



OPEN Impact of race and ethnicity on glaucoma progression detection by perimetry and optical coherence tomography

Luiz A. F. Beniz^{1,2}, Alessandro A. Jammal¹, Douglas R. da Costa¹, Eduardo B. Mariottoni², Swarup S. Swaminathan¹ & Felipe A. Medeiros¹✉

This study assessed the impact of race and ethnicity on longitudinal test variability and time to detect glaucoma progression using standard automated perimetry (SAP) and optical coherence tomography (OCT). The sample consisted of 47,003 SAP tests from 5402 eyes and 25,480 OCT tests from 4125 eyes, with 20% of participants self-identifying as Black or African American and 80% as White; 29% as Hispanic or Latino and 71% as Not Hispanic or Latino. Variability was measured using standard deviations of residuals from linear regression models for SAP mean deviation (MD) and OCT retinal nerve fiber layer (RNFL) thickness over time. Results showed significantly greater SAP variability in Black or African American (1.80 ± 1.30 dB) compared to White participants (1.56 ± 1.21 dB; $P < 0.001$) and in Hispanic or Latino (1.81 ± 1.46 dB) compared to Not Hispanic or Latino individuals (1.52 ± 1.10 dB; $P < 0.001$). OCT variability was higher in Black or African American (2.3 ± 1.5 μ m) compared to White (2.1 ± 1.3 μ m; $P < 0.001$) and in Not Hispanic or Latino (2.2 ± 1.3 μ m) compared to Hispanic or Latino (2.1 ± 1.2 μ m; $P = 0.029$). Increased SAP variability delayed progression detection, while OCT showed minimal differences. These findings suggest that higher perimetric variability in Black or African American and Hispanic or Latino may affect glaucoma progression detection using SAP.

Keywords Race, Ethnicity, Glaucoma progression, Perimetry, Visual field, Optical coherence tomography.

Abbreviations

| | |
|------|--|
| BPGR | Bascom palmer glaucoma repository |
| HFA | Humphrey visual field analyzer |
| ICD | International classification of diseases |
| MD | mean deviation |
| OAG | open-angle glaucoma |
| OCT | optical coherence tomography |
| RNFL | retinal nerve fiber layer thickness |
| SAP | standard automated perimetry |
| SD | standard deviation |
| SITA | Swedish interactive thresholding algorithm |

Glaucoma is a chronic optic neuropathy and the leading cause of irreversible blindness worldwide¹. While the development and progression of this disease can be influenced by multiple factors¹, the impact of racial and ethnic differences remains a complex and poorly understood aspect. The exact causes of these differences have not yet been elucidated².

Population-based studies in diverse settings have confirmed that glaucoma is more prevalent in Black or African American individuals compared to White individuals^{3–6}. In addition to increased prevalence, glaucoma can lead to a disproportionately higher rate of visual impairment in the former group^{7–10}. In a series of earlier investigations, we hypothesized that increased variability in standard automated perimetry (SAP) testing in Black or African American individuals might contribute to explain some of the observed disparities. This increased variability could lead to delays in detecting progressive damage, resulting in postponed interventions

¹Bascom Palmer Eye Institute, University of Miami, 900 NW 17th St, Miami, FL, USA. ²Department of Ophthalmology and Visual Sciences, Paulista School of Medicine, Federal University of São Paulo, São Paulo, Brazil. ✉email: fmedeiros@med.miami.edu

and consequently higher rates of visual impairment^{11,12}. Although the underlying causes of increased SAP variability among Black or African American subjects remain unclear, they could be associated with poorer socioeconomic conditions commonly experienced by this group and potential effects of systemic racism in test administration procedures^{11,12}.

Previous studies have shown that Hispanics and Latinos are also at higher risk for visual impairment from glaucoma^{10,13,14}. It is possible that issues related to test-retest variability may also affect the ability of clinicians to diagnose glaucoma progression in these groups, which could explain, at least in part, differences in outcomes. However, to the best of our knowledge, this hypothesis has not been tested.

The current study aims to extend our earlier research by examining longitudinal test-retest variability and predicted times to detect glaucoma progression across different racial and ethnic groups, including Hispanic and Latino populations. Additionally, we aimed to assess test-retest variability using structural evaluations performed with Optical Coherence Tomography (OCT), in addition to perimetry. We hypothesized that variability would be more pronounced in SAP compared to OCT, due to the inherently more objective nature of OCT measurements.

Results

Data was retrospectively collected from the Bascom Palmer Glaucoma Repository (BPGR). The dataset comprised of 47,003 SAP tests from 5,402 eyes of 3,643 patients and 25,480 peripapillary OCT tests from 4,125 eyes of 2,479 patients were included. Supplementary tables S1 and S2 list baseline demographic and clinical characteristics of the cohorts used in this study. SAP patients were followed for a mean (SD) of 9.6 (4.6) years with a mean (SD) number of 8.7 (4.1) SAP tests. From the 5,402 eyes, 1,171 (21.7%) were from 793 subjects self-identified as Black or African American and 4,231 (78.3%) from 2,850 self-identified as White; 1,778 (32.9%) were from 1,182 subjects self-identified as Hispanic or Latino and 3,624 (67.1%) from 2,461 self-identified as Not Hispanic or Latino. OCT patients were followed up for a mean (SD) of 7.0 (2.3) years with a mean (SD) number of 6.2 (1.4) OCT tests. From the 4,125 eyes, 700 (17.0%) were from 416 subjects self-identified as Black or African American and 3,425 (83.0%) from 2,063 self-identified as White; 1,009 (24.5%) were from 584 subjects self-identified as Hispanic or Latino and 3,116 (75.5%) from 1,895 self-identified as Not Hispanic or Latino.

The average rates of change for SAP MD over time were statistically significantly faster in White compared to Black or African American subjects (mean [SD], -0.27 [0.80] versus -0.18 [0.77] dB/year; $P=0.004$) and in Not Hispanic or Latino compared to Hispanic or Latino subjects (-0.28 [0.75] versus -0.19 [0.87] dB/year; $P=0.001$). For OCT retinal nerve fiber layer (RNFL) thickness, average rates of change were statistically significantly faster in Black or African American compared to White subjects (-0.5 [1.0] versus -0.4 [0.9] $\mu\text{m}/\text{year}$; $P=0.005$) and in Hispanic or Latino compared to Not Hispanic or Latino subjects (-0.5 [0.9] versus -0.4 [0.9] $\mu\text{m}/\text{year}$; $P=0.001$).

Test variability and race

Ordinary least squares linear regression models of SAP MD and OCT global RNFL thickness over time were fit to the sequence of perimetry and tomography tests for each eye of each individual from each one of racial and ethnic groups. The residuals from the trend lines were calculated and the SD of the residuals was used as an estimate of test-retest variability. The average SD of the residuals was statistically significantly greater in eyes of Black or African American individuals compared to those from White individuals for SAP (1.80 [1.30] versus 1.56 [1.21] dB, respectively; $P<0.001$) and also for OCT (2.3 [1.5] versus 2.1 [1.3] μm , respectively; $P<0.001$) (Table 1). The association between race and visual field variability remained statistically significant in the multivariable model ($P<0.001$, joint Wald test) (Table 2). There was a statistically significant interaction between race and severity of perimetric loss on visual field test variability, as seen in Table 2 by the coefficients associated with the interaction terms between race and MD splines ($P<0.001$, joint Wald test for interaction terms). The impact of this non-linear interaction, modeled by splines, is better visualized in Fig. 1A. The

| Parameter | Mean | SD | Median | p5 | p15 | p25 | p75 | p85 | p95 | P value |
|---------------------------|------|------|--------|------|------|------|------|------|------|---------|
| SAP (MD), by race | | | | | | | | | | |
| Black or African American | 1.80 | 1.30 | 1.45 | 0.54 | 0.78 | 0.96 | 2.17 | 2.83 | 4.30 | <0.001 |
| White | 1.56 | 1.21 | 1.25 | 0.47 | 0.70 | 0.85 | 1.84 | 2.33 | 3.79 | |
| SAP (MD), by ethnicity | | | | | | | | | | |
| Hispanic or Latino | 1.81 | 1.46 | 1.40 | 0.51 | 0.76 | 0.94 | 2.13 | 2.78 | 4.86 | <0.001 |
| Not Hispanic or Latino | 1.52 | 1.10 | 1.24 | 0.47 | 0.69 | 0.84 | 1.81 | 2.29 | 3.49 | |
| OCT (RNFL), by race | | | | | | | | | | |
| Black or African American | 2.3 | 1.5 | 2.0 | 0.8 | 1.2 | 1.4 | 2.8 | 3.4 | 4.9 | <0.001 |
| White | 2.1 | 1.3 | 1.8 | 0.7 | 1.1 | 1.3 | 2.6 | 3.1 | 4.3 | |
| OCT (RNFL), by ethnicity | | | | | | | | | | |
| Hispanic or Latino | 2.1 | 1.2 | 1.8 | 0.7 | 1.0 | 1.3 | 2.5 | 3.0 | 4.2 | 0.029 |
| Not Hispanic or Latino | 2.2 | 1.3 | 1.9 | 0.7 | 1.1 | 1.4 | 2.6 | 3.1 | 4.5 | |

Table 1. Summary statistics for the standard deviation (SD) of the residuals for standard automated perimetry (SAP) mean deviation and optical coherence tomography (OCT) retinal nerve fiber layer thickness, by race and ethnicity. OCT optical coherence tomography, SAP standard automated perimetry, SD standard deviation.

| Parameter | Coefficient | 95% CI | | P value |
|--------------------------------------|-------------|--------|--------|---------|
| By race | | | | |
| Black or African American race | 0.783 | 0.371 | 1.196 | <0.001 |
| Mean SAP MD spline 1, per 1 dB lower | 0.068 | 0.028 | 0.109 | 0.001 |
| Mean SAP MD spline 2, per 1 dB lower | -0.045 | -0.091 | 0.001 | 0.058 |
| Mean SAP MD spline 3, per 1 dB lower | 0.951 | 0.419 | 1.482 | <0.001 |
| Race × mean SAP MD spline 1 | -0.064 | -0.130 | 0.002 | 0.057 |
| Race × mean SAP MD spline 2 | 0.119 | 0.039 | 0.199 | 0.004 |
| Race × mean SAP MD spline 3 | -1.851 | -2.961 | -0.741 | 0.001 |
| Baseline age, per 10 years | -0.003 | -0.036 | 0.030 | 0.862 |
| Follow-up duration, per 1 year | 0.024 | 0.015 | 0.033 | <0.001 |
| By ethnicity | | | | |
| Hispanic or Latino ethnicity | 0.180 | 0.014 | 0.345 | 0.033 |
| Mean SAP MD spline 1, per 1 dB lower | 0.016 | -0.006 | 0.039 | 0.151 |
| Mean SAP MD spline 2, per 1 dB lower | 0.032 | 0.013 | 0.050 | 0.001 |
| Ethnicity × mean SAP MD spline 1 | 0.054 | 0.008 | 0.100 | 0.022 |
| Ethnicity × mean SAP MD spline 2 | -0.037 | -0.077 | 0.003 | 0.068 |
| Baseline age, per 10 years | -0.003 | -0.034 | 0.028 | 0.846 |
| Follow-up duration, per 1 year | 0.023 | 0.014 | 0.032 | <0.001 |

Table 2. Results of multivariable regression models evaluating the association of race and ethnicity with standard automated perimetry mean deviation variability (standard deviation [SD] of the residuals) adjusting for covariates. *CI* confidence interval, *SAP* standard automated perimetry, *MD* mean deviation, *SD* standard deviation.

difference in visual field variability between Black or African American and White eyes initially increased as visual field damage worsened, reaching its peak at an MD of approximately -12.0 dB, but decreased as the visual field damage became more advanced.

For OCT, race was also associated with greater test variability in the multivariable model ($P=0.002$, joint Wald test). The relationship between severity, as measured by mean RNFL thickness throughout follow-up, and variability was linear across the spectrum of disease severity (Fig. 2A). There was no statistically significant interaction between race and RNFL thickness measurement on test variability, as indicated in Table 3 ($P=0.129$).

Test variability and ethnicity

For SAP, the average SD of the residuals was statistically significantly greater in Hispanic or Latino compared to Not Hispanic or Latino subjects for MD (1.81 [1.46] versus 1.52 [1.10] dB; $P<0.001$) (Table 1). Ethnicity remained statistically significantly associated in the multivariable model ($P<0.001$, joint Wald test) (Table 2). There was an interaction between ethnicity and disease severity on test variability. This is seen in Table 2 by the coefficients associated with the interaction terms between ethnicity and MD splines ($P<0.001$, joint Wald test) (Fig. 1B). The difference in visual field variability between Hispanic or Latino and Not Hispanic or Latino subjects initially increased with worse visual field damage, with the greatest difference seen at an MD of approximately -11.0 dB, but decreased as the visual field damage became more advanced.

For OCT, the average SD of the residuals was statistically significantly greater in Not Hispanic or Latino compared to Hispanic or Latino subjects (2.2 [1.3] versus 2.1 [1.2] μm ; $P=0.029$) (Fig. 2B). Ethnicity remained statistically significant in the multivariable model ($P=0.005$, joint Wald test). There was also no statistically significant interaction between ethnicity and RNFL thickness measurement on OCT test variability ($P=0.280$) (Table 3).

Time to detect progression

From data on test variability, we simulated various scenarios to estimate the difference in time to detect progression between eyes of each racial and ethnic group. For SAP, we assumed MD values of -5 dB and -10 dB at baseline, with true rates of change of -0.25 dB/year (slow), -0.5 dB/year (moderate) and -1.0 dB/year (fast), in an annual testing regimen. For OCT, we assumed baseline RNFL thickness values of 90 μm and 70 μm , with true rates of change of -0.5 $\mu\text{m}/\text{year}$ (slow), -1.0 $\mu\text{m}/\text{year}$ (moderate) and -2.0 $\mu\text{m}/\text{year}$ (fast), with annual testing.

Tables 4 and 5 report mean predicted times to detect progression and the difference in predicted times to detect progression achieving 80% power (when 80% of the progressing eyes would be detected as progressing) for the simulated scenarios. Overall, greater variability led to delayed detection of SAP progression in all simulated scenarios. For example, in the scenario of baseline MD of -10 dB and moderate progression (slope of -0.50 dB/year), the mean difference in time to detect 80% of progressing eyes between Black or African American and White subjects was 1.8 years. Using the same simulated scenario, the mean difference in time between Hispanic or Latino and Not Hispanic or Latino subjects was 1.5 years. In contrast, for OCT, differences in time to detect progression were much smaller and generally less than 1 year for all comparisons between races and ethnicities.

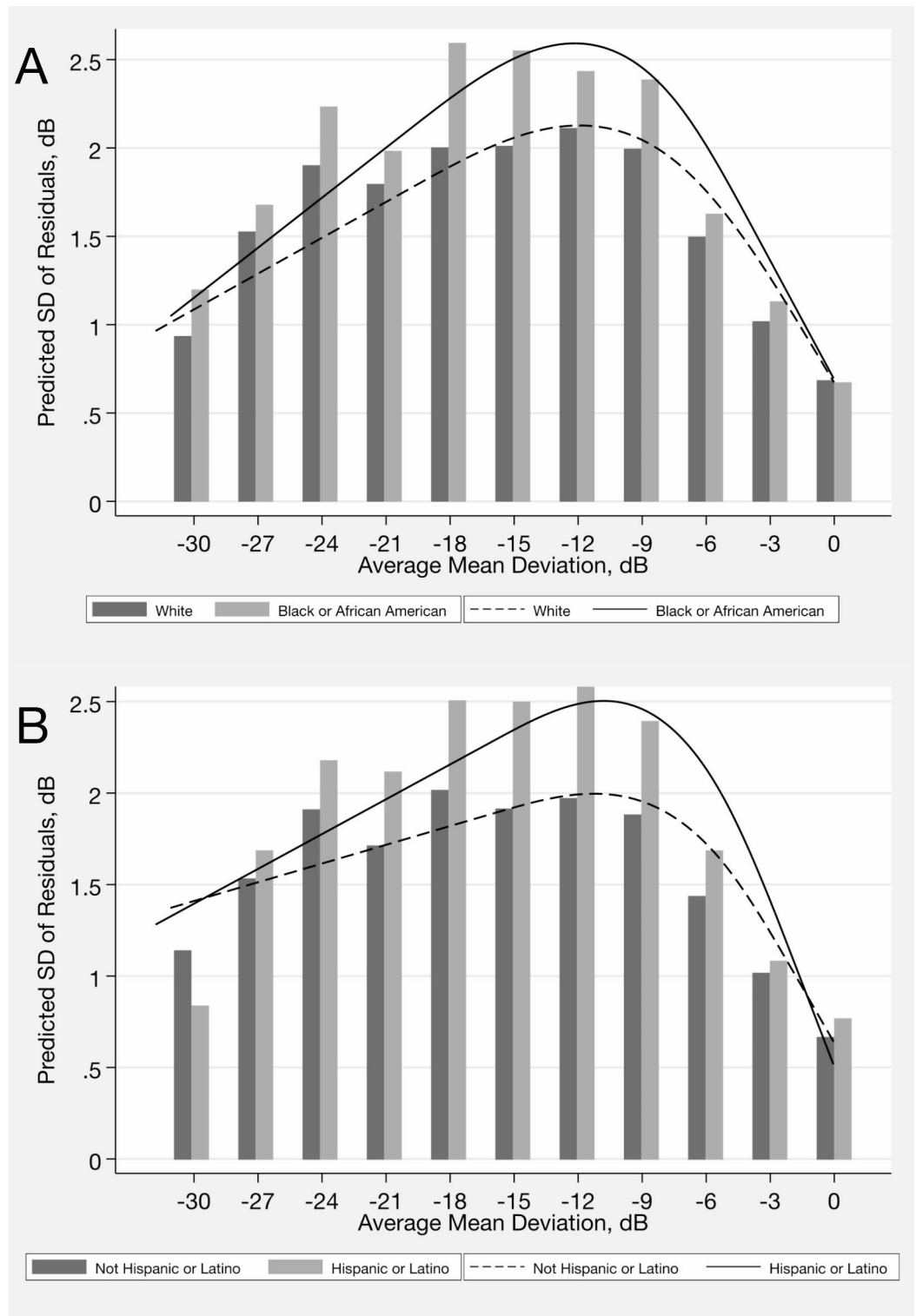


Fig. 1. Association between variability, as measured by the standard deviation of the residuals and visual field severity (average mean deviation during follow-up) by (A) race and by (B) ethnicity.

Discussion

In a large, diverse clinical sample, the current study confirms that Black or African American subjects with glaucoma exhibit greater SAP variability over time compared to White subjects and adds the new finding that Hispanic or Latino also exhibit increased SAP variability compared to their Not Hispanic or Latino counterparts. In contrast, structural testing with OCT revealed little differences in variability between the studied racial and ethnic groups. These disparities may result in delays in detecting progression when patients are monitored

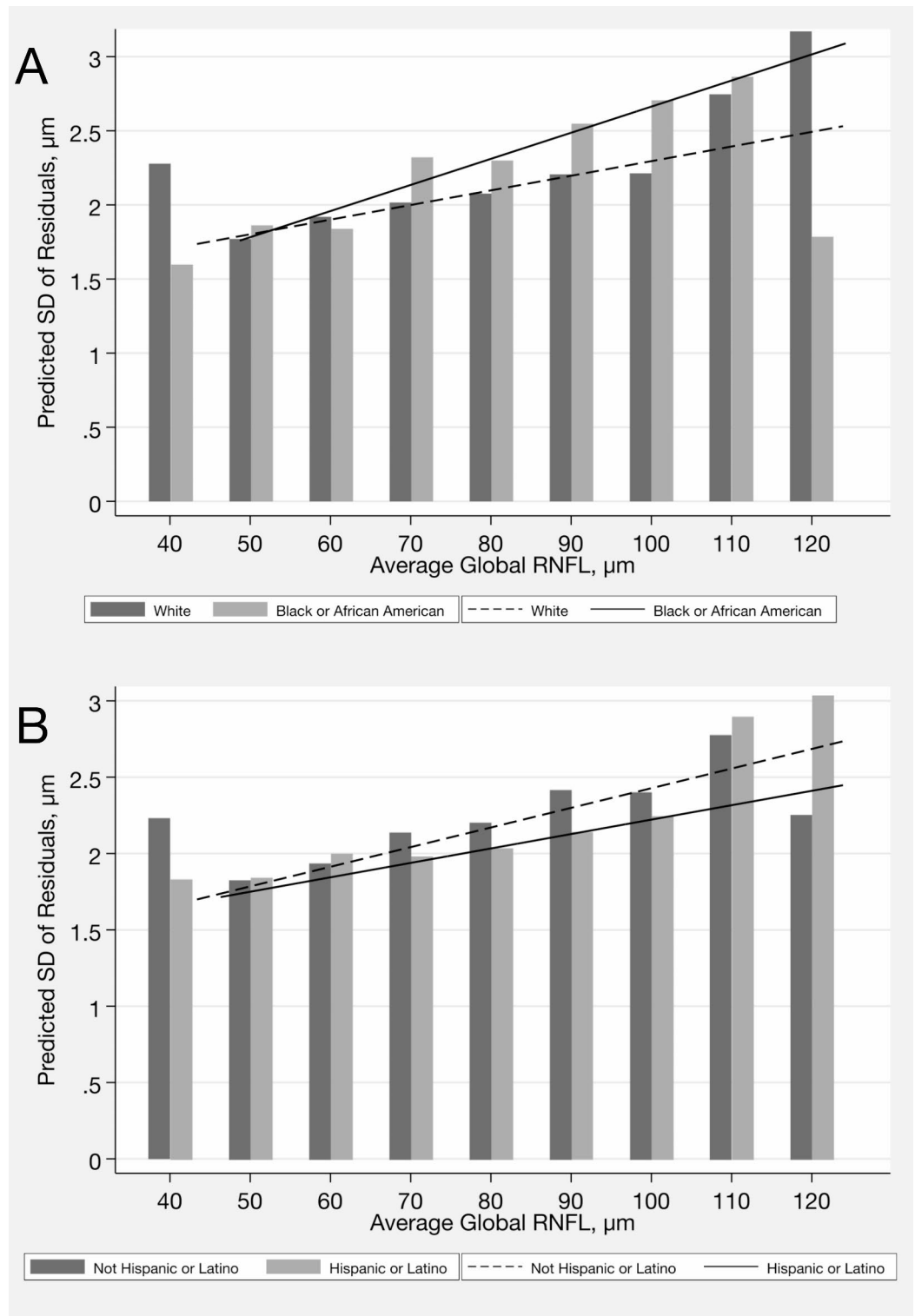


Fig. 2. Association between variability, as measured by the standard deviation of the residuals and optical coherence tomography severity (average global RNFL thickness during follow-up) by (A) race and by (B) ethnicity.

predominantly with perimetry, potentially contributing to the poorer clinical outcomes often observed in these minority groups. Such differences may be related to systemic biases in test administration procedures and deserve further investigation.

Our results are in agreement with previous findings from Gracitelli et al.¹¹, with data from participants enrolled in a multicenter prospective clinical study and Stagg et al.¹², with data from an EHR database. Both

| Parameter | Coefficient | 95% CI | | P value |
|---|-------------|---------|---------|---------|
| By race | | | | |
| Black or African American race | 0.193 | 0.075 | 0.311 | 0.001 |
| Mean OCT RNFL, per 1 μm higher | 0.012 | 0.008 | 0.016 | <0.001 |
| Race \times mean OCT RNFL | 0.007 | - 0.002 | 0.017 | 0.129 |
| Baseline age, per 10 years | 0.011 | - 0.032 | 0.054 | 0.630 |
| Follow-up duration, per 1 year | 0.031 | 0.013 | 0.050 | 0.001 |
| By ethnicity | | | | |
| Hispanic or Latino ethnicity | - 0.145 | - 0.243 | - 0.047 | 0.004 |
| Mean OCT RNFL, per 1 μm higher | 0.015 | 0.010 | 0.020 | <0.001 |
| Ethnicity \times mean OCT RNFL | - 0.004 | - 0.012 | 0.003 | 0.280 |
| Baseline age, per 10 years | - 0.013 | - 0.057 | 0.032 | 0.581 |
| Follow-up duration, per 1 year | 0.030 | 0.011 | 0.048 | 0.002 |

Table 3. Results of multivariable regression models evaluating the association of race and ethnicity with optical coherence tomography retinal nerve fiber layer thickness variability (standard deviation [SD] of the residuals) adjusting for covariates. *CI* confidence interval, *OCT* optical coherence tomography, *RNFL* retinal nerve fiber layer, *SD* standard deviation.

| Baseline disease severity, dB | “True” rates of change, dB/year | Mean time to detect progression, mean \pm SD, years | | Difference in time to detect progression in 80% of eyes (80% power), years |
|-------------------------------|---------------------------------|---|------------------------|--|
| | | Black or African American | White | Black or African American versus White Subjects |
| By race | | | | |
| - 5 | - 0.25 | 12.4 \pm 5.6 | 11.3 \pm 5.2 | 1.5 |
| | - 0.50 | 8.5 \pm 3.7 | 8.0 \pm 3.2 | 0.9 |
| | - 1.00 | 6.4 \pm 2.3 | 5.8 \pm 2.1 | 0.7 |
| - 10 | - 0.25 | 15.4 \pm 7.0 | 13.9 \pm 6.2 | 2.2 |
| | - 0.50 | 10.6 \pm 4.2 | 9.4 \pm 3.6 | 1.8 |
| | - 1.00 | 7.2 \pm 2.5 | 6.5 \pm 2.2 | 1.0 |
| Baseline disease severity, dB | “True” rates of change, dB/year | Mean time to detect progression, mean \pm SD, years | | Difference in time to detect progression in 80% of eyes (80% power), years |
| | | Hispanic or Latino | Not Hispanic or Latino | Hispanic or Latino versus not Hispanic or Latino |
| By ethnicity | | | | |
| - 5 | - 0.25 | 12.4 \pm 5.8 | 10.9 \pm 5.1 | 2.1 |
| | - 0.50 | 8.7 \pm 3.6 | 7.7 \pm 3.2 | 1.4 |
| | - 1.00 | 6.4 \pm 2.4 | 5.7 \pm 2.0 | 1.1 |
| - 10 | - 0.25 | 14.8 \pm 6.9 | 13.8 \pm 6.0 | 1.7 |
| | - 0.50 | 10.1 \pm 4.3 | 9.2 \pm 3.5 | 1.5 |
| | - 1.00 | 7.1 \pm 2.6 | 6.3 \pm 2.2 | 1.1 |

Table 4. Time to detect progression according to different scenarios of visual field loss over time, by race and ethnicity, assuming annual testing. *SD* standard deviation.

showed increased visual field variability and delayed detection of glaucomatous progression in Black or African American compared to White subjects. Additionally, our research provides novel insights by examining both race and ethnicity, in contrast to previous studies that focused solely on race. Few studies have specifically addressed the Hispanic or Latino population, a rapidly growing segment of the US population^{13,14}. Our results suggest potential ethnic differences in perimetric variability, uncovering a previously unexplored aspect in functional damage assessment. In previous studies, the average SD of the residuals were 1.45 versus 1.12 dB (mean difference: 0.33; $P < 0.001$)¹¹ and 1.53 versus 1.26 dB (mean difference: 0.27; $P < 0.001$)¹² for Black or African American versus White subjects, respectively. In the present study, we found 1.80 versus 1.56 dB (mean difference: 0.24; $P < 0.001$) for Black or African American versus White subjects and 1.81 versus 1.52 dB (mean difference: 0.29; $P < 0.001$) for Hispanic or Latino versus Not Hispanic or Latino subjects.

The ability to distinguish true change (the ‘signal’) from test-retest variability (the ‘noise’) is crucial for proper disease progression assessment. In our investigation, we observed more pronounced racial and ethnic differences in variability for MD values falling within the range of -10 to -15 dB. For this range of defect, typically classified as moderate or advanced disease, perimetric variability was approximately 25% greater in Black or

| Baseline disease severity, μm | “True” rates of change, $\mu\text{m}/\text{year}$ | Mean time to detect progression, mean \pm SD, years | | Difference in time to detect progression in 80% of eyes (80% power), years |
|--|---|---|------------------------|--|
| | | Black or African American | White | Black or African American versus White Subjects |
| By race | | | | |
| 90 | - 0.5 | 10.2 \pm 5.3 | 10.0 \pm 5.3 | 0.2 |
| | - 1.0 | 6.6 \pm 2.3 | 6.4 \pm 2.2 | 0.3 |
| | - 2.0 | 4.7 \pm 1.4 | 4.4 \pm 1.4 | 0.3 |
| 70 | - 0.5 | 10.0 \pm 5.6 | 9.6 \pm 5.7 | 0.2 |
| | - 1.0 | 6.4 \pm 2.1 | 5.9 \pm 1.8 | 0.7 |
| | - 2.0 | 4.4 \pm 1.2 | 4.1 \pm 1.2 | 0.3 |
| Baseline disease severity, μm | “True” rates of change, $\mu\text{m}/\text{year}$ | Mean time to detect progression, mean \pm SD, years | | Difference in time to detect progression in 80% of eyes (80% power), years |
| | | Hispanic or Latino | Not Hispanic or Latino | Hispanic or Latino versus not Hispanic or Latino |
| By ethnicity | | | | |
| 90 | - 0.5 | 9.6 \pm 5.4 | 10.6 \pm 5.7 | - 1.3 |
| | - 1.0 | 6.0 \pm 2.0 | 6.5 \pm 2.2 | - 0.7 |
| | - 2.0 | 4.4 \pm 1.2 | 4.5 \pm 1.4 | - 0.3 |
| 70 | - 0.5 | 9.4 \pm 5.6 | 9.5 \pm 5.3 | 0.2 |
| | - 1.0 | 5.9 \pm 2.1 | 5.9 \pm 1.9 | 0.0 |
| | - 2.0 | 4.2 \pm 1.1 | 4.2 \pm 1.2 | 0.0 |

Table 5. Time to detect progression according to different scenarios of optical coherence tomography retinal nerve fiber layer thickness loss over time, by race and ethnicity, assuming annual testing. *SD* standard deviation.

African American and in Hispanic or Latino subjects (Fig. 1). In computer simulations, we found a difference of 1.8 years in the time to detect progression in eyes of Black or African American compared to White subjects and 1.5 years for Hispanic or Latino compared to Not Hispanic or Latino subjects for detection of moderate progressors with annual testing. Even for detecting fast progressors, the difference between groups generally remained above 1 year (Table 4). The delayed recognition of progression could lead to delayed initiation or escalation of treatment and, consequently, irreversible vision loss. Additionally, it could also give the patient a misleading reassurance that the disease has not advanced and result in loss to follow-up^{11,15}. Greater variability may also result in false-positive events of progression and unnecessary changes in treatment¹¹.

For OCT, our findings demonstrated that differences in test variability between the different racial and ethnic groups were small and unlikely to be of clinical relevance. Mean difference in SD of the residuals was only 0.2 μm between Black or African American versus White subjects and 0.1 μm for Not Hispanic or Latino versus Hispanic or Latino subjects. In a study by Melchior et al.¹⁶, OCT variability was also found to be similar between individuals of African and European descent. Accordingly, times required to detect progression showed small differences between these groups in the simulated scenarios (Table 5).

Implicit biases among healthcare professionals have been recognized as contributors to disparities in health outcomes for minority groups. Existing research indicates that these biases, often manifesting unconsciously, are difficult to control and potentially impact patient-provider interactions and overall healthcare outcomes¹⁷. Perimetry, being a subjective test, is highly dependent on proper test instruction and supervision by the perimetrist or technician. The accuracy of perimetry can be influenced by the manner in which instructions are given, the patience and attentiveness of the technician and their ability to ensure that the patient remains focused and understands the test procedure. Any lapses in these areas can significantly affect test outcomes, leading to higher variability in results. This is particularly critical for minority groups who may already face communication barriers or receive less thorough instructions due to unconscious biases. Additionally, the subjective nature of perimetry requires the patient to respond consistently to visual stimuli, which can be influenced by their understanding of the test and the quality of interaction with the test administrator. Inconsistent or poor-quality interactions can lead to increased variability in test results. Conversely, OCT is an objective testing method that may be less susceptible to these biases. OCT provides quantitative data that does not rely on the patient's subjective responses or the quality of interaction with the test administrator. To investigate these biases further, studies could employ interviews of patients, test administrators and providers. Additionally, examining the training and interactions of healthcare professionals with minority patients could shed light on specific areas where biases may influence test administration and outcomes. By identifying these factors, targeted interventions can be developed to enhance training programs and improve the quality of patient-provider interactions, ultimately reducing health disparities.

It is crucial to emphasize that the results of this investigation should not be interpreted as indicating superiority of OCT over SAP in monitoring glaucoma progression. Previous studies have demonstrated significant discrepancies between structural and functional testing in detecting clinically relevant progression

over time, highlighting the necessity of both modalities for comprehensive patient monitoring. In fact, some eyes may show progression detectable by SAP but not by OCT, and vice versa^{18,19}.

Our study has limitations. The classification of race and ethnicity was self-reported by the subjects in the study. Race and ethnicity are dynamic entities shaped by cultural, geographic and sociopolitical aspects. However, the omission of these crucial factors in health and medical research dismisses the reality of social stratification, injustices and inequities²⁰. Additionally, studies utilizing self-reported information have proven valuable if this data is acquired in a standardized manner²¹. We were also not able to examine intersections of these categories or other subgroups due to sample size constraints. Future studies should investigate the impact of intersectionalities on differences in variability and test performance. Of note, our study lacked data on certain social determinants, such as educational and socioeconomic background. However, previous studies have shown that the association between race and test variability persisted even after adjusting for these factors, suggesting these variables alone do not fully account for the observed differences^{11,12}. As another limitation, the evaluation of visual field and OCT variability was based, exclusively, on trend analysis of MD and RNFL global thickness over time. Alternative methods for detecting change, including localized loss assessment and event-based approaches are also available. Nevertheless, it is conceivable that localized assessment could be even more susceptible to the impact of increased variability. Additionally, our study did not collect data on the number of tests excluded due to quality issues for each racial and ethnic group. Future studies should aim to gather such data to better understand potential differences in test quality across groups and to assess how these tests perform in diverse populations.

In conclusion, our results revealed that variability was greater in perimetric testing for Black or African American compared to White subjects and for Hispanic or Latino compared to Not Hispanic or Latino subjects, resulting in longer times to detect progression in the groups with greater variability. Assessment with OCT, due to the inherently more objective nature of its measurements, revealed little differences between the studied racial and ethnic groups. These disparities may contribute, at least partially, to potential delays in detecting disease progression and to the poorer clinical outcomes frequently observed in these minority groups. Future studies should investigate the presence of systematic biases in test administration procedures.

Methods

The research protocol received ethical approval from the Institutional Review Board at the University of Miami School of Medicine. The need to obtain informed consent was waived by the University of Miami Institutional Review Board due to the retrospective nature of the study. Data were de-identified before being used for statistical analyses. The study adhered to the tenets of the Declaration of Helsinki and followed the Health Insurance Portability and Accountability Act.

Data collection

The BPGR contains demographic and ophthalmic data of eyes with glaucoma or suspicion of glaucoma examined at the Bascom Palmer Eye Institute or its satellite clinics. The large self-identified Black or African American, Hispanic or Latino populations in South Florida contribute to the diversity of this database. Patients were identified using International Classification of Diseases (ICD) codes at baseline. For eyes that met inclusion criteria, any instance of key ocular diagnoses that could substantially confound testing were identified. Tests performed after any diagnosis of age-related macular degeneration (atrophic, exudative or late-stage), amblyopia, choroidal or retinal tumors, non-glaucomatous disorders of the optic nerve and visual pathways, retinal detachment, retinal venous or arterial occlusions, uveitis and proliferative diabetic retinopathy according to ICD codes were excluded. In addition, eyes that underwent glaucoma procedures (trabeculectomy, aqueous shunt insertion, cyclophotocoagulation, laser iridotomy, or micro-invasive glaucoma surgeries) were identified and tests following any of these procedures were also eliminated due to the potential impact of surgical intervention on the rates of change over time. Inclusion and exclusion criteria have been described in detail elsewhere²².

SAP testing

Standard Automated Perimetry data were extracted from the Zeiss Forum (Carl Zeiss Meditec Inc, Dublin, CA). All data between April 1997 and March 2022 were collected. Tests were performed using the 24–2 or 30–2 Swedish Interactive Threshold Algorithm (SITA) with size III white stimulus, from the Humphrey Visual Field Analyzer (HFA, versions II and III; Carl Zeiss Meditec Inc, Dublin, CA). Visual fields were excluded if they had fixation losses $\geq 33\%$ or false-positives $\geq 15\%$. Eyes included in this study were required to have confirmed glaucomatous field loss at baseline, based on the presence of repeatable (at least 2 consecutive) abnormal test results defined as a pattern standard deviation at a $P < 5\%$ or worse, or a glaucoma hemifield test result of “outside normal limits”. Eyes were also required to have at least 5 reliable SAP tests and a minimum of 1 year of follow-up.

OCT testing

Optical Coherence Tomography data from the Zeiss Cirrus system were extracted from Zeiss Forum (Carl Zeiss Meditec Inc, Dublin, CA). All data between April 2008 and February 2022 were collected. All scans were required to have signal strength $\geq 7/10$. Any scans with global RNFL thickness values $< 30 \mu\text{m}$ or $> 130 \mu\text{m}$ were excluded^{22,23}. Baseline global RNFL thickness was required to be at least $38 \mu\text{m}$ in order to allow for longitudinal trend given the “floor effect”²⁴. If multiple scans were performed on the same day, the mean of RNFL thickness values from the same day was utilized. Eyes included in this study were required to have at least 5 reliable OCT tests and a minimum of 1 year of follow-up.

Data analyses and computer simulations

This study focused on comparisons of longitudinal test-retest variability and time to detect progression between two racial groups, Black or African American versus White, as well as between two ethnic groups, Hispanic or Latino versus Not Hispanic or Latino. Ordinary least squares linear regression models of SAP MD and OCT global RNFL thickness over time were fit to the sequence of perimetry and tomography tests for each eye of each individual from each one of these groups. The residuals from the trend lines were calculated and the standard deviation (SD) of the residuals was used as an estimate of test-retest variability. This approach has been previously described^{11,12}. The SD of the residuals was compared between the 2 racial and the 2 ethnic groups using Generalized estimating equations with robust sandwich variance estimator²⁵, to account for correlations between two eyes of the same subject. We then evaluated the association of race and ethnicity with the SD of the residuals in multivariable models adjusting for baseline age and duration of follow-up. Since the relationship between variability and race/ethnicity could potentially be influenced by disease severity, we also incorporated variables indicative of severity during follow-up period (average SAP MD and average OCT RNFL thickness), as well as the interactions between these variables with categorical indicators of race or ethnicity. Since the association between MD variability and visual field sensitivity is nonlinear, it was modeled using restricted cubic splines^{26,27}, with the number of knots determined by cross-validation, replicating our previous approach^{26,27}. For RNFL thickness variability, this was not necessary and a simple interaction term was used.

We then used computer simulations to estimate time to detect SAP and OCT progression in the different racial and ethnic groups. The ordinary least squares residuals of MD and RNFL trends over time obtained from the original cohort were binned according to fitted levels of defect for each parameter (MD and RNFL). Empirical distributions of the residuals were then available for each level of MD and RNFL thickness, allowing reconstruction of trajectories of change over time by computer simulations, for expected “true” rates of glaucoma progression. A similar approach has been described previously for both SAP^{11,28,29} and OCT testing^{16,22,30}. Given a “true” MD or RNFL thickness value, the empirical distributions of MD or RNFL residuals contain the range of measured values that would be expected for each given test. Longitudinal sequences of SAP and OCT tests were then simulated by assuming a “true” baseline MD or baseline RNFL thickness, a “true” rate of change for each test and then sampling from the empirical distributions of the residuals to reconstruct what the test MD or RNFL thickness would be at each time. For example, assuming a “true” baseline MD of -5 dB and an annual rate of change of -1 dB/year, “true” MD measurements would be -5, -6, -7, -8 and -9 dB in the first 4 years of follow-up. Similarly, assuming a “true” baseline RNFL thickness of 90 μm and an annual rate of change of -2 $\mu\text{m}/\text{year}$, “true” RNFL thickness measurements would be 90, 88, 86, 84 and 82 μm in the first 4 years of follow-up. However, testing data are affected by noise, which in our simulations was added to the “true” values by sampling from the empirical distributions of the residuals for each corresponding level of MD or RNFL thickness. For example, simulated perimetry measurements for the described situation could be -5.3, -4.9, -7.5, -8.6 and -7.9 dB for the first 4 years of follow-up. Likewise, simulated OCT measurements for the could be 91.4, 87.7, 89.5, 91.0 and 89.2 μm . Testing data were simulated for each racial and ethnic group, taking into account specific empirical distributions of the residuals. We simulated 1,000 sequences of each test by each racial and ethnic group, assuming equivalent fixed-test intervals for each group. We then obtained the earliest time to detect progression for each testing sequence in each group.

For SAP, progression was defined as a statistically significant negative slope of MD over time ($P < 0.05$). For OCT, progression was defined as a statistically significant slope over time ($P < 0.05$), with a slope more negative than $-0.5 \mu\text{m}/\text{year}$, to account for age-related loss³⁰. This allowed us to construct cumulative probability functions of time to detect progression for each racial and ethnic group and estimate differences in time to detect progression under specific testing scenarios.

All statistical analyses and computer simulations were performed using Stata Version 18 (StataCorp LP, College Station, TX, USA).

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request. Access to the data is restricted due to privacy and ethical considerations.

Received: 5 September 2024; Accepted: 19 November 2024

Published online: 28 December 2024

References

- Weinreb, R. N., Aung, T. & Medeiros, F. A. The pathophysiology and treatment of glaucoma: a review. *JAMA* **311**, 1901–1911. <https://doi.org/10.1001/jama.2014.3192> (2014).
- Barquet-Piza, V. & Siegfried, C. J. Understanding racial disparities of glaucoma. *Curr. Opin. Ophthalmol.* **35**, 97–103. <https://doi.org/10.1097/ICU.0000000000001017> (2023).
- Tielsch, J. M. Racial variations in the prevalence of primary open-angle Glaucoma. *Jama* **266**, 369–374. <https://doi.org/10.1001/jama.1991.03470030069026> (1991).
- Leske, M. C., Connell, A. M., Schachat, A. P. & Hyman, L. The Barbados Eye Study. Prevalence of open angle glaucoma. *Arch. Ophthalmol.* **112**, 821–829. <https://doi.org/10.1001/archophth.1994.01090180121046> (1994).
- Buhrmann, R. R. et al. Prevalence of Glaucoma in a Rural East African Population. *Investig. Ophthalmol. Vis. Sci.* **41**, 40–48 (2000).
- Rotchford, A. P., Kirwan, J. F., Muller, M. A., Johnson, G. J. & Roux, P. Temba glaucoma study: a population-based cross-sectional survey in urban South Africa. *Ophthalmology* **110**, 376–382. [https://doi.org/10.1016/S0161-6420\(02\)01568-3](https://doi.org/10.1016/S0161-6420(02)01568-3) (2003).
- Sommer, A. et al. Racial differences in the cause-specific prevalence of blindness in east Baltimore. *N Engl. J. Med.* **325**, 1412–1417. <https://doi.org/10.1056/NEJM199111143252004> (1991).
- Munoz, B. et al. Causes of blindness and visual impairment in a population of older americans: the Salisbury Eye evaluation study. *Arch. Ophthalmol.* **118**, 819–825. <https://doi.org/10.1001/archophth.118.6.819> (2000).

9. Racette, L. et al. African descent and Glaucoma evaluation study (ADAGES): III. Ancestry differences in visual function in healthy eyes. *Arch. Ophthalmol.* **128**, 551–559. <https://doi.org/10.1001/archophthalmol.2010.58> (2010).
10. Swaminathan, S. S. & Medeiros, F. A. Socioeconomic disparities in Glaucoma severity at initial diagnosis: a Nationwide EHR Cohort Analysis. *Am. J. Ophthalmol.* **263**, 50–60. <https://doi.org/10.1016/j.ajo.2024.02.022> (2024).
11. Gracitelli, C. P. B. et al. Detection of Glaucoma progression in individuals of African descent compared with those of European descent. *JAMA Ophthalmol.* **136**, 329–335. <https://doi.org/10.1001/jamaophthalmol.2017.6836> (2018).
12. Stagg, B. et al. Longitudinal visual field variability and the ability to detect glaucoma progression in black and white individuals. *Br. J. Ophthalmol.* **106**, 1115–1120. <https://doi.org/10.1136/bjophthalmol-2020-318104> (2022).
13. Quigley, H. A. et al. The prevalence of glaucoma in a population-based study of hispanic subjects: Proyecto VER. *Arch. Ophthalmol.* **119**, 1819–1826. <https://doi.org/10.1001/archophth.119.12.1819> (2001).
14. Varma, R. et al. Prevalence of open-angle glaucoma and ocular hypertension in latinos: the Los Angeles latino Eye Study. *Ophthalmology* **111**, 1439–1448. <https://doi.org/10.1016/j.ophtha.2004.01.025> (2004).
15. Wang, F., Javitt, J. C. & Tielsch, J. M. Racial variations in treatment for glaucoma and cataract among Medicare recipients. *Ophthalmic Epidemiol.* **4**, 89–100. <https://doi.org/10.3109/09286589709057101> (1997).
16. Melchior, B. et al. Frequency of Optical Coherence Tomography Testing to detect progression in Glaucoma. *J. Glaucoma.* **31**, 854–859. <https://doi.org/10.1097/IJG.0000000000002101> (2022).
17. Hall, W. J. et al. Implicit Racial/Ethnic Bias among Health Care Professionals and its influence on Health Care outcomes: a systematic review. *Am. J. Public Health.* **105**, 60–76. <https://doi.org/10.2105/AJPH.2015.302903> (2015).
18. Banegas, S. A. et al. Agreement among spectral-domain optical coherence tomography, standard automated perimetry, and stereophotography in the detection of glaucoma progression. *Invest. Ophthalmol. Vis. Sci.* **56**, 1253–1260. <https://doi.org/10.1167/iovs.14-14994> (2015).
19. Medeiros, F. A., Zangwill, L. M., Bowd, C., Mansouri, K. & Weinreb, R. N. The structure and function relationship in glaucoma: implications for detection of progression and measurement of rates of change. *Invest. Ophthalmol. Vis. Sci.* **53**, 6939–6946. <https://doi.org/10.1167/iovs.12-10345> (2012).
20. Flanagan, A., Frey, T., Christiansen, S. L. & Committee, A. M. A. M. o. S. updated Guidance on the reporting of race and ethnicity in Medical and Science Journals. *JAMA* **326**, 621–627. <https://doi.org/10.1001/jama.2021.13304> (2021).
21. Rosenberg, N. A. et al. Genetic structure of human populations. *Science* **298**, 2381–2385. <https://doi.org/10.1126/science.1078311> (2002).
22. Swaminathan, S. S., Berchuck, S. I., Rao, J. S. & Medeiros, F. A. Performance of Linear mixed models in estimating structural rates of Glaucoma progression using Varied Random Effect distributions. *Ophthalmol. Sci.* **4**, 100454. <https://doi.org/10.1016/j.xops.2023.100454> (2024).
23. Jammal, A. A. et al. Rates of Glaucomatous Structural and Functional Change from a large Clinical Population: the Duke Glaucoma Registry Study. *Am. J. Ophthalmol.* **222**, 238–247. <https://doi.org/10.1016/j.ajo.2020.05.019> (2021).
24. Bowd, C., Zangwill, L. M., Weinreb, R. N., Medeiros, F. A. & Belghith, A. Estimating Optical Coherence Tomography Structural Measurement floors to improve detection of progression in advanced Glaucoma. *Am. J. Ophthalmol.* **175**, 37–44. <https://doi.org/10.1016/j.ajo.2016.11.010> (2017).
25. Zeger, S. L. & Liang, K. Y. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* **42**, 121–130 (1986).
26. Buse, A. & Lim, L. Cubic splines as a special case of restricted least squares. *J. Am. Stat. Assoc.* **72**, 64–68. <https://doi.org/10.1080/01621459.1977.10479907> (1977).
27. Smith, P. L. Splines as a useful and convenient statistical Tool. *Am. Stat.* **33**, 57–62. <https://doi.org/10.1080/00031305.1979.10482661> (1979).
28. Wu, Z., Saunders, L. J., Daga, F. B., Diniz-Filho, A. & Medeiros, F. A. Frequency of testing to detect Visual Field Progression Derived using a longitudinal cohort of Glaucoma patients. *Ophthalmology* **124**, 786–792. <https://doi.org/10.1016/j.ophtha.2017.01.027> (2017).
29. Swaminathan, S. S., Berchuck, S. I., Jammal, A. A., Rao, J. S. & Medeiros, F. A. Rates of Glaucoma progression derived from Linear mixed models using Varied Random Effect distributions. *Transl Vis. Sci. Technol.* **11**, 16–27. <https://doi.org/10.1167/tvst.11.2.16> (2022).
30. Wu, Z. et al. Impact of normal aging and progression definitions on the specificity of detecting retinal nerve Fiber layer thinning. *Am. J. Ophthalmol.* **181**, 106–113. <https://doi.org/10.1016/j.ajo.2017.06.017> (2017).

Author contributions

F.A.M. conceived and designed the study described; L.A.F.B., A.A.J., D.R.C., E.B.M., S.S.S., and F.A.M. collected and managed the data; L.A.F.B., A.A.J., and F.A.M. performed data analyses and interpretation; L.A.F.B., A.A.J., D.R.C., E.B.M., S.S.S., and F.A.M. drafted and revised the manuscript prior to submission.

Funding

Supported in part by National Institutes of Health (NIH)/National Eye Institute (NEI) grant EY029885 (FAM), K23 EY033831 (SSS), American Glaucoma Society Mentoring for the Advancement of Physician Scientists Award (SSS), and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Brazil) Finance Code 001 (LAFB). The funding organizations had no role in the design or conduct of this research.

Declarations

Competing interests

L.A.F.B.: none. A.A.J.: none. D.R.C.: none. E.B.M.: none. S.S.S.: Heidelberg Engineering (S), Lumata Health (C, E), Abbvie (C); Topcon (C). F.A.M.: AbbVie (C), Annexon (C); Carl Zeiss Meditec (C), Galimedix (C); Google Inc. (F); Heidelberg Engineering (F), nGoggle Inc. (P), Novartis (F); ONL Therapeutics (C), Stealth Biotherapeutics (C); Stuart Therapeutics (C), Thea Pharmaceuticals (C), Reichert (C, F).

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-80481-4>.

Correspondence and requests for materials should be addressed to F.A.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024