Comparison of Short- And Long-Term Variability in Standard Perimetry and Spectral Domain Optical Coherence Tomography in Glaucoma

CARLA N. URATA, EDUARDO B. MARIOTTONI, ALESSANDRO A. JAMMAL, NARA G. OGATA, ATALIE C. THOMPSON, SAMUEL I. BERCHUCK, TAIS ESTRELA, AND FELIPE A. MEDEIROS

• PURPOSE: To assess short- and long-term variability on standard automated perimetry (SAP) and spectral domain optical coherence tomography (SD-OCT) in glaucoma.
• DESIGN: Prospective cohort.
• METHODS: Ordinary least squares linear regression of SAP mean deviation (MD) and SD-OCT global retinal nerve fiber layer (RNFL) thickness were fitted over time for sequential tests conducted within 5 weeks (short-term testing) and annually (long-term testing). Residuals were obtained by subtracting the predicted and observed values, and each patient's standard deviation (SD) of the residuals was used as a measure of variability. Wilcoxon signed-rank test was performed to test the hypothesis of equality between short- and long-term variability.
• RESULTS: A total of 43 eyes of 43 glaucoma subjects were included. Subjects had a mean 4.5 ± 0.8 SAP and OCT tests for short-term variability assessment. For long-term variability, the same number of tests were performed and results annually collected over an average of 4.0 ± 0.8 years. The average SD of the residuals was significantly higher in the long-term than in the short-term period for both tests: 1.05 ± 0.70 dB vs. 0.61 ± 0.34 dB, respectively ($P < 0.001$) for SAP MD and 1.95 ± 1.86 μm vs. 0.81 ± 0.56 μm, respectively ($P < 0.001$) for SD-OCT RNFL thickness.
• CONCLUSIONS: Long-term variability was higher than short-term variability on SD-OCT and SAP. Because current event-based algorithms for detection of glaucoma progression on SAP and SD-OCT have relied on short-term variability data to establish their normative databases, these algorithms may be underestimating the variability in the long-term and thus may overestimate progression over time. (Am J Ophthalmol 2020;210:19–25. © 2019 Elsevier Inc. All rights reserved.)

Accepted for publication Oct 24, 2019.

From the Duke Eye Center and Department of Ophthalmology (C.N.U., E.B.M., A.A.J., N.G.O., A.C.T., S.I.B., T.E., F.A.M.), Duke University, Durham, North Carolina, USA; and the Department of Statistical Science and Forge (S.I.B.), Duke University, Durham, North Carolina, USA.

Inquiries to Felipe A. Medeiros, Visual Performance Laboratory, Duke Eye Center, Department of Ophthalmology, Duke University, 2310 Erwin Road, Durham, North Carolina 27710, USA; e-mail: felipe.medeiros@duke.edu

GLAUCOMA IS CHARACTERIZED BY A PROGRESSIVE optic neuropathy with corresponding patterns of visual field loss.[1] Monitoring and detection of glaucoma progression over time is paramount in management and clinical decision making, such as when to initiate or escalate therapy.[2] However, despite the availability of numerous functional and structural tests for monitoring glaucoma, such as standard automated perimetry (SAP) and optical coherence tomography (OCT), detection of progression remains a challenging aspect of clinical practice.

Effective detection of progression depends fundamentally on the ability to differentiate true change from test-retest variability. Because glaucoma is usually a slowly progressive disease, true changes are not expected to occur over relatively short time frames. This reasoning has been used as the basis for establishing normative databases of variability by conducting repeated testing over short periods of time in glaucomatous eyes, usually within a few weeks, and calculating confidence limits or tolerance intervals of variability.[3] If a patient is subsequently found to have a change that is greater than those confidence limits, the patient is deemed to have progressed. Such approach has been used by the so-called event-based algorithms for detecting progression such as the Guided Progression Analysis (GPA software; Carl Zeiss Meditec, Dublin, California) for SAP.[4] In the GPA, follow-up test results are compared to baseline test results, and if a number of points show a change that exceeds the expected variability, glaucoma is declared to be progressing. The GPA has been widely used in clinical practice and clinical trials and has also been recently extended for detecting structural progression on OCT.[4,5]

Establishing normative levels of variability based on short-term test-retest, however, may be problematic. Glaucoma patients and those suspected of having the disease are monitored over the course of many years, and there are reasons to believe that the long-term variability may be different than the short-term variability. Short-term studies of variability tend to enroll experienced patients who, not uncommonly, have participated in other studies and are thus usually highly cooperative and motivated.[6] Also, technicians tend to be skilled and remain that way throughout the study. In contrast, in "real-world" long-term monitoring, much less motivated patients are likely

0002-9394/$36.00
https://doi.org/10.1016/j.ajo.2019.10.034

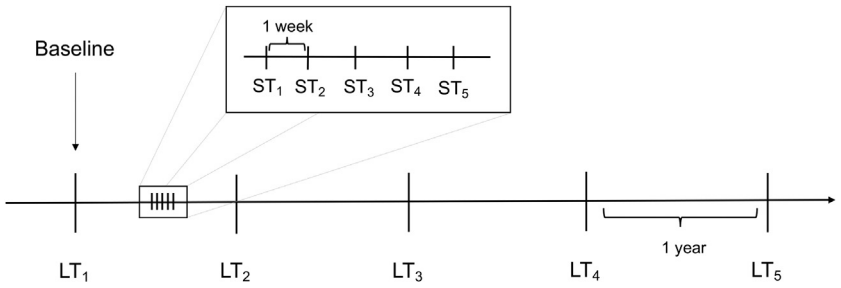

FIGURE 1. Timeline illustrating an example of ST and LT visits typical of the study patients. In the example, 5 ST tests (weekly) were compared with 5 LT tests (annually) of the same eye. LT visits were selected to match the number of ST visits. ST = short-term visit. LT = long-term visit.

to be encountered who may also have intercurrent conditions affecting test result quality. Long-term testing is likely to be done by different technicians showing a variety of degrees of training and expertise. If long-term variability is significantly different compared to short-term variability, then the algorithms for detection of progression that rely on confidence limits of variability from short-term test-retest results may provide spurious assessments of whether true change has occurred or not.

In this study, the test-retest estimates of short-term variability were compared with long-term variability of SAP and spectral domain OCT (SD-OCT) measurements in a cohort of glaucoma patients followed over time.

SUBJECTS AND METHODS

PARTICIPANTS FROM THIS STUDY WERE CONSECUTIVELY recruited from the clinic and were enrolled in a prospective longitudinal study designed to evaluate functional impairment in glaucoma. The Institutional Review Board approved all methods, and written informed consent was obtained from all participants. The methodology complied with the Declaration of Helsinki guidelines for human subject research, and this study adhered to the Health Insurance Portability and Accountability Act.

Patients underwent a comprehensive ophthalmologic examination, including medical history, visual acuity, slit-lamp biomicroscopy, intraocular pressure measurement using Goldmann applanation tonometry, gonioscopy and dilated fundoscopy using a 78-diopter (D) lens every 6 months. In addition, all patients included in this study were required to have open angles, visual acuity of ≥20/40, and spherical equivalent of <3.0 D throughout the study. Subjects with coexisting retinal disease, uveitis, or any systemic disease that could affect the optic nerve head, or the visual field, were excluded. Subjects who had undergone cataract surgery during the follow-up period were also excluded.

All patients underwent SAP tests, using the 24-2 Swedish interactive threshold algorithm standard of the Humphrey field analyzer II (Carl Zeiss Meditec). Only reliable visual fields with less than 15% false positives and less than 33% fixation losses were included, and the first 2 reliable examinations were excluded in order to avoid learning effects. SAP examinations with the presence of eyelid artifacts, rim artifacts, or other evidence of artifactual visual field defects not related to glaucoma were also excluded.

Patients also were tested using the Spectralis SD-OCT (software version 5.4.7.0; Heidelberg Engineering, Heidelberg, Germany) to measure the peripapillary retinal nerve fiber layer (RNFL) thickness. For SD-OCT, axial length and corneal curvature measurements were entered into the instrument's software to ensure accurate scaling of all measurements, and the device's eye-tracking capability was used during image acquisition to ensure that the same location of the retina was scanned over time. Images were excluded if the signal strength was <15 dB or if they were inverted or clipped. The global circumpapillary RNFL thickness was used as the study metric and corresponded to the 360° average measurement of the 1,535 A-scan points acquired from a circle of 3.45 mm centered on the optic disc, which was automatically calculated by the SD-OCT software. In this study, the pool of technicians performing perimetry and SD-OCT consisted of 5 experienced and trained technicians. However, each subject was not necessarily tested by the same technician over the course of the study.

Glaucoma diagnosis was defined as the presence of at least 2 consecutive reliable SAP test results with abnormalities at baseline (pattern standard deviation with a P value of <0.05 and/or glaucoma hemifield test results outside normal limits) with corresponding optic nerve damage (i.e., neuroretinal rim thinning, cupping, notching, or characteristic RNFL defects). Only patients with open angle glaucoma in at least 1 eye were included in the study. If both eyes of the same patient met the criteria, one eye was randomly chosen for the analysis.

• **ESTIMATION OF LONG-TERM AND SHORT-TERM VARIABILITY:** Figure 1 illustrates the timeline of the visits included in the determination of long- and short-term

TABLE 1. Demographic and Clinical Characteristics of the Study Patients at Baseline

Variable	43 Subjects (43 Eyes)
Age (y)	71.2 ± 9.7
Females	19 (44)
Race	
White	24 (56)
(African-American descendent, %)	14 (32)
(Asian, %)	4 (9)
(American Indian or Alaska native, %)	1 (2)
IOP (mm Hg)	14.9 ± 4.9
SAP 24-2 baseline MD (dB)	−8.4 (−25.1, 0.3)[a]
RNFL global thickness at baseline (μm)	69.8 ± 20.8

IOP = intraocular pressure; SAP = standard automated perimetry; MD = mean deviation; RNFL = retinal nerve fiber layer.
Values are mean ± SD or n (%), unless otherwise noted.
[a]Values are median (interquartile range).

TABLE 2. Comparison of Short-Term and Long-Term Variability for SAP and SD OCT

Parameter	SD of Residuals Short-Term	SD of Residuals Long-Term	*P* Value
SAP MD (dB)	0.61 ± 0.34	1.05 ± 0.70	<0.001[a]
SD OCT global RNFL thickness (μm)	0.81 ± 0.56	1.95 ± 1.86	<0.001[a]

MD = mean deviation; RNFL = retinal nerve fiber layer; SD = standard deviation; SD OCT = spectral-domain optical coherence tomography.
Values are mean ± SD.
[a]Wilcoxon signed-rank test.

variabilities. Annual SAP and SD-OCT visits were used to estimate long-term variability. To estimate short-term variability, subjects were invited to perform a sequence of 5 additional weekly visits at some point during follow-up. The number of short- and long-term visits were matched for each subject. The same method was used to estimate variability for both the long-term as well as short-term testing, consisting of fitting ordinary least squares (OLS) linear regression models of the parameter of interest over time and then using the standard deviation (SD) of the residuals of the OLS model as an estimate of variability. This approach has been previously described[7–14] and was applied in the current study for SAP MD as well as for SD-OCT global RNFL thickness. The SD of residuals was used to determine short- and long-term variability because it gives a measure of variability that is less affected by the possibility of progression over time, assuming that any progression within the observed period would be linear. For the long-term variability, only the annual visits were used for the OLS model. For the short-term variability, only the weekly visits were used.

• **STATISTICAL ANALYSIS:** To test the hypothesis that long- and short-term variability are different, the differences in SD of the residuals over long- and short-term visits for both SAP MD and SD-OCT RNFL thickness were analyzed. To make the comparison, the Wilcoxon signed-rank test was used, because the data were paired and not normally distributed (confirmed by a Shapiro-Wilk test).

We investigated the relationship between the differences in SD of residuals for long- and short-term variability and disease severity for each test. Because the relationships were not linear, a quadratic curve was fitted. In addition, Spearman rank correlation was used to analyze the correlation between long- and short-term variability for all eyes and the correlation of the difference between long- and short-term variability and age. All statistical analyses were performed using Stata version 15.1 software (StataCorp, College Station, Texas). The α level (type I error) was set at 0.05.

RESULTS

THE STUDY INCLUDED 43 EYES OF 43 SUBJECTS WITH A MEAN age of 71.2 ± 9.7 years old and an average follow-up time of 4.0 ± 0.8 years. Subjects had 4.5 ± 0.8 short-term visits matched with the same number of long-term visits during the study period. Demographic and clinical characteristics of the enrolled subjects are displayed in Table 1.

Results for short- and long-term variability for each test are summarized in Table 2. The average SD of the residuals was significantly greater in the long-term test results than in the short-term results for both SAP MD (1.05 ± 0.70 dB vs. 0.61 ± 0.34 dB, respectively; mean difference, 0.44 dB; 95% confidence interval [CI], 0.26-0.63; $P < 0.001$) and SD-OCT global RNFL thickness (1.95 ± 1.86 μm vs. 0.81 ± 0.56 μm, respectively; mean difference, 1.15 μm; 95% CI, 0.58-1.71; $P < 0.001$). Figure 2 illustrates the distribution of the SD of the residuals for both short- and long-term testing in both modalities. There was a greater spread in the distribution of the SD of residuals in the long-term results compared to the short-term results for both SAP and SD-OCT. Figure 3 illustrates the distributions of the differences between the long- and short-term SD values of the residuals for SAP MD and SD-OCT. There was a moderate correlation between short- and long-term SD of residuals for SAP MD (rho = 0.39; $P = 0.009$) but a weak correlation for SD-OCT global RNFL thickness (rho = 0.28; $P = 0.064$).

Next the relationship between differences in short- and long-term variability and disease severity were investigated. A quadratic model was used to describe the relationship between the difference of the SD of residuals and

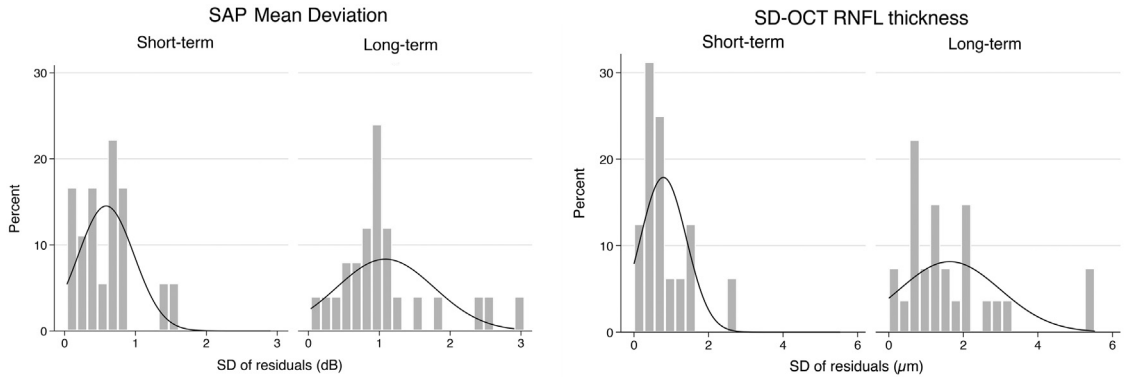

FIGURE 2. Distribution of the standard deviation of the residuals for both short- and long-term visits of SAP mean deviation (left) and SD-OCT retinal nerve fiber layer thickness (right). SAP = standard automated perimetry; SD-OCT = spectral-domain optical coherence tomography.

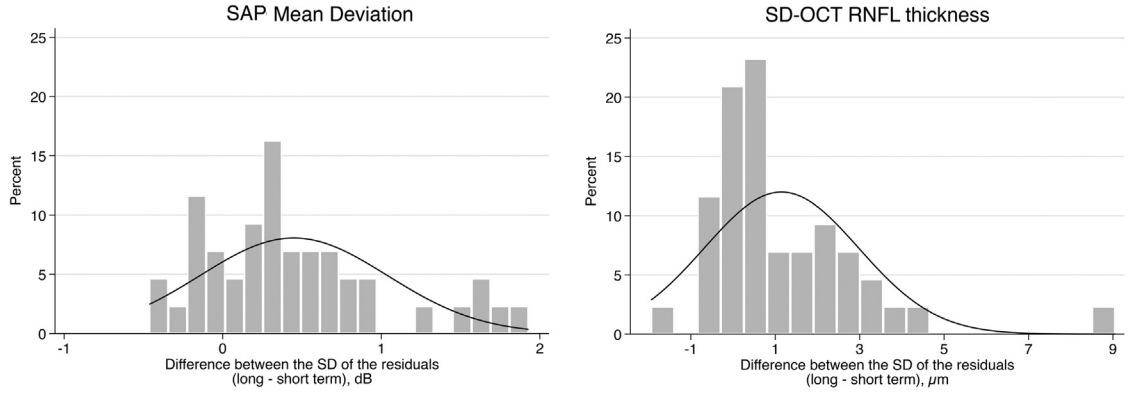

FIGURE 3. Distribution of the difference over long- and short-term visits in the standard deviation of the residuals for both SAP mean deviation (left) and SD-OCT RNFL thickness (right). RNFL = retinal nerve fiber layer; SAP = standard automated perimetry; SD-OCT = spectral-domain optical coherence tomography.

either the average of all MDs or the average of all RNFL thickness measurements for each eye, as these relationships were not linear (Figure 4 and 5). There was a significant relationship between the difference in short- and long-term SD of residuals and average MD ($R^2 = 0.174$; $P = 0.021$). The differences in the SD of the residuals appeared to be greatest for eyes with average MD of approximately −12 dB (Figure 4). For SD-OCT, the differences in the SD of the residuals in long- and short-term testing did not vary significantly with the average global RNFL thickness ($R^2 = 0.046$; $P = 0.385$) (Figure 5). In addition, the correlation of the differences in long- and short-term variability and the subjects' average age during the study period were studied. It was found that the differences in long- and short-term SD of the residuals had a weak correlation with age for SAP MD (rho = 0.24; $P = 0.033$) and a moderate correlation with RNFL thickness (rho = 0.45; $P = 0.014$).

DISCUSSION

BECAUSE GLAUCOMA IS USUALLY A SLOWLY PROGRESSIVE disease, short-term testing has been used to help clinicians identify thresholds of expected normal variability.[15,16] In the present study, the structural and functional variability seen in measurements acquired during long-term testing was shown to be significantly greater than the variability seen in measurements taken during short-term follow-up. This may have significant implications in determining whether or not true progression has occurred when using event-based algorithms to assess progression. To the best of the present authors' knowledge, this is the first study to evaluate the differences in long- and short-term variability for both SAP and SD-OCT in glaucoma patients.

In this study, long-term variability was 1.7 and 2.4 times higher than short-term variability for SAP and SD-OCT, respectively. A previous study using computer simulations

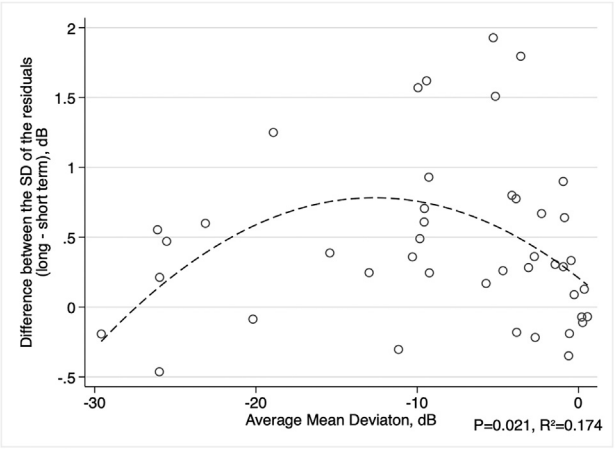

FIGURE 4. Scatterplot between the average SAP MD for each eye and the average difference in the standard deviation of the residuals (long-term minus short-term variability). The dashed line represents a quadratic fit. MD = mean deviation; SAP = standard automated perimetry.

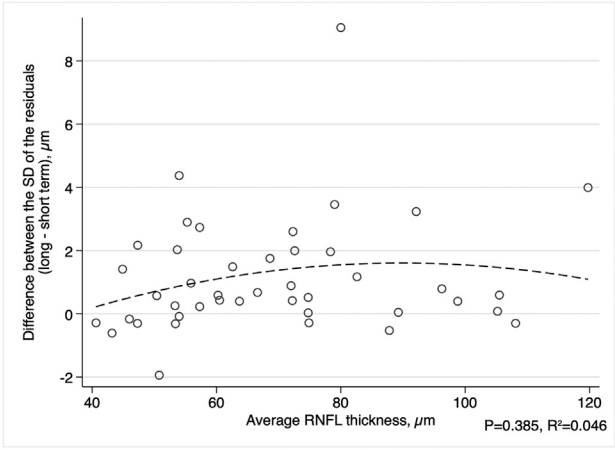

FIGURE 5. Scatterplot between the average RNFL thickness for each eye and the average difference in the standard deviation of the residuals (long-term minus short-term variability). The dashed line represents a quadratic fit. RNFL = retinal nerve fiber layer.

of MD variability over time suggested that the SAP variability must be reduced by approximately 20% for a clinically appreciable improvement in detection of visual field change.[17] Therefore, using a similar reasoning, an increase of 70% in variability would likely result in a clinically appreciable worsening in the ability to detect visual field change in the long-term. For SD-OCT, it is likely that a variability that is more than 2 times when assessed in the long-versus the short-term is also likely to be significant in the ability to detect change with this instrument.

The present findings confirm that the difference between long- and short-term variability of SAP MD tends to increase as MD values worsen, although not in a linear way.[9,18] Figure 4 shows that the differences between long- and short-term variability increased with increasing visual field loss, peaking at values of approximately −12 dB and then declining as visual field damage becomes more severe and closer to the floor level. For SD-OCT, although the differences between short- and long-term variability peaked at approximately 90 μm, the relationship with disease severity was not statistically significant. The factors explaining differences between long- and short-term variability according to levels of damage are likely to be related to the precision of the instruments at different levels of disease and the dynamic range of the tests.

Previous studies have similarly found that long-term variability exceeds short-term variability, but they used different approaches and instruments. Medeiros and associates[19] evaluated the estimates of long-term variability in stable glaucoma patients by using a different instrument (GDx VCC Retinal Scanner, Zeiss, Oberkochen, Germany) to measure RNFL thickness. Long-term variability was calculated as 1.96 times the inter-visit SD and the short-term variability as 1.96 times the intra-visit SD. The long-term variability estimates ranged from 3.21 to 4.97 μm, whereas the short-term ranged from 2.45 to 3.89 μm. The present results for long-term variability are comparable to those in the study by Gardiner and associates.[7] They studied the longitudinal signal-to-noise ratios in structural and functional tests, using the SD of the residuals from the OLS over time to measure the long-term variability of RNFL thickness using the Spectralis SD-OCT and SAP MD. Short-term variability was not studied, however. For RNFL thickness, the SD of the residuals was 1.76 μm and 0.58 dB for SAP MD. However, Gardiner and associates[7] used a shorter testing interval of 6 months rather than 12 months, which may explain why the long-term variability is slightly smaller in their study than in the present study. In another study, Katz and associates[20] found that, for normal eyes, visual field tests acquired over longer intervals also had greater variability than tests taken only 1 week apart. However, their results cannot be directly compared to those in the present study because they tested only healthy subjects.

Regardless of the data modality studied for progression detection, structural or functional, the present study showed that long- and short-term variabilities are different. The reasons for this are likely multifactorial.[21] Subjects enrolled in studies that undergo visual field and imaging testing in short periods of time are likely to be better test takers than patients routinely followed in clinical practice.[3,22] Furthermore, it is likely that multiple technicians with different skills, experience, and supervision will perform the tests in the long term. Subjects followed in the long term may be less motivated to undergo routine examinations and are more likely to be tested under different

conditions that can affect test performance and increase variability. For example, changes in environmental and ocular conditions (e.g., media opacities due to dry eyes), psychological factors and physical ability (e.g., difficulty in properly positioning themselves without tilting the head during the test) are all likely to affect test performance in the long run.[14,21,23–25] Even though the study included only subjects with visual acuity better than 20/40 throughout the study period, it is possible, and expected, that some of these eyes might have developed cataract during the follow-up. This could play a role in the larger variability found in the long-compared to the short-term. However, this replicates the clinical practice scenario as patients are being followed over time.

Our study has limitations. The residuals of OLS regression were used to estimate variability. This assumes that any progression, if occurring, would be linear. This may not be strictly true, especially for SAP MD.[26] However, although changes over the full course of the disease are unlikely to be truly linear, the assumption of linearity is a very sensible one for periods encompassing just a few years,[27] as in the current study. Notably, clinical management decisions are usually made taking into account test results collected over similar periods of time in clinical practice. As another limitation, SAP pointwise sensitivities or sectoral RNFL changes were not analyzed, as is done in many event-based algorithms.[3,4,28,29] However, it is likely that differences in short- and long-term variability would be even greater if localized sectors were analyzed, as these have been shown overall to have greater variability.[30,31]

In conclusion, this study showed that long-term variability is significantly greater than short-term variability for both structural and functional tests in glaucoma. These results underscore the importance of accounting for the greater variability that occurs during long-term testing when developing algorithms that detect progression in glaucoma, as algorithms that use short-term testing to establish normative levels of variability will tend to overestimate progression over time and could lead to inappropriate escalation of therapy in patients with clinically stable disease.

ALL AUTHORS HAVE COMPLETED AND SUBMITTED THE ICMJE FORM FOR DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST and none were reported.

FUNDING/SUPPORT: This work was supported in part by US National Institutes of Health/National Eye Institute grants EY029885 (F.A.M.), EY027651 (F.A.M.), and EY021818 (F.A.M.).

Financial Disclosures: F.A.M. received support from Alcon Laboratories, Allergan, Bausch & Lomb, Carl Zeiss Meditec, Heidelberg Engineering, Merck, nGoggle Inc., Sensimed, Topcon, Reichert, and National Institutes of Health/National Eye Institute. All other authors indicate no financial support or financial conflict of interest. The coauthors have seen and agree with each of the changes made to the manuscript in this revision and to the way his or her name is listed.

REFERENCES

1. Weinreb RN, Aung T, Medeiros FA. The pathophysiology and treatment of glaucoma: a review. *JAMA* 2014;311:1901–1911.
2. Prum BE Jr, Lim MC, Mansberger SL, et al. Primary open-angle glaucoma suspect preferred practice pattern guidelines. *Ophthalmology* 2016;123:P112–P151.
3. Artes PH, O'Leary N, Nicolela MT, Chauhan BC, Crabb DP. Visual field progression in glaucoma: what is the specificity of the Guided Progression Analysis? *Ophthalmology* 2014;121:2023–2027.
4. 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;247:391–397.
5. Lee EJ, Kim TW, Weinreb RN, Park KH, Kim SH, Kim DM. Trend-based analysis of retinal nerve fiber layer thickness measured by optical coherence tomography in eyes with localized nerve fiber layer defects. *Invest Ophthalmol Vis Sci* 2011;52:1138–1144.
6. Chauhan BC, Johnson CA. Test-retest variability of frequency-doubling perimetry and conventional perimetry in glaucoma patients and normal subjects. *Invest Ophthalmol Vis Sci* 1999;40:648–656.
7. Gardiner SK, Fortune B, Demirel S. Signal-to-noise ratios for structural and functional tests in glaucoma. *Transl Vis Sci Technol* 2013;2:3.
8. Wu Z, Medeiros FA. Impact of different visual field testing paradigms on sample size requirements for glaucoma clinical trials. *Sci Rep* 2018;8:4889.
9. Wu Z, Medeiros FA. Development of a visual field simulation model of longitudinal point-wise sensitivity changes from a clinical glaucoma cohort. *Transl Vis Sci Technol* 2018;7:22.
10. Gracitelli CPB, Zangwill LM, Diniz-Filho A, et al. Detection of glaucoma progression in individuals of african descent compared with those of European descent. *JAMA Ophthalmol* 2018;136:329–335.
11. Strouthidis NG, Scott A, Peter NM, Garway-Heath DF. Optic disc and visual field progression in ocular hypertensive subjects: detection rates, specificity, and agreement. *Invest Ophthalmol Vis Sci* 2006;47:2904–2910.
12. Howard CL, Wallace C, Abbas J, Stokic DS. Residual standard deviation: validation of a new measure of dual-task cost in below-knee prosthesis users. *Gait Posture* 2017;51:91–96.
13. Ledolter J, Kardon RH. Assessing trends in functional and structural characteristics: a survey of statistical methods with an example from ophthalmology. *Transl Vis Sci Technol* 2018;7:34.

14. Diniz-Filho A, Delano-Wood L, Daga FB, Cronemberger S, Medeiros FA. Association between neurocognitive decline and visual field variability in glaucoma. *JAMA Ophthalmol* 2017;135:734–739.
15. Artes PH, Hutchison DM, Nicolela MT, LeBlanc RP, Chauhan BC. Threshold and variability properties of matrix frequency-doubling technology and standard automated perimetry in glaucoma. *Invest Ophthalmol Vis Sci* 2005;46:2451–2457.
16. Wadhwani M, Bali SJ, Satyapal R, et al. Test-retest variability of retinal nerve fiber layer thickness and macular ganglion cell-inner plexiform layer thickness measurements using spectral-domain optical coherence tomography. *J Glaucoma* 2015;24:e109–e115.
17. Turpin A, McKendrick AM. What reduction in standard automated perimetry variability would improve the detection of visual field progression? *Invest Ophthalmol Vis Sci* 2011;52:3237–3245.
18. Russell RA, Garway-Heath DF, Crabb DP. New insights into measurement variability in glaucomatous visual fields from computer modelling. *PLoS One* 2013;8:e83595.
19. Medeiros FA, Doshi R, Zangwill LM, Vasile C, Weinreb RN. Long-term variability of GDx VCC retinal nerve fiber layer thickness measurements. *J Glaucoma* 2007;16:277–281.
20. Katz J, Sommer A. A longitudinal-study of the age-adjusted variability of automated visual-fields. *Arch Ophthalmol* 1987;105:1083–1086.
21. Montolio FGJ, Wesselink C, Gordijn M, Jansonius NM. Factors that influence standard automated perimetry test results in glaucoma: test reliability, technician experience, time of day, and season. *Invest Ophthalmol Vis Sci* 2012;53:7010–7017.
22. Wu Z, Medeiros FA. Comparison of visual field point-wise event-based and global trend-based analysis for detecting glaucomatous progression. *Transl Vis Sci Technol* 2018;7:20.
23. Henson DB, Emuh T. Monitoring vigilance during perimetry by using pupillography. *Invest Ophthalmol Vis Sci* 2010;51:3540–3543.
24. Kocabeyoglu S, Mocan MC, Bozkurt B, Irkec M. Effect of artificial tears on automated visual field testing in patients with glaucoma and dry eye. *Can J Ophthalmol* 2013;48:110–114.
25. Rozanski C, Haythornthwaite JA, Dagnelie G, Bittner AK. Applying theories and interventions from behavioral medicine to understand and reduce visual field variability in patients with vision loss. *Med Hypotheses* 2014;83:190–195.
26. Pathak M, Demirel S, Gardiner SK. Nonlinear, multilevel mixed-effects approach for modeling longitudinal standard automated perimetry data in glaucoma. *Invest Ophthalmol Vis Sci* 2013;54:5505–5513.
27. Gardiner SK, Demirel S, De Moraes CG, et al. Series length used during trend analysis affects sensitivity to changes in progression rate in the ocular hypertension treatment study. *Invest Ophthalmol Vis Sci* 2013;54:1252–1259.
28. Abe RY, Gracitelli CP, Medeiros FA. The use of spectral-domain optical coherence tomography to detect glaucoma progression. *Open Ophthalmol J* 2015;9:78–88.
29. Leung CK, Yu M, Weinreb RN, Lai G, Xu G, Lam DS. Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: patterns of retinal nerve fiber layer progression. *Ophthalmology* 2012;119:1858–1866.
30. Budenz DL, Fredette M-J, Feuer WJ, Anderson DR. Reproducibility of peripapillary retinal nerve fiber thickness measurements with stratus OCT in glaucomatous eyes. *Ophthalmology* 2008;115:661–666.
31. Saunders LJ, Russell RA, Crabb DP. Measurement precision in a series of visual fields acquired by the standard and fast versions of the Swedish interactive thresholding algorithm: analysis of large-scale data from clinics. *JAMA Ophthalmol* 2015;133:74–80.