ONLINE FIRST

Influence of Visual Field Testing Frequency on Detection of Glaucoma Progression With Trend Analyses

Kouros Nouri-Mahdavi, MD, MSc; Reza Zarei, MD; Joseph Caprioli, MD

Objective: To explore whether increased frequency of visual field testing leads to earlier detection of glaucoma progression with trend analyses.

Methods: The visual fields of 468 eyes (381 patients) from the Advanced Glaucoma Intervention Study with 10 or more reliable visual field tests and 3 or more years of follow-up were studied. Starting at year 1, every other visual field examination was deleted to create a low-frequency data set, and the original group was kept as the high-frequency data set. The proportion of progressing eyes and the time to progression were compared between the 2 data sets with global and pointwise linear regression criteria.

Results: The median number of visual field examinations was 20 and 12 for the high- and low-frequency data sets, respectively. Based on primary mean deviation criteria, 204 eyes (43.6%) in the high-frequency data set and 160 eyes (34.2%) in the low-frequency data set progressed ($P<.001$), whereas 185 eyes (39.5%) in the high-frequency data set and 167 eyes (35.7%) in the low-frequency data set progressed according to pointwise linear regression ($P=.02$). The high-frequency data set was more likely to detect progression with mean deviation (hazard ratio [HR], 1.69 [95% confidence interval {CI}, 1.36-2.10]) or pointwise linear regression criteria (HR, 1.52 [95% CI, 1.21-1.90]). A similar number of improving eyes were detected with mean deviation criteria (HR, 0.95 [95% CI, 0.58-1.60]), but pointwise linear regression criteria were more likely to detect improvement in the high-frequency data set (HR, 2.27 [95% CI, 1.43-3.62]). The results did not significantly change after censoring data at 5 years.

Conclusions: Increasing the frequency of visual field testing leads to earlier detection of glaucoma progression, especially with global trend analyses. This finding has significant implications for the care of patients with glaucoma.

Arch Ophthalmol. 2011;129(12):1521-1527.
Published online August 8, 2011.
doi:10.1001/archophthalmol.2011.224

Author Affiliations: Glaucoma Division, Jules Stein Eye Institute, UCLA (University of California, Los Angeles).

VISUAL FIELD TESTING REmains one of the main methods to monitor patients with glaucoma for progression. Standard achromatic automated perimetry provides a large amount of quantitative data amenable to various statistical analyses for the detection of glaucoma progression over time. Visual field outcomes have classically been the principal outcome used for detection of glaucoma progression in clinical trials.[1-4] One of the issues with regard to visual field testing is the need to repeat tests to have adequate statistical power to detect significant glaucoma progression rates.[5,6] Estimating rates of progression can help identify patients who are progressing at a faster pace and are therefore at serious risk of developing visual disability during their lifetime. These patients can be subjected to more aggressive treatment or more frequent follow-up. However, not all patients need frequent visual field testing, and in only a minority of patients, glaucoma progresses so fast to warrant frequent testing.[7] Based on theoretical considerations, it has been suggested that a large number of visual field examinations are necessary to detect smaller, yet clinically significant, progression rates, especially in eyes demonstrating a high degree of long-term fluctuation.[8] Also, it has been suggested, based on simulation data, that a minimum of 3 examinations per year are required for achieving optimal sensi-

CME available online at www.jamaarchivescme.com and questions on page 1520

tivity and specificity for detection of clinically significant rates of progression according to pointwise linear regression (PLR) analysis.[9] However, there is no evidence based on real patient data in the literature to support such recommendations. The goal of our investigation is to report the performance of global and pointwise trend analyses for the detection of progression in a subset of patients from the Advanced Glaucoma Intervention Study (AGIS) under 2 different testing scenarios: (1) when all the visual field data are included compared with (2) when only half of the follow-up visual fields are included. We hypothesized that glaucoma progression would be detected earlier or in a larger proportion of patients with the more frequent testing schedule.

METHODS

The AGIS design and methods are described in detail elsewhere[10] and are summarized herein. Phakic patients aged 35 to 80 years with primary open-angle glaucoma no longer controlled by maximally tolerated medical treatment were recruited in AGIS. Between 1988 and 1992, investigators at 12 participating AGIS clinical centers enrolled 789 eyes of 591 patients. Our study's data are based on a database closure of March 31, 2001. The institutional review board at UCLA approved our study, and the research was performed in accordance with the tenets of the Declaration of Helsinki.

A total of 468 eyes of 381 patients from AGIS[10] who met the following criteria were selected for our study: a baseline visual field AGIS score of 16 or less, 3 years or more of follow-up, and 10 or more visual field examinations with a reliability score of 2 or less. Visual field tests were performed with a Humphrey Visual Field Analyzer I (Carl Zeiss Meditec Inc, Dublin, California) set for the central 24-2 threshold test, size III white stimulus, and full threshold strategy. Visual field measurements (a minimum of 2) were originally made at baseline, 3 months after initial intervention, and at each 6-month follow-up examination and were repeated when clinically indicated during the follow-up period.

For the purpose of our study, the last available visual field test performed before the first surgical procedure was used as the baseline to reduce learning effects. The high-frequency data set refers to the original group with the entire series of visual field tests after exclusion of all but the last preintervention visual field tests. The low-frequency data set refers to the group in which nearly half of the visual field tests were deleted to create an approximately yearly visual field testing frequency. For both high-frequency and low-frequency data sets, all the visual field tests performed within the first year of follow-up (usually baseline, 3-month, 6-month, and 1-year visual field tests) and the very last available visual field tests were included. Starting from the second year of follow-up, every other visual field test in the original database was deleted to allow the low-frequency data set to have about half as many visual field tests as the high-frequency data set performed at roughly yearly intervals for the remainder of the follow-up.

OUTCOMES

Mean Deviation Slopes

Visual field mean deviation (MD) was regressed against time starting at the first postintervention year. Sequential regression analyses were performed yearly with the cutoff for inclusion of visual field tests set at 1.2 years, 2.2 years, and so on to allow inclusion of the visual field tests performed at scheduled yearly examinations. Progression at time t was defined as the presence of a negative MD slope with $P < .05$ confirmed at least once. Bland-Altman plots were used to compare concordance of the 2 data sets with regard to MD slopes.[11] Kaplan-Meier survival curves derived from Cox proportional hazard models were also compared between the 2 data sets. Time to progression was defined as the time to the first significant negative slope ($P < .05$) that was confirmed at least once (primary outcome). Those who never progressed during follow-up were considered censored observations. Because there is no external standard for comparing the performance of the 2 data sets, rates of improvement in the 2 data sets were also compared. To this aim, eyes that had a confirmed statistically significant ($P < .05$) positive slope were considered to be improving during follow-up. As an alternative definition for change based on MD, we also explored the 2-omitting criteria as described by Gardiner and Crabb.[12]

PLR Criteria

Pointwise linear regression analysis was similarly performed starting at the 1-year follow-up period and yearly thereafter. The criteria for the definition of worsening at a single test location by PLR consisted of a regression slope of less than or equal to –1.0 dB/y along with $P \leq .01$. These criteria have been generally accepted and used in the literature for the definition of visual field progression.[13] Test locations representing the blind spot were excluded from such analysis. At least 3 test locations were required to be progressing in a visual field series with 1 confirmation before that series was considered to be worsening. The proportion of progressing eyes and the number of test locations demonstrating significant negative slopes were compared in the 2 data sets. Similar criteria were applied to define improvement with PLR in the 2 data sets (pointwise slope $\geq$1.0 dB/y along with $P \leq .01$).

We compared the error standard deviations for the regression analyses, as a measure of the amount of longitudinal noise, and the standard error for the MD or PLR slopes, as a measure of the reliability of the slope estimates, between the 2 data sets. For all regression criteria, 2-sided hypothesis testing was performed.

We also compared the results of regression analyses with those of the AGIS criteria (in the high-frequency data set) for detection of progression. A change of more than 3 points in the AGIS score with 2 confirmations was considered evidence for progression.

STATISTICAL METHODS

A Cox proportional hazard model using the frailty option of the *stcox* procedure in the Stata statistical software (version 11.0; StataCorp LP, College Station, Texas) was used to compare time to progression in the 2 data sets. The frailty option was used to account for the paired analyses. It reflects the random eye effect (ie, the 2 measurements from each eye were considered as repeated measurements of the same eye). Only 23% of the patients had both eyes included in the study; therefore, no correction for inclusion of both eyes of some patients was performed. Qualitative evaluation of observed vs expected survival plots and the Schonfeld's goodness-of-fit test were used to test for the assumption of proportionality of hazard rates. The paired t test and the Wilcoxon rank sum test were used to compare normally distributed and nongaussian data, respectively. The McNemar test was used to compare paired proportions. Cohen κ statistic was used to measure the agreement of the regression analyses with the AGIS method.

Table 1. Baseline Characteristics of the Study Cohort of 381 Patients (468 Eyes) From the Advanced Glaucoma Intervention Study

Characteristic	Patients or Eyes
Age at baseline, mean (SD), y	72.6 (10.1)
Eye laterality, No. of eyes (%)	
Right	218 (46.6)
Left	250 (53.4)
Sex, No. of eyes (%)	
Male	196 (41.8)
Female	272 (58.2)
Race, No. of eyes (%)	
White	196 (41.9)
African American	264 (56.5)
Hispanic	8 (1.8)
Visual field examinations, median (range), No.	
High frequency	20 (10-32)
Low frequency	12 (6-19)
Baseline visual field MD, mean (SD), dB	–10.7 (5.4)
Baseline visual field PSD, mean (SD), dB	7.1 (3.0)
Cataract surgery, No. of eyes (%)	
Yes	187 (39.8)
No	281 (60.2)

Abbreviations: MD, mean deviation; PSD, pattern standard deviation.

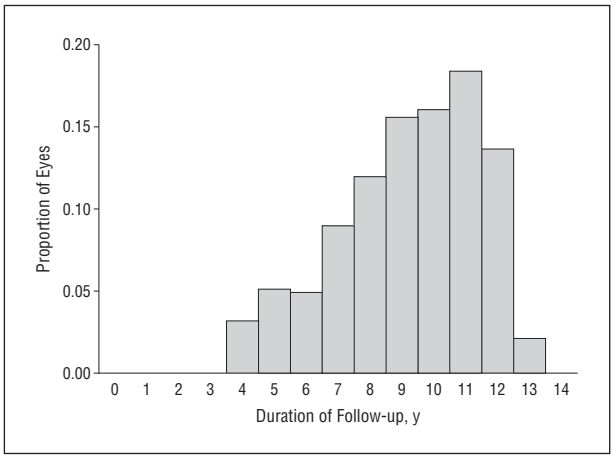

Figure 1. Distribution of follow-up times in the study cohort, which consisted of 468 eyes (381 patients) from the Advanced Glaucoma Intervention Study with 10 or more reliable visual field tests and 3 or more years of follow-up.

RESULTS

A total of 468 eyes from 381 patients were included. **Table 1** describes the baseline characteristics of the study cohort. The median follow-up period was 9.0 years (range, 3.2-13.0 years) (**Figure 1**). The median number of visual field examinations was 20 (range, 10-32) for the high-frequency data set and 12 (range, 6-19) for the low-frequency data set. The mean (SD) baseline MD was –10.7 (5.4) dB. **Figure 2** demonstrates the frequency distribution of the time intervals between sequential visual field examinations in the 2 data sets. The median time interval was 6.0 months in the high-frequency data set and 11.3 months in the low-frequency data set.

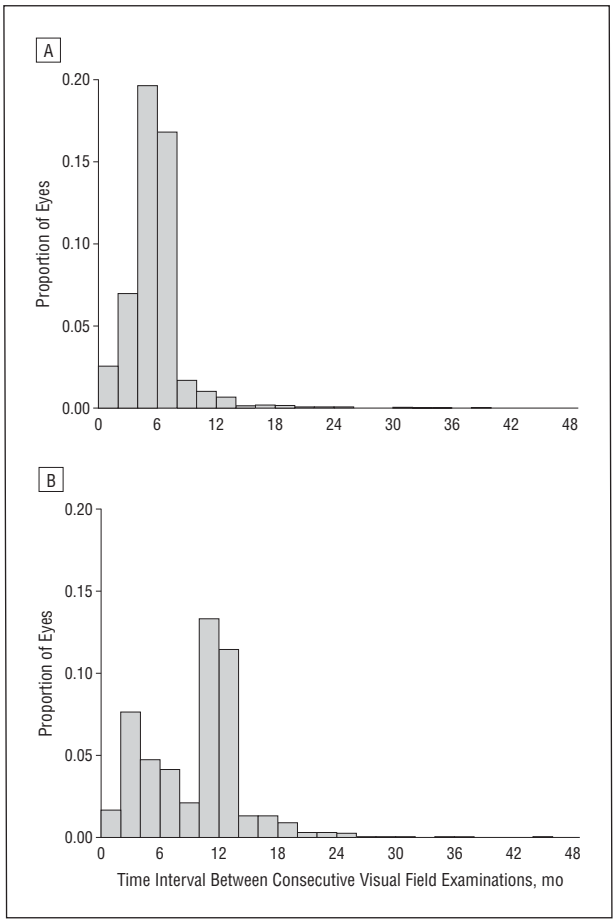

Figure 2. Distribution of time intervals between consecutive visual field examinations in the high-frequency (A) and low-frequency (B) data sets, which consisted of 468 eyes (381 patients) from the Advanced Glaucoma Intervention Study with 10 or more reliable visual field tests and 3 or more years of follow-up.

MD SLOPE RESULTS

A higher proportion of eyes in the high-frequency data set demonstrated a significant slope at each yearly follow-up period compared with the low-frequency data set starting in the third follow-up year ($P=.27$ for the second year, $P=.007$ for the third year, and $P<.001$ for all the remaining comparisons through year 12; McNemar test) (**Figure 3**). The differences in MD slopes between the 2 data sets were close to 0 for all follow-up periods (**Figure 4**). However, there was significant variability during the earlier follow-up periods, with the variability progressively decreasing as the follow-up period increased.

The results for detection of change with various MD criteria are reported in **Table 2**. Potential differences between the 2 data sets with regard to time to detection of progression according to MD slopes were also evaluated with survival analyses derived from the Cox proportional hazard model (Table 2 and **Figure 5**). Because the 2 data sets were similar in terms of all the potential covariates, only the influence of baseline MD on the performance of the 2 data sets was explored. The 2 data sets were divided into 3 strata according to baseline MD (3 equal terciles). The resulting 3 strata had an MD of greater

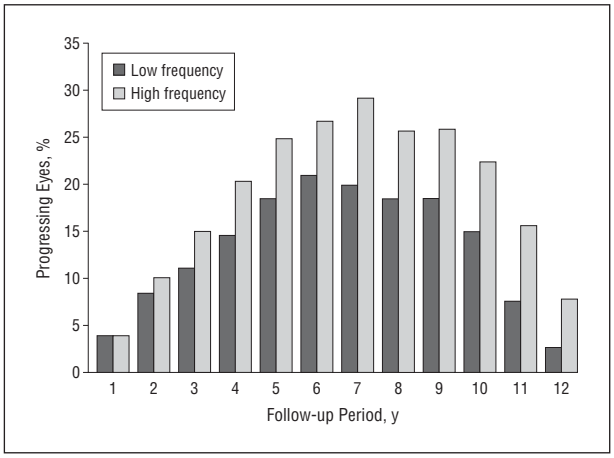

Figure 3. Proportion of eyes with worsening mean deviation (negative mean deviation slope with $P<.05$) at yearly follow-up periods.

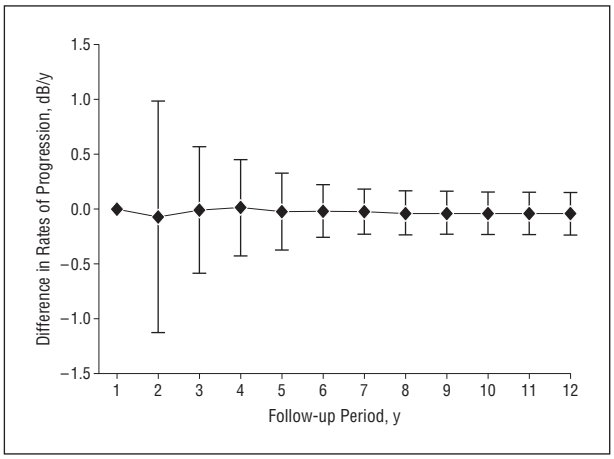

Figure 4. Difference in mean deviation slope estimates detected in the high- vs low-frequency data sets (high slope – low slope) measured at yearly intervals through 12th year of follow-up. The whiskers represent ±1 SD.

than –7.82 dB (mild loss), an MD between –7.82 and –13.5 dB (moderate loss), and an MD less than –13.5 dB (severe loss). Because the assumption of proportionality of hazards was not met when baseline MD was entered into the Cox model, 3 separate regression models were built for each baseline MD stratum. There was a trend for the difference between the 2 data sets to increase with worsening baseline MD based on the primary MD outcome (Table 2). However, the performance of MD criteria did not vary as a function of baseline severity with the 2-omitting criteria. The hazard ratio for improvement was not significantly different between the 2 data sets with the primary MD criteria, whereas the high-frequency data set detected less improvement with the 2-omitting criteria compared with the low-frequency data set (Table 2).

Table 2. Hazard Ratios for Detection of Change in the High-Frequency Data Set Compared With the Low-Frequency Data Set (Reference) According to 3 Different Mean Deviation Criteria

Direction of Change	HR (95% CI)
One Confirmation	
Progressing (206 vs 164 eyes)	1.69 (1.36-2.10)
Early	1.40 (0.98-1.98)
Moderately advanced	1.77 (1.18-2.65)
Advanced	2.05 (1.38-3.06)
Improving (29 vs 30 eyes)	0.95 (0.58-1.60)
Early	1.00 (0.35-2.85)
Moderately advanced	0.86 (0.33-2.24)
Advanced	1.00 (0.47-2.10)
One Confirmation (Data Censored at 5 Years)	
Progressing (121 vs 101 eyes)	1.45 (1.10-1.91)
Early	1.17 (0.76-1.79)
Moderately advanced	1.52 (0.92-2.50)
Advanced	1.84 (1.10-3.09)
Improving (29 vs 30 eyes)	0.95 (0.58-1.60)
Early	1.00 (0.35-2.85)
Moderately advanced	0.86 (0.33-2.24)
Advanced	1.00 (0.47-2.10)
Two-Omitting Criteria	
Progressing (211 vs 142 eyes)	2.50 (2.00-3.17)
Early	2.56 (1.77-3.70)
Moderately advanced	2.53 (1.64-3.91)
Advanced	2.36 (1.57-3.55)
Improving (27 vs 49 eyes)	0.56 (0.35-0.91)
Early	0.67 (0.26-1.73)
Moderately advanced	0.43 (0.20-0.91)
Advanced	0.77 (0.34-1.75)

Abbreviations: CI, confidence interval; HR, hazard ratio.

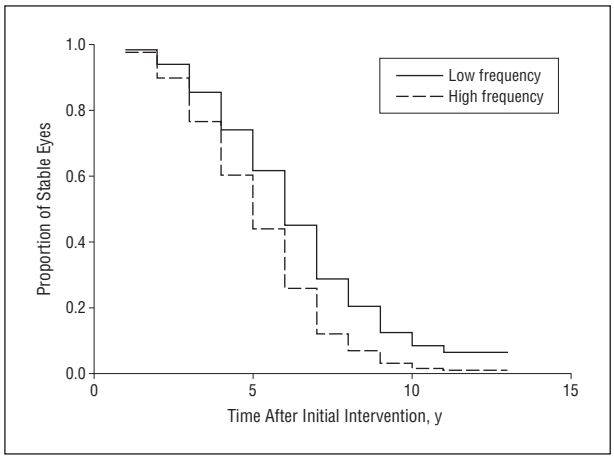

Figure 5. Proportion of stable eyes ("progression-free survival") according to mean deviation progression criteria (negative mean deviation slope with $P<.05$ confirmed once) based on the Cox proportional hazard model ($P<.001$ for the difference of the 2 curves). The solid and dashed curves represent the low- and high-frequency data sets, respectively.

PLR OUTCOMES

Overall, of 468 eyes, 185 (39.5%) in the high-frequency data set and 167 (35.7%) in the low-frequency data set progressed according to PLR criteria during follow-up (hazard ratio, 1.52 [95% confidence interval, 1.2-1.90]) (**Table 3** and **Figure 6**). The performance of PLR analysis did not vary as a function of baseline MD. The high-frequency data set was also more likely to detect improvement (Table 3). A total of 148 eyes (31.6%) from both data sets progressed according to PLR criteria. The median number of progressing test locations at initial detection of glaucoma in the subgroup of patients who progressed according to both methods was 5 (range, 3-24) in the high-

Table 3. Hazard Ratios for Detection of Change in the High-Frequency Data Set Compared With the Low-Frequency Data Set (Reference) According to Pointwise Linear Regression Criteria

Direction of Change	HR (95% CI)
One Confirmation (All Data)	
Progressing (185 vs 167 eyes)	1.52 (1.21-1.90)
Early	1.41 (0.99-2.00)
Moderately advanced	1.70 (1.12-2.58)
Advanced	1.52 (1.03-2.22)
Improving (51 vs 33 eyes)	2.27 (1.43-3.62)
Early	2.02 (0.58-7.00)
Moderately advanced	3.42 (1.54-7.60)
Advanced	1.70 (0.90-3.31)
One Confirmation (Data Censored at 5 Years)	
Progressing (105 vs 90 eyes)	1.68 (1.25-2.26)
Early	1.82 (1.13-2.92)
Moderately advanced	1.90 (1.12-3.27)
Advanced	1.32 (0.76-2.93)
Improving (32 vs 15 eyes)	2.22 (1.20-4.13)
Early	2.07 (0.51-8.30)
Moderately advanced	2.53 (0.90-7.10)
Advanced	2.08 (0.82-5.25)

Abbreviations: CI, confidence interval; HR, hazard ratio.

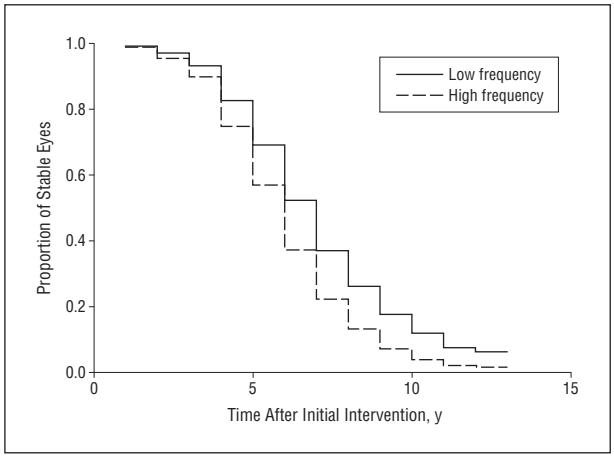

Figure 6. Proportion of stable eyes ("progression-free survival") according to pointwise linear regression progression criteria (presence of ≥3 test locations with a slope of ≤–1.0 dB/y and $P \leq .01$ with at least 1 confirmation) based on the Cox proportional hazard model ($P < .001$ for the difference of the 2 curves). The solid and dashed curves represent the low- and high-frequency data sets, respectively.

frequency data set and 4 (range, 3-26) in the low-frequency data set ($P = .11$, Wilcoxon signed rank test).

To evaluate the potential influence of long-term variability or noise on the performance of the 2 data sets, we compared the average error standard deviations and the slope standard errors between the 2 data sets. Herein, we present only the error standard deviation comparison for PLR at the final follow-up. The average error standard deviations for the yearly MD regression analyses were very similar in the 2 data sets ($P < .04$ for all, with the standard deviations for the high-frequency data set being slightly smaller). The same was true for the difference between the error standard deviations from PLR (median difference = –0.027 dB/y, with the high-frequency error standard deviations being, on average, lower; $P = .05$). The standard errors for the MD and pointwise slope estimates at the final follow-up were smaller for the high-frequency data set than for the low-frequency data set (median, 0.145 vs 0.188 dB/y for MD criteria and 0.286 vs 0.363 dB/y for PLR criteria; $P < .001$ for both).

Based on the AGIS criteria, 160 eyes (34.2%) progressed at the end of follow-up. The percentage agreement and Cohen κ statistics for agreement of regression criteria with the AGIS method were as follows. For MD criteria, percentage agreement was 73.5% vs 72.2%, and the Cohen κ statistic (SE) was 0.415 (0.046) vs 0.422 (0.046) for the low-frequency and high-frequency data sets, respectively. For PLR, the percentage agreement was 73.7% vs 72.4% and the Cohen κ statistic (SE) was 0.425 (0.046) vs 0.412 (0.046) for the low-frequency and high-frequency data sets, respectively.

COMMENT

We found that more frequent testing of the visual field (approximately every 6 months compared with every 12 months) led to a higher proportion of significant MD slopes over time (described as $P < .05$) and a shorter time to progression when 1 confirmation was required to define time to progression. Progression was 69% more likely to be detected in the high-frequency data set. When a PLR model was used to define progression, the difference between the 2 data sets was somewhat less prominent (hazard ratio, 1.52). The hazard ratios for progression are large enough, especially for MD criteria, for the benefits to be considered worth the extra time and expense required for earlier detection of glaucoma progression, at least in a subset of patients at higher risk of progression. A higher number of progressing test locations at the time of initial detection of visual field progression was found in the high-frequency data set compared with the low-frequency data set (5 vs 4 test locations), although the difference was not significant at the $P < .05$ level. The fact that the hazard ratio was higher for MD outcomes compared with PLR outcomes may be due to the fact that no criteria for the amount of change were applied with MD criteria.

Although the baseline MD seemed to have some influence on the difference in performance between the 2 data sets with our primary MD criteria, this influence was not observed with the 2-omitting criteria. Eyes with moderate to advanced glaucoma tend to have larger long-term fluctuations.[14] Eyes with moderately advanced glaucoma had the highest error standard deviations in our study in both data sets ($P < .001$, Kruskal-Wallis test; data not shown). Despite this, the performance of all regression analyses was not particularly worse in this group.

Three major parameters determine the performance of any method to detect progression: the rate of progression, the testing frequency, and the amount of variability (long-term fluctuation). Bland-Altman plots failed to show any significant trend in performance of the 2 data sets with changing rates of progression (data not shown). We hypothesized that a larger number of measurements would potentially result in a more precise detection of rates. However, a larger number of measurements could

theoretically lead to increased variability of the estimated slopes and, hence, cause increased false-positive detection. We compared the regression error standard deviations and the standard errors for the fitted slopes between the 2 data sets to test our hypothesis. The error standard deviations and the standard errors for slope estimates were very similar in the 2 data sets and tended to be smaller in the high-frequency data set. We believe that, given the similar levels of noise and standard error for the slopes, these 2 potential sources of false-positive detection could be ruled out. Given the higher number of examinations available for the high-frequency data set, it was not unexpected to find a smaller standard error for this group. The fact that the agreement of both regression criteria with AGIS method did not vary as a function of the data set used also supports the notion that a higher false-positive rate is likely not contributing to higher progression rates in the high-frequency data set. The difference in performance between the 2 testing algorithms (likely with various levels of long-term fluctuation in this real-patient visual field data set) did not seem to be as large as predicted based on the theoretical considerations by Chauhan et al.[8] This discrepancy merely reflects the differences between real data and the incomplete theoretical assumptions that are currently used to characterize longitudinal visual field data.

We also compared the proportion of improving eyes between the 2 data sets. Results showed that the number of improving eyes in the high-frequency data set was actually similar or less with the MD criteria, whereas more improvement was also detected in the high-frequency data set with the PLR criteria. We speculate that averaging of noise likely led to a smaller overall improvement rate with the MD criteria. From a purely statistical standpoint, it would be expected that a higher number of measurements would lead to more frequent detection of change in either direction. We also reanalyzed the data with MD and PLR criteria after censoring data at 5 years. The results of all analyses were consistent with the results of our primary MD criteria.

Another significant finding was that, regardless of statistical significance, the average estimates for MD slopes were very similar in the 2 data sets. However, there was significant variability during the first few years of follow-up, reaching a low plateau about 6 to 7 years into the follow-up period. This supports our hypothesis that the high-frequency data set is simply detecting change earlier because of the tighter confidence intervals made possible by a higher number of data points available.

We used raw threshold sensitivities to perform PLR analyses because it has been shown that corrected threshold sensitivities are less sensitive at detecting progression with PLR.[15,16] However, the influence of worsening media opacity or cataract extraction cannot be accounted for with our methods. Because most of the data are based on the pairwise comparison of data, there is no reason to believe that this could have affected the results.

The influence of the frequency of testing on detection of progression has been previously explored with simulation data. Gardiner and Crabb[9] found that, with clinically significant rates of progression (2 dB/y), more frequent testing improved the sensitivity, albeit at the expense of a lower specificity. The best compromise of sensitivity and specificity was found with a frequency of about 3 visual field examinations per year. However, it must be emphasized that such a frequent testing strategy is cumbersome and is not practical under most clinical circumstances. Our data provide evidence for the advantages of a more practical 6-month schedule for visual field testing.

Our results need to be interpreted in light of the following facts. We decided to keep all field examinations from the first year for both data sets. This was done to create a scenario similar to clinical patient care in which 2 to 3 visual field examinations are performed during the first year of follow-up to rule out faster progression rates. It is possible that, had a smaller number of visual field examinations been available during the first year of follow-up, it might have led to an even better relative performance of the high-frequency data set. It is also important to note that our results are not directly generalized to event analyses such as the AGIS scoring system or the Guided Progression Analysis of the Humphrey Field Analyzer (Carl Zeiss Meditec Inc, Dublin, California). It can be safely assumed, however, that, with those methods, performing more frequent tests would potentially lead to earlier detection of progression, albeit at the expense of decreased specificity, because the 2 required confirmations are more likely to be observed in a shorter time period. In clinical care, data are reviewed in real time. When suspected progression is observed with event analyses, such as Guided Progression Analysis, or with trend analyses, clinicians tend to seek confirmation within a short time regardless of previous testing frequency. A higher testing frequency is more relevant when detecting smaller amounts of progression is desirable, such as in patients with advanced glaucoma in which even a slower rate of progression needs to be detected as early as possible. It may be argued that, because eyes in the low-frequency data set were in reality tested more frequently, this might have led to a seemingly better performance of the low-frequency data set. This potential bias in performance would lead to a larger difference, if anything, between the 2 testing frequency schedules in the real world.

In summary, we found that a twice-yearly schedule of visual field testing resulted in earlier detection of glaucoma progression compared with a yearly schedule, especially with global trend analyses. Validation of these findings in other patient populations would be desirable. Our results have significant health care policy implications with regard to determining the frequency of visual field testing in patients with glaucoma.

Submitted for Publication: August 13, 2010; final revision received April 14, 2011; accepted May 18, 2011.
Published Online: August 8, 2011. doi:10.1001/archophthalmol.2011.224
Correspondence: Kouros Nouri-Mahdavi, MD MSc, Glaucoma Division, Jules Stein Eye Institute, UCLA, 100 Stein Plaza, Los Angeles, CA 90095 (nouri-mahdavi@jsei.ucla.edu).
Author Contributions: Dr Nouri-Mahdavi had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Financial Disclosure: None reported.

Funding/Support: This study was supported in part by an unrestricted grant from Research to Prevent Blindness.
Additional Contributions: We thank Jeffrey Gornbein, DrPH, Department of Biomathematics, UCLA's David Geffen School of Medicine, for his invaluable statistical assistance.

REFERENCES

1. Collaborative Normal-Tension Glaucoma Study Group. Comparison of glaucomatous progression between untreated patients with normal-tension glaucoma and patients with therapeutically reduced intraocular pressures. *Am J Ophthalmol*. 1998;126(4):487-497.
2. The AGIS Investigators. The Advanced Glaucoma Intervention Study (AGIS), 7: the relationship between control of intraocular pressure and visual field deterioration. *Am J Ophthalmol*. 2000;130(4):429-440.
3. Lichter PR, Musch DC, Gillespie BW, et al; CIGTS Study Group. Interim clinical outcomes in the Collaborative Initial Glaucoma Treatment Study comparing initial treatment randomized to medications or surgery. *Ophthalmology*. 2001; 108(11):1943-1953.
4. Weinreb RN, Kaufman PL. The glaucoma research community and FDA look to the future: a report from the NEI/FDA CDER Glaucoma Clinical Trial Design and Endpoints Symposium. *Invest Ophthalmol Vis Sci*. 2009;50(4):1497-1505.
5. Boden C, Blumenthal EZ, Pascual J, et al. Patterns of glaucomatous visual field progression identified by three progression criteria. *Am J Ophthalmol*. 2004; 138(6):1029-1036.
6. Keltner JL, Johnson CA, Quigg JM, Cello KE, Kass MA, Gordon MO; Ocular Hypertension Treatment Study Group. Confirmation of visual field abnormalities in the Ocular Hypertension Treatment Study. *Arch Ophthalmol*. 2000;118(9): 1187-1194.
7. Heijl A, Bengtsson B, Hyman L, Leske MC; Early Manifest Glaucoma Trial Group. Natural history of open-angle glaucoma. *Ophthalmology*. 2009;116(12):2271-2276.
8. Chauhan BC, Garway-Heath DF, Goñi FJ, et al. Practical recommendations for measuring rates of visual field change in glaucoma. *Br J Ophthalmol*. 2008; 92(4):569-573.
9. Gardiner SK, Crabb DP. Frequency of testing for detecting visual field progression. *Br J Ophthalmol*. 2002;86(5):560-564.
10. The AGIS Investigators. The Advanced Glaucoma Intervention Study (AGIS), 1: study design and methods and baseline characteristics of study patients. *Control Clin Trials*. 1994;15(4):299-325.
11. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307-310.
12. Gardiner SK, Crabb DP. Examination of different pointwise linear regression methods for determining visual field progression. *Invest Ophthalmol Vis Sci*. 2002; 43(5):1400-1407.
13. Nouri-Mahdavi K, Caprioli J, Coleman AL, Hoffman D, Gaasterland D. 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.
14. Heijl A, Lindgren A, Lindgren G. Test-retest variability in glaucomatous visual fields. *Am J Ophthalmol*. 1989;108(2):130-135.
15. Artes PH, Nicolela MT, LeBlanc RP, Chauhan BC. Visual field progression in glaucoma: total versus pattern deviation analyses. *Invest Ophthalmol Vis Sci*. 2005; 46(12):4600-4606.
16. Manassakorn A, Nouri-Mahdavi K, Koucheki B, Law SK, Caprioli J. Pointwise linear regression analysis for detection of visual field progression with absolute versus corrected threshold sensitivities. *Invest Ophthalmol Vis Sci*. 2006;47 (7):2896-2903.