics, Institute of Population Health, Faculty of Medical and Human Sciences, University of Manchester, Manchester, UK

(3) Optometry and Visual Science, City University London, UK

**Correspondence** Paul H. Artes [paul@dal.ca]
Ophthalmology and Visual Sciences, Dalhousie University
Room 2035, West Victoria, 1276 South Park St,
Halifax, Nova Scotia, B3H 2Y9, Canada

**Support** Glaucoma Research Foundation (PHA & DPC)
Grant # MOP-11357, Canadian Institutes for Health Research (BCC)

**Disclosures** None

**Précis** In patients with glaucoma examined many times over a short period of time, we show that the specificity of the Guided Progression Analysis of the Humphrey Field Analyzer varies substantially between patients.

# Abstract

**Purpose:** To estimate the specificity of the Guided Progression Analysis (GPA, Carl Zeiss Meditec, CA), in individual glaucoma patients.

**Design:** Observational cohort study.

**Participants:** Thirty patients with open-angle glaucoma.

**Methods:** In 30 patients with open-angle glaucoma, one eye (median Mean Deviation [MD], -2.5 dB, interquartile range -4.4 to -1.3 dB) was tested 12 times over 3 months (Humphrey Field Analyzer, Carl Zeiss Meditec; SITA Standard, 24-2). "Possible progression" and "likely progression" were determined with the Guided Progression Analysis (GPA). These analyses were repeated after the order of the tests had been randomly re-arranged (1000 unique permutations).

**Main Outcome Measures:** Rate of false-positive alerts of "possible progression" and "likely progression" with the Guided Progression Analysis.

**Results:** On average, the specificity of the GPA "likely progression" alert was high—for the entire sample, the mean rate of false-positive alerts after 10 follow-up tests was 2.6%. With "possible progression", the specificity was considerably lower (false-positive rate, 18.5%). Most importantly, the cumulative rate of false-positive alerts varied substantially among patients, from <1% to 80% with "possible progression", and from <0.1% to 20% with "likely progression". Factors associated with false-positive alerts were visual field variability (standard deviation of MD, Spearman's rho=0.41, $p<0.001$) and the reliability indices (proportion of false-positive and false-negative responses, fixation losses, rho>0.31, $p\leq0.10$).

**Conclusions:** On average, progression criteria currently employed in the GPA have high specificity, but some patients are much more likely to show false-positive alerts than others. This must be considered when the GPA is used in clinical practice, where specificity needs to be controlled for individuals rather than for large groups of patients.

# Introduction

In patients with glaucoma, accurate decisions on visual field progression are a prerequisite of good clinical management.[1, 2] Visual fields have complex properties, and therefore progression is best judged with the help of software such as the Guided Progression Analysis (GPA) of the Humphrey Field Analyser (HFA, Carl Zeiss Meditec, Dublin, CA). [3 4-6]

The GPA compares each test result, point by point, to values from two earlier baseline tests. Points are highlighted on a probability plot if changes exceed the typical measurement variability derived from a group of stable glaucoma patients. If such changes occur at 3 or more points, and in 2 consecutive follow-up tests, the GPA raises an alert of "possible progression"; if they occur in 3 consecutive tests an alert of "likely progression" is raised. Criteria similar to "likely progression" were used in the Early Manifest Glaucoma Trial, and the GPA has subsequently been widely adopted in clinical practice and research.[7, 8] Previous studies have shown that the analysis agrees reasonably closely with the subjective judgement of expert clinicians,[9, 10] and some authors have used the GPA as a reference standard for functional change in glaucoma.[11, 12]

Owing to the fundamental role of visual field progression in the clinical management of glaucoma, it is important to know how often the GPA raises alerts of "possible progression" and "likely progression" in the absence of genuine change, i.e. false-positives. We previously demonstrated that the GPA is likely to have high specificity, on average.[13] However, the analysis is based on a statistical model of typical variability inferred from a group of stable patients—it does not take into account that some patients are more reliable test-takers than others.[14] Because the reproducibility of visual fields varies more than 2-fold between individuals with the same degree of damage (Artes et al, Invest. Ophthalmol. Vis. Sci. 54: E-Abstract 2630), the limits for significant change of the GPA are likely to be too wide for patients who are have relatively low variability, and too narrow for those with relatively larger variability.

In this study, we aim to investigate how the specificity of the GPA varies between individual patients. For this purpose, we tested a group of patients multiple times, over a short period of time during which a clinically meaningful change was unlikely to have taken place.

# Methods

## Patients

Thirty patients were recruited from the glaucoma clinics at the Queen Elizabeth Health Sciences Centre in Halifax, Nova Scotia. Inclusion criteria were a clinical diagnosis of open-angle glaucoma, a Mean Deviation (MD) better than -15.0 dB in at least one eye, absence of ocular or systemic pathology known to reduce visual field sensitivity, and the ability and willingness to participate for 12 consecutive weekly sessions. All patients were experienced with static automated perimetry and had performed at least 5 visual field tests before the study started. They had well controlled levels of intraocular pressure as judged by their physician (MTN). In accordance with the Declaration of Helsinki, the institutional research ethics board approved the protocol, and all patients gave written informed consent.

## Tests

Patients attended 12 weekly sessions over a period of 3 months. During each session, the study eye was examined with program 24-2 SITA-Standard of the HFA.

## Analysis

*Guided Progression Analysis (GPA)*

The GPA is based on principles previously described as Glaucoma Change Probability.[14] In brief, a visual field baseline is calculated from the first two tests, and each subsequent test is then compared, point by point, to this baseline. If the difference in pattern deviation exceeds the retest variability estimated from a group of stable patients, the corresponding location is flagged on a probability map by an open triangle. Half-filled and solid triangles signify change on two or three consecutive follow-up tests, respectively. The GPA gives alerts of “possible progression” and “likely progression” when there are three or more locations with half-filled or solid triangles, respectively.

## *Permutation*

The premise of our study was that a meaningful change was unlikely to have taken place during the short period of 3 months during which the 12 tests were performed. Under this assumption, a GPA alert of "possible progression" or "likely progression" in the series of 12 tests could be regarded as a false-positive event. Furthermore, by assuming that the order of the tests could be treated as arbitrary, a large number of permutations could be generated from the originally observed series, by randomly changing the order of the tests in the sequence. In this way, the probability of observing a false-positive "possible progression" or a "likely progression" alert could be derived specifically for each individual patient. For each patient, we submitted 1000 permuted series to Carl Zeiss Meditec who generated the GPA results as they would appear on the instrument's output.

## *Analysis*

Individually for each patient, we determined the proportion of series in which at least one alert of "possible progression" or "likely progression" had been raised, at the 4$^{th}$ through the 12$^{th}$ test, across the 1000 permuted series. Similarly, we determined the cumulative probability of encountering at least one "possible progression" or "likely progression" alert in a patient's series of 12 tests (2 baseline and 10 follow-up tests). Confidence intervals for the mean proportion of false-positive alerts across the group of patients were determined by bootstrap (n=10,000 samples). We also investigated the association between the cumulative probability of encountering at least one progression alert after 12 tests (2 baseline and 10 follow-up tests) and the MD, the standard deviation (SD) of the MD, and to indices of patient reliability (false-positive and false-negative response errors, fixation losses, averaged across the entire series of 12 tests). All analyses were performed in the open-source programming language R (R Foundation for Statistical Computing, Vienna, Austria; http://www.R-project.org; last accessed 20 January 2014).

# Results

The median age of the patients was 69.1 years (interquartile range [IQR], 64.4 to 70.7 years). Patients had early to moderate visual field damage (median MD, -2.5 dB, IQR -4.4 to -1.3 dB) as illustrated in Figure 1 (available at http://aaojournal.org). All patients were experienced test-takers, and there were no clinically important learning- or practice effects—the mean MD of the 30 patients changed by <0.1 dB between the first and last tests (Fig. 2, (available at http://aaojournal.org). However, the variability of the MD varied by a factor >3 between patients (Fig. 3).

The analysis of the randomly re-ordered test results confirmed that, on average, the specificity of the GPA "likely progression" alert was high—after 10 follow-up tests (12 tests in total, including the 2 baselines), the mean false-positive alert rate across the 30 patients was 2.6% (95% confidence interval: 1.2%, 4.4%). The specificity of the "possible progression" alert was considerably lower—after 10 tests the mean false-positive rate was 18.5% (95% confidence interval: 11.5%, 26.5%) (Figs. 4, 5). Most importantly, however, the false-positive rate of the GPA varied substantially between patients. In 11 patients (37%), no "likely progression" alerts were detected in any of the 1000 reordered series, and 4 of these patients also did not have a "possible progression" alert. On the other hand, in one patient 80% of the reordered series contained alerts of "possible progression", and 18% contained alerts of "likely progression".

"Possible progression" and "likely progression" alerts were more closely associated with the patient reliability indices (false-positive and -negative response errors, fixation losses) and with visual field variability (SD of MD) than with visual field damage as measured with MD and Pattern Standard Deviation (PSD). However, none of these associations were sufficiently strong to predict to a useful level of accuracy in which patients the GPA would be prone to false-positive progression alerts. (Table 1, Figs. 6, 7; available at http://aaojournal.org).

# Discussion

The aim of our study was to investigate the specificity of the Glaucoma Progression Analysis, i.e. the likelihood of encountering a "possible progression" or "likely progression" alert in a series of visual fields in which no meaningful change has taken place. Stable series were established by testing patients frequently over a short period during which disease progression was unlikely, such that any GPA progression alert could be regarded as a false-positive event. Under the assumption that the order of the tests could be randomly exchanged, we were able to estimate the rate of false-positive GPA progression alerts from a large number of random permutations of the original visual field series, for each individual patient.

Our results corroborate earlier reports of high *average* specificity with the GPA[13, 15] — after 12 tests, the average false-positive rate of "likely progression" alerts was <5%. With tests conducted at intervals of 6 months, a series of 12 tests would translate to approximately 5 years of follow-up, and this level of specificity appears sufficiently high for most clinical applications. However, the large variation in the GPA false-positive rates between individual patients confirmed our hypothesis that some patients are much more prone to show false-positive progression alerts than others. The high average specificity of the GPA observed in a group of patients does not apply equally to all patients.

The GPA uses a statistical model to establish, point by point, whether the differences between a follow-up test and two earlier baseline tests exceed the limits of measurement variability typically observed in patients with glaucoma. This model aims to account for the amount of baseline damage at individual test locations, for the location within the visual field, and for the overall damage of the visual field as measured by the MD index.[14] The lack of a relationship between the GPA false-positive rate and visual field damage (as measured by MD and PSD, Fig. 6, available at http://aaojournal.org) indicates that the GPA adequately compensates for the larger threshold variability in damaged areas of the visual field. However, the level of damage explains less than half of the variability in visual field measurements.[16, 17] Clearly, there are patient-related factors unrelated to visual field damage that influence variability, for example the ability to sustain attention and to provide consistent responses. Because the GPA uses the variability estimated from a reference group of patients, the analysis is overly conservative (i.e., highly specific, but less sensitive) in patients who are highly reliable test-takers, and not sufficiently conservative (i.e., more sensitive, but less specific) in patients with relatively larger between-test variability.

While there were statistically significant relationships between overall visual field variability (measured by the SD of the MD), the reliability indices (false-positive and false-negative response errors, and fixation losses), and the likelihood of false-positive GPA progression alerts, these associations were too weak to be practically useful for predicting in which patients the GPA is most likely to produce false-positive progression alerts (Fig. 6, 7; available at http://aaojournal.org).

One alternative to the Glaucoma Change Probability model of the GPA is pointwise linear regression (PLR), a method that has been widely discussed elsewhere.[18-20] PLR establishes statistical significance of change at individual visual field locations by least-squares linear regression of sensitivity (or deviation) over time. Other statistical models for deriving rate of change and its statistical significance at single test locations have also been proposed.[21-23] Common to all of these techniques is that the patient's own variability is estimated, obviating the need to rely on variability estimates from other patients. O'Leary et al. have recently introduced a method (Permutation of Pointwise Linear Regression, PoPLR) in which the statistical significance of deterioration over the entire visual field is derived solely from random re-ordering (permutation) of the individual patient's data, without reference to population-based reference values.[24] This method provides an individualised statistical test of the null hypothesis that there is no negative change at any visual field location, removing any between-patient variation in specificity. We believe that this method may provide a useful alternative to the Glaucoma Change Probability model of the GPA, particularly when more than 5 tests are available for analysis and when specificity needs to be controlled at the level of the individual patient, as it must be in clinical practice.

Two assumptions of our study are a) that visual fields obtained over a short period of time are representative of those obtained over a longer period, and b) that any re-ordered sequence of tests could have occurred with the same likelihood as the originally observed sequence. It is likely that visual field data violate both assumptions. Variability, for example, may be higher in the long term than observed during the 12-week period of our study, and the differences between two tests obtained one after the other may be smaller than between tests at the start and the end of the sequence (serial correlation). However, while these violations may affect our estimates of specificity, they are unlikely to have a substantial effect on the finding that the specificity of the GPA varies considerably between patients.

In summary, we have shown that the GPA criterion of "likely progression" has high specificity on average, but that some patients are much more prone to false-positive alerts than others. Rather than discouraging clinicians from using the GPA, we hope that this report helps to avoid false-positive decisions on progression in patients with larger-than-average variability and frequent response errors.


# Acknowledgements

The authors thank Gary Lee and Mary Durbin of Carl Zeiss Meditec for generating the GPA output for the permuted visual fields.

# References

1. Weinreb RN, Khaw PT. Primary open-angle glaucoma. Lancet 2004;363:1711-20.
2. Hitchings RA. Glaucoma: an area of darkness. Eye (Lond) 2009;23:1764-74.
3. Viswanathan AC, Crabb DP, McNaught AI, et al. Interobserver agreement on visual field progression in glaucoma: a comparison of methods. Br J Ophthalmol 2003;87:726-30.
4. Bengtsson B, Lindgren A, Heijl A, et al. Perimetric probability maps to separate change caused by glaucoma from that caused by cataract. Acta Ophthalmol Scand 1997;75:184-8.
5. Nouri-Mahdavi K, Nassiri N, Giangiacomo A, Caprioli J. Detection of visual field progression in glaucoma with standard achromatic perimetry: a review and practical implications. Graefes Arch Clin Exp Ophthalmol 2011;249:1593-616.
6. Katz J. A comparison of the pattern- and total deviation-based Glaucoma Change Probability programs. Invest Ophthalmol Vis Sci 2000;41:1012-6.
7. Heijl A, Leske MC, Bengtsson B, et al. Measuring visual field progression in the Early Manifest Glaucoma Trial. Acta Ophthalmol Scand 2003;81:286-93.
8. Garway-Heath DF, Lascaratos G, Bunce C, et al. The United Kingdom Glaucoma Treatment Study: a multicenter, randomized, placebo-controlled clinical trial: design and methodology. Ophthalmology 2013;120:68-76.
9. 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;115:1557-65.
10. Tanna AP, Budenz DL, Bandi J, et al. Glaucoma Progression Analysis software compared with expert consensus opinion in the detection of visual field progression in glaucoma. Ophthalmology;119:468-73.
11. Vizzeri G, Bowd C, Weinreb RN, et al. Determinants of agreement between the confocal scanning laser tomograph and standardized assessment of glaucomatous progression. Ophthalmology 2010;117:1953-9.
12. Alencar LM, Zangwill LM, Weinreb RN, et al. Agreement for detecting glaucoma progression with the GDx guided progression analysis, automated perimetry, and optic disc photography. Ophthalmology;117:462-70.
13. Artes PH, Nicolela MT, LeBlanc RP, Chauhan BC. Visual field progression in glaucoma: total versus pattern deviation analyses. Invest Ophthalmol Vis Sci 2005;46:4600-6.
14. Heijl A, Lindgren G, Lindgren A, et al. Extended empirical statistical package for evaluation of single and multiple fields in glaucoma: Statpac 2. Perimetry update 1990;91:303-15.
15. Anton A, Pazos M, Martin B, et al. Glaucoma progression detection: agreement, sensitivity, and specificity of expert visual field evaluation, event analysis, and trend analysis. Eur J Ophthalmol 2013;23:187-95.
16. Wall M, Woodward KR, Doyle CK, Artes PH. Repeatability of automated perimetry: a comparison between standard automated perimetry with stimulus size III and V, matrix, and motion perimetry. Invest Ophthalmol Vis Sci 2009;50:974-9.
17. Henson DB, Chaudry S, Artes PH, et al. Response variability in the visual field: comparison of optic neuritis, glaucoma, ocular hypertension, and normal eyes. Invest Ophthalmol Vis Sci 2000;41:417-21.

18. McNaught AI, Crabb DP, Fitzke FW, Hitchings RA. Visual field progression: Comparison of Humphrey Statpac and pointwise linear regression analysis. Graefe's archive for clinical and experimental ophthalmology 1996;234:411-8.
19. Gardiner SK, Crabb DP. Examination of different pointwise linear regression methods for determining visual field progression. Investigative ophthalmology & visual science 2002;43:1400-7.
20. Nouri-Mahdavi K, Caprioli J, Coleman AL, et al. Pointwise linear regression for evaluation of visual field outcomes and comparison with the advanced glaucoma intervention study methods. Archives of ophthalmology 2005;123:193.
21. Caprioli J, Mock D, Bitrian E, et al. A method to measure and predict rates of regional visual field decay in glaucoma. Invest Ophthalmol Vis Sci 2011;52:4765-73.
22. Russell RA, Crabb DP. On alternative methods for measuring visual field decay: Tobit linear regression. Invest Ophthalmol Vis Sci 2011;52:9539-40.
23. Bryan SR, Vermeer KA, Eilers PH, et al. Robust and Censored Modeling and Prediction of Progression in Glaucomatous Visual Fields. Invest Ophthalmol Vis Sci 2013.
24. O'Leary N, Chauhan BC, Artes PH. Visual Field Progression in Glaucoma: Estimating the Overall Significance of Deterioration with Permutation Analyses of Pointwise Linear Regression (PoPLR). Invest Ophthalmol Vis Sci 2012;53:6776-84.