



Published in final edited form as:

Ophthalmology. 2012 March ; 119(3): 468–473. doi:10.1016/j.ophtha.2011.08.041.

Glaucoma Progression Analysis Software Compared to Expert Consensus Opinion in the Detection of Visual Field Progression in Glaucoma

Angelo P. Tanna¹, Donald L. Budenz^{2,3}, Jagadeesh Bandi^{2,3}, William J. Feuer², Robert M. Feldman⁴, Leon W. Herndon⁵, Douglas J. Rhee⁶, Julia Whiteside-de Vos⁷, Joyce Huang¹, and Douglas R. Anderson²

¹Department of Ophthalmology, Northwestern University Feinberg School of Medicine, Chicago, IL

²Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, Miami, FL

³Department of Epidemiology and Public Health, University of Miami Miller School of Medicine, Miami, FL

⁴Department of Ophthalmology and Visual Science, The University of Texas Medical School at Houston, Houston, TX

⁵Duke University Eye Center, Durham, NC

⁶Massachusetts Eye and Ear Infirmary, Boston, MA

⁷Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

Abstract

Purpose—To compare the results of Glaucoma Progression Analysis (GPA, Carl Zeiss Meditec, Dublin, CA) to subjective expert consensus in the detection of glaucomatous visual field progression.

Design—Retrospective, observational case series.

Participants—100 eyes of 83 glaucoma patients.

© 2011 American Academy of Ophthalmology, Inc. Published by Elsevier Inc. All rights reserved

Corresponding Author: Donald L. Budenz, M.D., M.P.H. Bascom Palmer Eye Institute 900 N.W. 17th St Miami, FL 33136
Telephone: (305) 326-6384 Facsimile: (305) 326-6337 dbudenz@med.miami.edu.

Conflicts of Interest: APT: None

DLB: None

JB: None

WJF: None

RMF: None

LWH: Alcon (C, L), Allergan (L), Ista, Inc (C)

DJR: Alcon (C), Allergan (C), Merck (C), Santen (C)

JW: None

JH: None

DRA: Carl Zeiss Meditec, Inc. (C), Pfizer, Inc (C)

Meeting Presentation: Presented in part at the Association for Research in Vision and Ophthalmology Annual Meeting, Fort Lauderdale, FL, April 2008 and May 2009

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Methods—Five serial Humphrey visual fields from 100 eyes of 83 glaucoma patients were evaluated by five masked glaucoma subspecialists for determination of progression. Four months later, with a randomly reordered patient sequence, the same visual field series were re-evaluated by the same graders, at which time they had access to the Glaucoma Progression Analysis (GPA) printout.

Main Outcome Measures—The level of agreement between majority expert consensus and GPA, both before and after access to GPA data, was assessed using kappa statistics.

Results—On initial review and on re-evaluation with access to the GPA printout, the level of agreement between majority expert consensus and GPA was fair (kappa = 0.52, 95% confidence interval [CI] = 0.35 – 0.69 and 0.62, 95% CI = 0.46 – 0.78, respectively). Expert consensus was more likely to classify a series of fields as showing progression than was GPA ($p \leq 0.002$). There was good agreement between expert consensus on initial review and reevaluation 4 months later (kappa = 0.77, 95% CI = 0.65 – 0.90).

Conclusions—The level of agreement between majority expert consensus of subjective determination of visual field progression and GPA is fair. In cases of disagreement with GPA, the expert consensus classification was usually progression. Access to the results of GPA did not significantly change the level of agreement between expert consensus and the GPA result; however, expert consensus did change in 11 of 100 cases.

INTRODUCTION

The detection of visual field progression is of critical importance in the management of patients with glaucoma. In the Early Manifest Glaucoma Trial (EMGT), for example, in the vast majority of subjects who progressed, the change was recognized only by visual field evaluation, rather than optic disc progression.¹ Despite the clinical importance of detecting visual field progression, there is no accepted standard that defines it. Furthermore, there are a variety of approaches in clinical use that range from subjective review of visual field printouts to event-based and trend-based statistical algorithms.^{2–5}

For the purpose of detecting visual field progression in the EMGT,^{6,7} empiric databases were developed to determine the 95% confidence intervals of the magnitude of fluctuation in pattern deviation of stable glaucoma patients at various locations of the visual field as a function of the severity of baseline damage. In the EMGT, visual field progression was defined as a deterioration exceeding this 95% confidence interval in the pattern deviation at the same ≥ 3 locations on three consecutive visual field tests. This method has become available as commercial software for use with the Humphrey Field Analyzer (HFA) II (Carl Zeiss Meditec, Inc., Dublin, California), called Glaucoma Progression Analysis at the time of this study and subsequently renamed Guided Progression Analysis to distinguish it from the previous glaucoma progression software and to remind the users that diseases other than glaucoma may cause visual field deterioration (GPA, Carl Zeiss Meditec, Inc.).

Although the methodology employed by Glaucoma Progression Analysis (GPA) is very similar to that used in the EMGT, there are only a few published studies that have studied how well the result of this approach matches other methods used for the detection of visual field progression.^{8–11} The purpose of the current study is to compare the GPA analysis of serial visual fields with the majority expert consensus by a group of glaucoma subspecialists, and to determine how frequently subsequent access to the GPA printout results alters the expert consensus regarding the determination of visual field progression.

METHODS

This study was approved by the Human Subjects Committee of the Institutional Review Board (IRB) of the University of Miami-Miller School of Medicine and followed the tenets of the Declaration of Helsinki. This study is in compliance with the Health Insurance Portability and Accountability Act of 1996.

One hundred series of five serial Humphrey visual fields of glaucoma patients managed at the Bascom Palmer Eye Institute were selected by the principal investigator (DLB) such that approximately half of the eyes demonstrated progression based on clinical impression, which included a review of the GPA printout. This resulted in a sufficient number of visual field series that demonstrated progression to allow meaningful statistical analysis of the data.

The vast majority of visual field tests were obtained with the HFA II with use of the Standard version of the Swedish Threshold Interactive Algorithm (SITA-Standard) and the 24-2 pattern of test locations. Of the five hundred visual field tests, approximately 1% had been performed with the Full Threshold strategy, and approximately 1% were performed with the 30-2 pattern of test locations. The first two visual field tests were required to have evidence of damage, defined as the presence of a cluster of ≥ 3 locations depressed at the $P < 5\%$ level, one of which had to be depressed at the $P < 1\%$ level, an abnormal Glaucoma Hemifield Test, or a pattern standard deviation value that was significant at the $P < 5\%$ level.^{12,13}

Expert Review

Each of the visual field series was composed of five reliable visual field tests. Patient names, dates of birth, medical record numbers, and dates of tests were removed from the 100 overview printouts of the five visual fields in each series. The expert readers had no clinical information about the subjects who contributed visual fields, or about the proportion of eyes that were deemed to demonstrate progression. The overview printouts of the visual field series were randomly ordered and sent to five glaucoma subspecialists (RMF, LWH, DJR, APT, JW) who had at least five years of clinical experience after having completed glaucoma fellowships at different institutions. They were masked to the purpose of the study and to the identities of the other readers. Expert readers classified each visual field series using the following options: 'No Progression,' 'Questionable Progression,' 'Probable Progression,' or 'Definite Progression.' The glaucoma specialists were not given any guidelines for evaluating the visual field series for progression. This is referred to as round 1 of expert review.

One month after the initial review, the same visual field series were returned to the readers for repeat assessment. The data from this reading were used to determine intra-observer reproducibility, the subject of another report,¹⁴ and are not included in this manuscript. Then, four months after round 1 of expert review, the overview printouts *and GPA printouts* of the same visual field series were sent to the readers for additional assessment (in the current study, this is referred to as round 2 of expert review). Each time the series of visual fields were returned for reassessment, the visual field series were re-ordered, the graders had been unaware that they were going to receive the fields for re-review, and the graders had been required to return the prior set of visual fields and readings, so that they could not refer to their prior assessments.

Dichotomized determinations of majority expert consensus were made for rounds 1 and 2 of expert review. Visual field series for which ≥ 3 of the 5 expert readers made a determination of 'definite progression' or 'probable progression' were defined to have been classified as 'progression' by majority expert consensus. Eyes for which ≥ 3 expert readers made a

determination of 'no progression' were defined to have been classified as 'no progression' by majority expert consensus. The classification 'questionable progression' was considered to represent uncertainty, and was not counted as a vote for either progression or stability.

The cases in which majority consensus was not achieved in rounds one or two were reviewed by the readers 12 months later using a forced choice system, with the only two choices being progression or no progression. Unresolved cases from round 1 were reviewed first, with only overview printouts available. Subsequently, the same procedure was applied to the unresolved cases from round 2, except that GPA printouts were available.

Glaucoma Progression Analysis

GPA printouts were obtained with software version A10.1/12.6. The GPA methodology is described in detail elsewhere.⁷ The GPA printouts were based only on the 5 fields included in this study for each study eye. Briefly, glaucoma change probability maps (GCPM) were constructed from a database of serial visual fields obtained during a short time period in subjects with stable glaucoma. These were used to define the range of random anticipated fluctuation in the pattern deviation values at each location in the visual field as a function of baseline severity of damage and eccentricity from fixation. The mean pattern deviation value from the first two baseline visual fields at each visual field location is used for comparison against each subsequent visual field test. If on subsequent testing the pattern deviation value at a particular location has deteriorated outside the 95% confidence interval of the GCPM, that location is considered to have progressed.

The location(s) of the progressed points are flagged on the GPA printout with open, half-black, or solid black triangles, based on whether progression from the baseline values had occurred on one, two, or three consecutive follow-up visual field tests, respectively. If the same ≥ 3 locations (not necessarily contiguous or in the same hemifield) have progressed on two or three consecutive visual field tests, the GPA printout indicates that there is "Possible Progression" or "Likely Progression," respectively.

Agreement Between Majority Consensus and GPA

In the current study, the GPA result was dichotomized for the purposes of comparison with majority expert consensus. The GPA classification "Likely Progression" was considered to be in agreement with the majority expert consensus of 'progression.' The absence of a GPA progression message was considered to be in agreement with the majority expert consensus of 'no progression.' Visual field series for which there was agreement between majority consensus and GPA classification are referred to as concordant and those for which there was disagreement are referred to as discordant.

Eyes with a GPA classification of "Possible Progression" on the final examination (the fifth visual field in the series) were excluded from the main analysis because this classification is considered to indicate the need for repeat, confirmatory visual field testing. In the EMGT, the analogous result necessitated repeat visual field testing.

For another analysis, general height adjustment was determined for the first and fifth visual fields of each series by calculating the difference between the total deviation and pattern deviation values for the same visual field testing location. Since the HFA II calculates total deviation and pattern deviation measurements to the tenths of a decibel, but reports those values to the nearest whole decibel, precise values were not obtained.

Statistical Analysis

Kappa statistics and 95% confidence intervals were used to estimate the level of agreement between the dichotomized results for majority expert consensus and GPA. Kappa values were classified into ranges of fair, good, or excellent as suggested by Fleiss.¹⁵ Kappa values were compared with a test by Fleiss.¹⁵ Statistical analyses were performed using SPSS version 18 (SPSS Inc, 2010, Chicago, IL). McNemar's test was used to determine if discordant progression assessments were evenly divided between expert consensus and GPA. To determine if there was an association between change in general height adjustment between the first and fifth visual field and agreement between progression detection methods, bivariate logistic regression was performed using majority consensus as the dependent variable and the GPA result and the change in general height adjustment as the independent variables.

For both round 1 and round 2, the kappa level of agreement between GPA and majority consensus was separately analyzed for the subsets of visual field series with baseline MD values < -10 dB vs. ≥ -10 dB to compare agreement between GPA and expert consensus in fields with and without very severe damage.

RESULTS

Five hundred visual fields from 100 eyes of 83 patients were included in this study. The mean deviation (MD) of the baseline visual fields was -7.39 ± 5.02 , 1.65 to -25.57 dB (mean \pm SD, range). The frequency distribution of the baseline MD is shown in Figure 1.

Majority consensus was initially achieved in 80 of 100 cases (80%) in round 1 and 87 cases (87%) in round 2. The remaining cases were resolved by forced choice re-evaluation. The GPA results of the cases in which majority consensus was not initially achieved are summarized in Table 1. GPA classification was "Possible Progression" in 18 cases. Only three of these were among the cases that required a forced choice re-evaluation in either round 1 or 2.

The agreement of final majority consensus for rounds 1 and 2 of expert review with the GPA classifications of progression are summarized in Table 2. The kappa coefficients of agreement of rounds 1 and 2 with GPA were 0.52 and 0.62, respectively, in the range usually considered "fair" agreement. The attendant increase in kappa when the expert readers were unmasked to the GPA results was not statistically significant ($p=0.43$). For both rounds 1 and 2 the discordant cases were more likely to be identified as progression by expert consensus than by GPA ($p = 0.002$, round 1; $p < 0.001$, round 2). The kappa coefficient of agreement of expert opinion in rounds 1 and 2 was 0.77, usually considered "good to excellent" agreement.

There were 13 cases in which expert consensus was discordant with GPA in both rounds of the study. In all of these cases and in both rounds, majority consensus was 'progression'. There were 11 cases in which expert consensus changed between rounds 1 and 2. In 6 of these cases, majority consensus switched to 'progression' in round 2. In 7 cases, majority consensus became concordant with GPA, in 3 cases, it became discordant, and in one of these 11 cases the GPA classification was "Possible Progression." Among the 18 cases classified by GPA as "Possible Progression," experts classified 15 and 16 of these as demonstrating progression in rounds 1 and 2, respectively.

Bivariate logistic regression analysis showed that change in general height between the first and fifth visual field test of each series approached but did not achieve statistical significance in predicting majority consensus ($p=0.12$ for round 1 and $p=0.071$ for round 2).

Its inclusion in the model with GPA did not influence the statistical significance of the association between expert consensus and GPA ($p < 0.001$ for both round 1 and 2).

For both round 1 and round 2, the kappa level of agreement between GPA and majority consensus was lower in the subsets of visual field series with baseline MD values < -10 dB compared to those with MD ≥ -10 dB (Table 3). The small sample size of the group with very severely affected visual fields precludes determination of whether there is less agreement between GPA and experts in this subset than in fields with better MD.

DISCUSSION

The identification of glaucoma progression is a critical step in the management of the disease and presents a diagnostic challenge. Progression can be detected on the basis of worsening optic disc excavation or retinal nerve fiber layer atrophy or on the basis of deterioration in visual function, usually as assessed by standard automated perimetry

One commonly utilized approach for the detection of visual field progression, in clinical practice, appears to be subjective assessment of serial visual field printouts. The magnitude of anticipated variation in measured values of threshold sensitivity at each location in the field depends on a number of factors, including the baseline severity of damage and the eccentricity of the location in question, making the subjective determination of progressive worsening at a location or overall an important clinical challenge.^{16–20} Furthermore, threshold sensitivity declines as a function of increasing media opacity, even in the absence of worsening glaucomatous injury.^{21–24} Given these complexities, it seems likely that the use of statistical techniques may be able to account for these various relevant influences that would be difficult for a clinician to take into account intuitively when attempting to detect or quantify progression.

One of the event-based techniques, GPA, is now readily available for routine clinical use. Although the same principles were utilized for the detection of progression in the EMGT, and the statistical principles are sound, there have been few studies published that have compared GPA to other approaches for the detection of visual field progression,^{8–11} and in particular, compared to the interpretation of experienced clinicians.

We compared the results of GPA with the subjective evaluation of serial visual fields for the detection of progression by a group of experienced glaucoma experts. We also explored the extent to which the ultimate assessment by these experienced specialists was influenced by the availability of the GPA results themselves.

We observed fair agreement between majority expert consensus and GPA in both the rounds of the study. In only 11 cases (11%) did the majority consensus differ between the two rounds of the study. In those cases, consensus shifted to agreement with GPA in 7 cases.

Among the cases in which GPA did not agree with the expert consensus (discordant cases), majority expert consensus was usually 'progression', with GPA detecting no progression. There were 13 cases in which expert consensus remained discordant with GPA in both rounds of the study. In *all* of those cases, majority consensus was 'progression'. Furthermore, among the discordant cases in round 1, majority consensus usually did not change in round 2, with majority consensus usually declaring 'progression' in the face of a GPA classification of no progression.

In order to minimize the impact of progressive cataract as a confounding factor in the determination of progression, GPA has been designed to utilize pattern deviation data instead of threshold sensitivity or total deviation measurements. In some cases of glaucoma

progression, generalized loss of sensitivity can occur due to the glaucomatous disease process itself. In such cases, GPA would not detect progression, but expert observers may be able to. In this study, we found that, although there was a tendency toward significance, the association between change in general height adjustment between the first and fifth visual field and agreement between progression detection methods did not reach the level of significance.

Another potential problem with GPA occurs in areas with already severe damage, where a floor effect exists for pattern deviation values. However, while kappa values assessing agreement between expert consensus and GPA were smaller in the subset of fields with baseline MD < -10 dB, the differences did not achieve statistical significance.

The GPA classification “Possible Progression” implies a level of uncertainty regarding the stability of the visual fields that, in a clinical setting, would usually result in closer-than-usual monitoring of the visual field prior to a decision whether to intensify treatment. Similarly, in the EMGT, such a visual field result necessitated repeat, confirmatory visual field evaluation prior to declaration that a subject progressed. Accordingly, the 18 visual field series for which the final GPA result was “Possible Progression” were excluded from the main analysis of the current study.

Among the eyes classified by GPA as “Possible Progression,” most were judged to have progressed by majority consensus, again suggesting that the GPA software algorithm was more conservative about declaring progression than the majority of experts in this study. Because there is no external standard, it is unknown whether the readers were able to correctly detect progression before GPA verification on three consecutive visual field tests or if they prematurely and inaccurately judged progression in some eyes.

Previous studies comparing methods for judging glaucoma progression have found disappointing levels of agreement. Katz and colleagues performed a study in which they used criteria from the Advanced Glaucoma Intervention Study (AGIS), the Collaborative Initial Glaucoma Treatment Study (CIGTS), and the EMGT to determine which of 67 eyes of 56 patients followed for 10 years demonstrated visual field progression.²⁵ They found that the AGIS criteria determined that progression had occurred in 11% of eyes compared to 22% with CIGTS and 23% with EMGT. However, only half of the eyes identified as having progressed by one set of criteria were identified by either of the others, and there were only five eyes that demonstrated visual field progression by all three sets of criteria.

The visual field index (VFI) is an index of the severity of visual field loss that relies on both the pattern deviation probability maps and total deviation data and assigns greater weight to more central locations in the visual field.²⁶ A previous study that compared GPA with linear regression analysis of MD and of VFI demonstrated that GPA classified more visual field series as having progressed compared to the latter two trend analyses.⁸ The issue of whether the GPA event analysis is more sensitive than the VFI trend analysis or represents false positive determinations of progression is unclear.

With nonparametric progression analysis (NPA) visual field progression is defined on the basis of a reproducible deterioration in the MD.^{9,27} A prospective study that compared GPA with NPA demonstrated that the level of agreement between the two approaches was moderate (kappa = 0.50), with the NPA identifying more visual field series as having progressed.⁹

Another trend-based approach for progression detection, Threshold Noiseless Trend (TNT), was compared to GPA and glaucoma experts in visual field series from 56 eyes, only 6 of which were classified by GPA as showing “Likely Progression.”¹⁰ The study demonstrated

slight to moderate agreement between experts and GPA; however, there was stronger agreement between TNT and experts.

In another study, GPA results were compared to the findings of an expert observer who utilized both the GPA result itself and a series of defined criteria for the diagnosis of progression.¹¹ The kappa index of agreement between the two approaches was 0.87 when 2 consecutive tests were required to demonstrate progression by GPA (“Possible Progression”) and 0.64 when 3 consecutive tests were required (“Likely Progression”). This study utilized only a single reader who had access to the GPA results and was guided by the objectively applied statistical criteria of the GPA. Since there is no external standard, it is difficult to draw a comparison to the present and the above-mentioned studies.

Lester, et al. reported agreement among nine expert readers when they had access only to overview printouts and, on separate occasions after reordering of the 38 visual field series, only GPA 1 and GPA 2 printouts (without overview printouts and their total deviation maps).²⁸ They reported a median kappa of 0.58 for intra-observer agreement between the session when only the overview printout was available and the session when only the GPA 1 printout was available. They did not report the level of agreement among the expert readers' opinions and the GPA result itself.

In conclusion, the current study demonstrates fair agreement between GPA and majority consensus by a group of experienced glaucoma subspecialists. There is no established independent standard for judging glaucoma progression, so the current study does not suggest that GPA is better or worse than the expert readers. There were some cases in which majority expert consensus disagreed with GPA, typically in the direction that experienced specialists designated progression when the statistical analysis of GPA did not. Further research is required to identify the reasons for such disagreement and to determine whether expert readers are correctly detecting progression prior to GPA in such cases.

Acknowledgments

Financial Support: Supported by an unrestricted grant from Research to Prevent Blindness (APT, DLB, JB, WJF, DRA); NEI grant P30 EY-014801 (DLB, JB, WJF, DRA); an unrestricted grant from Carl Zeiss Meditec, Inc; and National Institutes of Health, National Eye Institute Training Grant EY 07127, Clinical Trials Training Program in Vision Research (JW)

REFERENCES

1. Heijl A, Leske MC, Bengtsson B, et al. Early Manifest Glaucoma Trial Group. Reduction of intraocular pressure and glaucoma progression: results from the Early Manifest Glaucoma Trial. *Arch Ophthalmol*. 2002; 120:1268–79. [PubMed: 12365904]
2. Viswanathan AC, Fitzke FW, Hitchings RA. Early detection of visual field progression in glaucoma: a comparison of PROGRESSOR and STATPAC 2. *Br J Ophthalmol*. 1997; 81:1037–42. [PubMed: 9497460]
3. Fitzke FW, Hitchings RA, Poinosawmy D, et al. Analysis of visual field progression in glaucoma. *Br J Ophthalmol*. 1996; 80:40–8. [PubMed: 8664231]
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. [PubMed: 9197570]
5. Gardiner SK, Crabb DP. Examination of different pointwise linear regression methods for determining visual field progression. *Invest Ophthalmol Vis Sci*. 2002; 43:1400–7. [PubMed: 11980853]
6. Leske MC, Heijl A, Hyman L, Bengtsson B, Early Manifest Glaucoma Trial Group. Early Manifest Glaucoma Trial: design and baseline data. *Ophthalmology*. 1999; 106:2144–53. [PubMed: 10571351]

7. Heijl A, Leske MC, Bengtsson B, et al. EMGT Group. Measuring visual field progression in the Early Manifest Glaucoma Trial. *Acta Ophthalmol Scand*. 2003; 81:286–93. [PubMed: 12780410]
8. Casas-Llera P, Rebolleda G, Munoz-Negrete FJ, et al. Visual field index rate and event-based glaucoma progression analysis: comparison in a glaucoma population. *Br J Ophthalmol*. 2009; 93:1576–9. [PubMed: 19535357]
9. Wesselink C, Heeg GP, Jansonius NM. Glaucoma monitoring in a clinical setting: glaucoma progression analysis vs nonparametric progression analysis in the Groningen Longitudinal Glaucoma Study. *Arch Ophthalmol*. 2009; 127:270–4. [PubMed: 19273789]
10. Diaz-Aleman VT, Anton A, de la Rosa MG, et al. Detection of visual-field deterioration by Glaucoma Progression Analysis and Threshold Noiseless Trend programs. *Br J Ophthalmol*. 2009; 93:322–8. [PubMed: 18621790]
11. Arnalich-Montiel F, Casas-Llera P, Munoz-Negrete FJ, Rebolleda G. Performance of glaucoma progression analysis software in a glaucoma population. *Graefes Arch Clin Exp Ophthalmol*. 2009; 247:391–7. [PubMed: 18982343]
12. Katz J, Sommer A, Gaasterland DE, Anderson DR. Comparison of analytic algorithms for detecting glaucomatous visual field loss. *Arch Ophthalmol*. 1991; 109:1684–9. [PubMed: 1841576]
13. Anderson, DR. Automated Static Perimetry. Mosby Year Book; St. Louis, MO: 1992. p. 123
14. Tanna AP, Bandi JR, Budenz DL, et al. Interobserver agreement and intraobserver reproducibility of the subjective determination of glaucomatous visual field progression. *Ophthalmology*. 2011; 118:60–5. [PubMed: 20723992]
15. Fleiss, JL. Statistical Methods for Rates and Proportions. 2nd ed.. Wiley; New York: 1981. p. 222-3.
16. Heijl A, Lindgren G, Olsson J. Normal variability of static perimetric threshold values across the central visual field. *Arch Ophthalmol*. 1987; 105:1544–9. [PubMed: 3675288]
17. Werner EB, Petrig B, Krupin T, Bishop KI. Variability of automated visual fields in clinically stable glaucoma patients. *Invest Ophthalmol Vis Sci*. 1989; 30:1083–9. [PubMed: 2732022]
18. Lee AC, Sample PA, Blumenthal EZ, et al. Infrequent confirmation of visual field progression. *Ophthalmology*. 2002; 109:1059–65. [PubMed: 12045044]
19. Schulzer M. Errors in the diagnosis of visual field progression in normal-tension glaucoma. *Ophthalmology*. 1994; 101:1589–94. [PubMed: 8090461]
20. Nouri-Mahdavi K, Brigatti L, Weitzman M, Caprioli J. Comparison of methods to detect visual field progression in glaucoma. *Ophthalmology*. 1997; 104:1228–36. [PubMed: 9261308]
21. Lam BL, Alward WL, Kolder HE. Effect of cataract on automated perimetry. *Ophthalmology*. 1991; 98:1066–70. [PubMed: 1891215]
22. Smith SD, Katz J, Quigley HA. Effect of cataract extraction on the results of automated perimetry in glaucoma. *Arch Ophthalmol*. 1997; 115:1515–9. [PubMed: 9400784]
23. Chen PP, Budenz DL. The effects of cataract extraction on the visual field of eyes with chronic open-angle glaucoma. *Am J Ophthalmol*. 1998; 125:325–33. [PubMed: 9512149]
24. Hayashi K, Hayashi H, Nakao F, Hayashi F. Influence of cataract surgery on automated perimetry in patients with glaucoma. *Am J Ophthalmol*. 2001; 132:41–6. [PubMed: 11438052]
25. Katz J, Congdon N, Friedman DS. Methodological variations in estimating apparent progressive visual field loss in clinical trials of glaucoma treatment. *Arch Ophthalmol*. 1999; 117:1137–42. [PubMed: 10496384]
26. Bengtsson B, Heijl A. A visual field index for calculation of glaucoma rate of progression. *Am J Ophthalmol*. 2008; 145:343–53. [PubMed: 18078852]
27. Jansonius NM. Bayes' theorem applied to perimetric progression detection in glaucoma: from specificity to positive predictive value. *Graefes Arch Clin Exp Ophthalmol*. 2005; 243:433–7. [PubMed: 15578199]
28. Iester M, Capris E, De Feo F, et al. Agreement to detect glaucomatous visual field progression by using three different methods: a multicentre study. *Br J Ophthalmol*. In press.

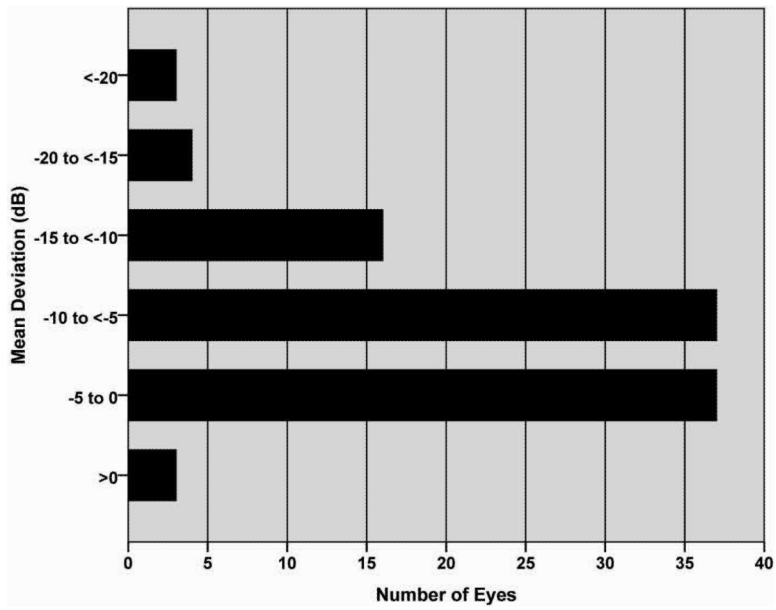


FIGURE 1. Frequency distribution of mean deviations of the first baseline visual field of the 100 study eyes.

Table 1

Glaucoma Progression Analysis results for visual field series for which majority consensus was not initially achieved.

	Failure to achieve initial majority consensus	
	Round 1	Round 2
N	20	13
GPA Progression	3	0
GPA No Progression	15	12
GPA Possible Progression	2	1

GPA = Glaucoma Progression Analysis.

Table 2

Agreement between majority expert consensus and Glaucoma Progression Analysis and reproducibility of expert consensus between rounds 1 and 2.

A. Agreement of Round 1 (expert consensus masked to GPA result) with GPA (excludes 18 cases for GPA result was possible progression)			
		GPA result	
		Progressed	Not Progressed
Expert Consensus	Progressed	27 (33%)	17 (21%)
	Not Progressed	3 (4%)	35 (43%)
Percent agreement = 76%			
Kappa = 0.52 (95% CI: 0.35 – 0.69)			

B. Agreement of Round 2 (expert consensus unmasked to GPA result) with GPA (excludes 18 cases for GPA result was possible progression)			
		GPA result	
		Progressed	Not Progressed
Expert Consensus	Progressed	29 (35%)	15 (18%)
	Not Progressed	1 (1%)	37 (45%)
Percent agreement = 81%			
Kappa = 0.62 (95% CI: 0.46 – 0.78)			

C. Agreement of Round 1 (expert opinion masked to GPA result) with Round 2 (expert opinion unmasked to GPA result)			
		Consensus when unmasked to GPA	
		Progressed	Not Progressed
Consensus when masked to GPA	Progressed	54 (54%)	5 (5%)
	Not Progressed	6 (6%)	35 (35%)
Percent agreement = 89%			
Kappa = 0.77 (95% CI: 0.64 – 0.90)			

GPA = Glaucoma Progression Analysis; CI = confidence interval.

Table 3

Agreement between majority consensus and Glaucoma Progression Analysis for rounds 1 and 2 in eyes with baseline MD < -10 dB vs. \geq -10 dB.

Baseline MD	N	Kappa (SE)	
		Round 1	Round 2
< -10 dB	21	0.36 (0.17)	0.43 (0.18)
\geq -10 dB	61	0.55 (0.10)	0.65 (0.09)
P-value		0.349	0.281

GPA = Glaucoma Progression Analysis; MD = mean deviation; SE = standard error.