



HHS Public Access

Author manuscript

Am J Ophthalmol. Author manuscript; available in PMC 2018 April 01.

Published in final edited form as:

Am J Ophthalmol. 2017 April ; 176: 148–156. doi:10.1016/j.ajo.2017.01.013.

Detecting change using standard global perimetric indices in glaucoma

Stuart K Gardiner* and Shaban Demirel

Devers Eye Institute, Legacy Research Institute, Portland, OR, USA

Abstract

Purpose—Various global indices are available to summarize results from standard automated perimetry. This study asks which index can detect significant deterioration earliest, for a fixed specificity.

Design—Comparison of prognostic indices.

Methods—Two cohorts were tested. A test-retest cohort contained 5 reliable visual fields, within a short interval, from 45 eyes of 23 participants with glaucoma and/or likelihood of developing glaucoma. A separate longitudinal cohort contained 508 eyes from 330 participants, tested on average 13 times. Three global indices were extracted; Mean Deviation (MD), Pattern Standard Deviation (PSD) and Visual Field Index (VFI). For each index we defined a critical p -value $\text{Crit}_{\text{Index}}$, such that 5% of test-retest series showed significant deterioration with $p < \text{Crit}_{\text{Index}}$, using artificial “test dates” in random order. Therefore these criteria have 95% specificity over series of 5 tests. The times to detect significant deterioration in the longitudinal cohort were compared using a survival analysis model.

Results—The median time to detect significant deterioration with MD was 7.3 years (95% confidence interval 6.8–7.9 years). For VFI, the median was 8.5 years (7.9–9.0); this comparison had $p=0.088$. For PSD, the median was 10.5 years (9.3–11.7); slower than MD with $p < 0.001$. Within the first five years of a series, MD detected significant deterioration in 138 eyes, versus 104 for VFI ($p=0.0013$) and 107 for PSD ($p=0.029$).

Conclusions—MD detected significant deterioration sooner than VFI or PSD. In particular, MD detected more eyes in the first five years of their follow-up, which were presumably undergoing more rapid progression.

Introduction

Functional measurement remains an essential part of the standard of care for patients with glaucoma. Despite advances in imaging technology, there is considerable variation in function for eyes with the same structural measurements.^{1, 2} While the rate of retinal nerve

*Corresponding Author Address: Dr Stuart Gardiner, Devers Eye Institute, Legacy Research Institute, 1225 NE 2nd Ave, Portland, OR 97232, USA, Tel: +1 503 413 1199, sgardiner@deverseye.org.

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

fiber layer thinning does correlate with the rate of functional change,³ the ability to detect glaucomatous progression using structural measurements alone has been reported to be limited.⁴⁻⁶ Since function is an important outcome measure for the patient because it relates to quality of life and activities of daily living, perimetry will remain a necessary clinical management tool for the foreseeable future.⁷ Most commonly, this is performed using white-on-white standard automated perimetry (SAP).

SAP produces estimates of the pointwise sensitivity at different locations in the visual field, with various test patterns being available on different instruments. “Contrast sensitivity” is conventionally defined in clinical perimetry as the contrast level of a visual stimulus that the observer will respond to (by pressing a button) on 50% of stimulus presentations, and is generally reported in decibels (dB; 1dB = 0.1 log units contrast) either as raw sensitivity or as deviation from age-matched normal. It is common practice to then summarize this information in the form of one or more global indices. This facilitates interpretation, not least by reducing the substantial variability present in pointwise measurements. In addition, global indices are often used to assess the rate of progression in an eye. Indeed, many translational and clinical studies have used the rate of change of global indices as their main outcome measures.⁸⁻¹²

The software that accompanies the Humphrey Field Analyzer (HFA; Carl Zeiss Meditec Inc, Dublin, CA, USA) currently produces three quantitative global indices (the Glaucoma Hemifield Test is not included here because it categorizes eyes by damage level, rather than providing a continuous quantification). Mean Deviation (MD) represents a centrally-weighted average of the pointwise deviations from normal, and is designed to quantify the overall or generalized level of sensitivity loss averaged across the visual field.¹³ It becomes more negative with severity of visual field loss, with MD=0dB being the expected average normative value for the patient’s age. Pattern Standard Deviation (PSD) represents the interlocation variability, by subtracting the MD from the deviation value at each location (the difference from normal), and is designed to quantify localized changes such as those that are typically observed in glaucoma. It increases with severity in early and moderate stages of visual field damage, before becoming non-monotonic and decreasing again once MD is below around -20dB.¹⁴

More recently, the Visual Field Index (VFI) has been introduced. This index aims to reduce the effect that cataract may have on MD,¹⁵ although the magnitude of the effect of cataract has been questioned.¹⁶ VFI reflects the pattern deviation values restricted to only those locations that are outside normal limits.¹⁴ The VFI is expressed on a scale as “percent of normal vision”. It should be noted that “100%” does not necessarily correspond with completely normal vision, since visual field locations can retain contrast sensitivity that is “within normal limits” even in the presence of substantial retinal ganglion cell loss,¹⁷ making VFI suboptimal for detection of early disease.¹⁸ It should also be noted that “0%” does not exactly correspond with blindness, since 0dB is defined by the limits of the perimeter rather than any physiologic or pathophysiologic limit, and sensitivities below 15–19dB are too unreliable to be useful for most purposes.¹⁹ There is also a discontinuity in the way VFI is calculated in the presence of severe visual field loss,¹⁴ which can increase variability.²⁰ The rate of change of VFI is now incorporated in the HFA’s Guided

Progression Analysis software, and so it has become widely used both clinically and in recent research studies.^{21–23}

The VFI was explicitly designed to improve monitoring of visual field progression in eyes with cataract. These are an important sub-population, given the ages of patients with glaucoma encountered clinically. However, most patients encountered in clinics in developed countries have no worse than mild cataract, as moderate or severe cataracts are routinely removed. Therefore in this study, we ask which of these three global indices detects glaucomatous progression soonest in eyes without severe cataract. We focus on eyes with no more than moderate glaucomatous visual field loss. We aim to test the hypothesis that VFI is more sensitive for detecting progression than previous global indices, when all are compared at the same specificity.

Methods

Participants – Longitudinal Cohort

This was a retrospective cohort study comparing prognostic indices. Participants in the Portland Progression Project (P3) were recruited to a tertiary glaucoma clinic at Devers Eye Institute. Inclusion criteria were simply a diagnosis of primary open-angle glaucoma and/or likelihood of developing glaucomatous damage, as determined by each participant's physician. Exclusion criteria were an inability to perform reliable visual field testing, best-corrected visual acuity worse than 20/40, cataract or media opacities likely to significantly increase light scatter, or other conditions or medications that may affect the visual field. All protocols for this study were approved and monitored by the Legacy Health Institutional Review Board, and adhered to the Health Insurance Portability and Accountability Act of 1996 and the tenets of the Declaration of Helsinki. All participants provided written informed consent once all of the risks and benefits of participation were explained to them.

Participants in the longitudinal cohort are tested every six months with a variety of structural and functional tests. SAP was performed using an HFAII perimeter, using the 24-2 test pattern, a size III white-on-white stimulus, and the SITA Standard algorithm.²⁴ Only tests with <15% false positives were used. Tests with >33% fixation losses were included only if the technician performing and monitoring the test noted that fixation remained stable as observed on the instrument's operator screen, since fixation losses can erroneously be reported if the blind spot is inaccurately mapped at the start of the test. Since PSD becomes non-monotonic once MD is below approximately –20dB,¹⁴ such visual fields were also excluded from the analysis. For this study, only eyes with series of at least four reliable tests by these criteria were included in the analysis.

Participants – Test-retest Cohort

A test-retest cohort was also assessed. The same inclusion and exclusion criteria were used as for the longitudinal cohort. The testing protocol was also the same, except that participants were tested five times within a short time period (approximately 1–2 months), instead of every six months.

Analysis – Test-retest Variability

Visual field tests performed on participants in the test-retest cohort were assigned artificial “test dates” at six-month intervals, to match the typical inter-test interval in the longitudinal cohort. Then, the rates of change of each global index (MD, PSD and VFI) were calculated by linear regression. If the rate of change was in the direction of apparent worsening of the field over time (i.e. becoming more negative for MD and VFI, becoming more positive for PSD), then the significance of the rate of change (the p-value from a two-sided t-test) was recorded; otherwise, the series was assigned $p=1.000$.

This process was repeated for all 120 possible permutations of the five tests per eye. Hence, for each global index, there are $(120 * N)$ p-values, where N represents the number of eyes in the test-retest cohort. The 5th percentile of these p-values was defined as the “critical value” for that index, labeled $Crit_{MD}$, $Crit_{PSD}$ and $Crit_{VFI}$. Therefore, 5% of these artificial series for MD were worsening with a $p < Crit_{MD}$ in this test-retest dataset in which we know that no true change is likely to have occurred, and similarly for PSD and VFI, such that all three criteria have specificity equal to 95% in this dataset.

The primary analysis described here set the specificity at 95%, as this represents a reasonable compromise between specificity and sensitivity in many clinical situations. However, the analysis was also repeated setting specificity at 99%, 98%, 90%, 85% and 80%, since the relative performance of the indices could vary between criteria.

Analysis – Detection of Change

For each eye in the longitudinal cohort, change in each of the three global indices was assessed by linear regression over the first four visual fields in their series, then the first five fields, the first six fields etc. The eye was labeled as “significantly deteriorating” by MD on the earliest test date at which this slope was negative (consistent with worsening) with $p < Crit_{MD}$; if this criterion was never attained then the last visit in the series was recorded as the censoring date. The date at which PSD and VFI were first found to be significantly deteriorating were defined in the same manner, using $p < Crit_{PSD}$ and $p < Crit_{VFI}$ respectively. These dates were recorded as “years since baseline”, using the first test date in the series for an eye, to facilitate comparisons.

In order to explore the likelihood that change would be confirmed at the next visit, the analysis was repeated adding one more field to the series after the initial date at which the series showed “significant deterioration”, in order to see whether this extended series still met the criterion for “significant deterioration”. Only series with at least one such field available were included in this analysis. This exercise was also performed adding two extra fields to the series, in eyes for which they were available.

Analysis – Comparison of Global Indices

Kaplan-Meier survival curves were plotted to show how soon each index showed “significant deterioration” for eyes in the longitudinal cohort. From these curves, we extracted the median time (the first date at which 50% of eyes had shown significant deterioration), with 95% confidence interval using standard errors based on Greenwood’s

formula.²⁵ The statistical significance of differences between the survival curves for the three indices was assessed using a stratified Cox proportional hazards model,²⁶ with strata identifying clusters for fellow eyes of the same individual in order to adjust the p-values for inter-eye correlations. Including strata is equivalent to using a generalized estimating equation in a linear model to account for clustering and obtain more accurate p-values.²⁷ This process was repeated for each of the different specificities described above.

In order to test which global index has higher sensitivity for detecting rapid change, the survival analysis was repeated using only the first five years of the series for each eye, on the basis that more rapidly changing eyes should show “significant deterioration” earlier in their series. The analysis was also repeated after splitting the cohort into two equal parts based on whether the most recent MD in their series was below or above the median value in the cohort. All analyses in this study were performed using the R statistical programming language.²⁸

Even though eyes with severe cataract were excluded, it has been suggested that even mild cataract could cause a deterioration in MD. However, this would not be expected to affect PSD or VFI. Therefore, to ensure that our results comparing MD with VFI were not caused purely by the worsening of mild cataracts, two sub-analyses were performed. Firstly, a comparison between the indices was performed restricted to those eyes in the longitudinal cohort that demonstrated “significant deterioration” in PSD before the end of their series. Secondly, a comparison was performed censoring data series on the date that an eye either deteriorated by two lines of visual acuity from baseline, or underwent cataract surgery.

Results

The test-retest cohort contained series of 5 reliable visual fields from 45 eyes of 23 participants. 74% were female; 18 had MD (on the last visit) greater than 0dB; 87% were Caucasian. 20 had MD between 0dB and -6dB; and seven were worse than -6dB. 30 of these eyes had visual acuity 20/20 or better.

The longitudinal cohort consisted of 508 eyes from 330 participants, once unreliable fields and those with MD < -20dB had been excluded. 58% were female; 89% were Caucasian. 226 eyes had MD > 0dB; 229 had MD between 0dB and -6dB; and 53 were worse than -6dB. 378 of the eyes had visual acuity 20/20 or better on the date of their initial test. Table 1 summarizes other characteristics of the two cohorts.

Using all 120 permutations of the test order for each of the 45 eyes in the test-retest cohort gives 5400 series of visual fields. Of these, 3.1% had MD significantly decreasing with $p < 0.05$; i.e. “MD worsening with $p < 0.05$ ” would have 96.9% specificity. The equivalent criteria with $p < 0.05$ would give specificities of 98.1% for VFI, and 97.1% for PSD. Without equalizing these specificities (i.e. determining change based on the first time at which the index was deteriorating with $p < 0.05$), the median time to detect change using MD was 8.5 years (95% confidence interval 7.8 – 10.2), compared with 11.6 years (10.5 – 12.7) for VFI, and 12.9 years (11.5 – ∞) for PSD.

In order to equalize specificities and enable fairer comparisons between the sensitivities of the three global indices for detecting change, we henceforth defined “significant deterioration” to give 95% specificity in the test-retest cohort. This was achieved by setting the p-value cutoffs at $\text{Crit}_{\text{MD}}=0.078$; $\text{Crit}_{\text{VFI}}=0.097$; and $\text{Crit}_{\text{PSD}}=0.084$.

Table 2 shows the median time to detect “significant deterioration” by each global index for eyes in the longitudinal cohort, using the criteria just described above that ensure 95% specificity over five visits in the test-retest cohort. Figure 1 shows the Kaplan-Meier survival curves for each index. VFI detected significant deterioration on average 1.2 years later than MD, roughly equivalent to two visits given our testing schedule, although this difference was not statistically significant ($p=0.088$). PSD detected significant deterioration on average 3.2 years later than MD ($p<0.001$).

Of the 254 eyes that showed “significant deterioration” by MD and also had at least one subsequent visual field available before the end of their series, the change was confirmed on that next test date using the same criterion in 191 eyes (75%). For VFI, this proportion was 186 out of 267 (70%, comparison $p=0.171$ using Fisher’s exact test); for PSD the proportion confirmed was 140 out of 213 (66%, comparison with MD $p=0.032$). Of the eyes that showed “significant deterioration” and had at least two subsequent fields available after that date, changes were confirmed in 171 of 230 eyes (74%) for MD; 177 of 254 eyes (70%) for VFI; and 112 of 200 eyes (56%) using PSD.

It can be seen in Figure 1 that among those eyes that were detected as “significantly deteriorating” earliest in their series, and hence presumably had the most rapid glaucomatous progression, MD detected this deterioration sooner than VFI. The gap between the two indices becomes less pronounced, and eventually disappears, after around 9 years. By this time more than half of the eyes had “significantly deteriorated” according to our criteria. This suggests that MD may be preferable for detecting rapid progression. When restricting the survival analysis to just the first five years of the series for each eye, MD detected “significant deterioration” sooner than VFI, with $p=0.0013$; and sooner than PSD with $p=0.029$. The numbers of eyes that were detected as “significantly deteriorating” by each index within the first five years is shown in Figure 2. The differences between MD and VFI, and between MD and PSD, were both significant with $p<0.001$ (Fisher’s exact test).

Figure 3 shows the Kaplan-Meier survival curves for each index after splitting the cohort in half, according to whether the most recent MD in the series for an eye was above or below the median in the cohort, which was -0.505dB . Not surprisingly, eyes with worse MD were more likely to have shown significant deterioration up to that time.

In eyes with most recent MD $> -0.505\text{dB}$, MD tended to detect deterioration sooner than VFI, but this was not significant; $p=0.700$. However over the first five years of the series, MD did detect deterioration sooner than VFI with $p=0.006$, suggesting again that MD is preferable for detecting rapid progression. There was no significant difference between MD and PSD in this group, either for the entire series ($p=0.230$) or for the first five years ($p=0.740$).

In eyes with most recent MD < -0.505dB, MD tended to detect significant deterioration sooner than VFI, with $p=0.015$ using the entire series, and $p=0.070$ within the first five years. MD also detected deterioration sooner than PSD, with $p<0.001$ using the entire series and $p=0.007$ within the first five years.

Figure 4 shows the median time to detect “significant deterioration” using each index at different specificities. At each of the six tested specificities (99%, 98%, 95%, 90%, 85% and 80%), the median time to detect significant deterioration was sooner for MD than VFI or PSD. For specificities of 98% and 99%, significant deterioration was detected using PSD in fewer than 50% of eyes, hence there is no median time shown. Using the entire series, the difference between MD and VFI was only significant for a specificity of 80% ($p=0.029$; all other specificities $p>0.05$). Using just the first five years, this difference was not significant for specificities of 99% ($p=0.720$) or 98% ($p=0.420$), likely due to the low numbers of eyes reaching these criteria so quickly; however MD did detect deterioration significantly sooner than VFI for specificities of 95% or lower.

Among the 224 eyes that showed “significant deterioration” by PSD, the median time to detect “significant deterioration” by MD was 6.1 years (95% confidence interval 5.9 – 7.0), compared with 6.3 years for VFI (6.0 – 7.0). This difference was statistically significant, with $p=0.037$. The median time using PSD in this group was 5.0 years (4.1 – 5.7), shorter than above because eyes that never deteriorated by PSD are excluded.

104 eyes underwent cataract surgery, and 195 eyes lost 2 lines of visual acuity, during the study period (with 72 eyes meeting both of these criteria); although many had already shown “significant deterioration” by one or more indices prior to these events. Figure 5 shows the Kaplan-Meier survival curves for each index, as in Figure 1, but after censoring series on the date at which either of these criteria were first met. This should minimize any influence of development of cataract on the results. The results were not substantially different. The median time to detect “significant deterioration” in this sub-analysis was 7.5 years for MD (7.0 – 8.6). For VFI, the median was 8.5 years (8.0 – 9.1), with the comparison having $p=0.175$. For PSD, the median was 10.9 years (9.6 – 12.1), and the comparison with MD had $p=0.001$.

Discussion

It is not uncommon for patients with glaucoma to also develop cataract. Both conditions become more common with aging, and as the average life expectancy increases, this situation may become the norm in secondary and tertiary eye care clinics. Therefore, it is reasonable to consider global indices that are less affected by such media opacities.¹⁵ This motivated the development and deployment of the VFI. However, severe cataracts are routinely removed in clinical practice in developed countries, and so the benefits of such indices in this patient population are unclear. Perhaps the most important test is whether VFI improves the ability to detect glaucomatous progression, since changes in ocular media caused by both the development, and removal of, cataracts could occur over the course of glaucoma management/monitoring. Our results show that in our cohort, which resembles a typical population in a tertiary glaucoma clinic and has similar rates of functional change,²⁹

VFI did not detect significant deterioration any sooner than MD. Indeed for an average eye, more often than not, MD actually detected significant deterioration sooner than VFI, as evidenced by the shorter median time in Table 2. While not all of the results presented here attain the conventional criterion for statistical significance, the magnitudes of the differences are considerable. For an average eye in our cohort (with testing occurring every six months), MD detected deterioration more than a year sooner than VFI, and more than three years sooner than PSD.

Comparisons between the indices varied with severity of the disease. In less damaged eyes (most recent MD > -0.505dB), MD and PSD both detected deterioration equally well (with MD detecting it possibly slightly sooner than PSD but with this difference not being significant). However, VFI did not perform well in this subgroup, especially for detecting more rapid deterioration. This could be caused by the ceiling effect on VFI, whereby it continues to report VFI=100% even after some damage may have occurred.^{17, 18} By contrast, in moderately damaged eyes (most recent MD between -0.505dB and -19.5dB) MD and VFI were almost equivalent, with the advantage of MD over VFI not being statistically significant. At this stage, MD and VFI are closely correlated.¹⁷ MD did detect deterioration sooner than PSD in this more damaged group; although it has been suggested that PSD is useful until MD reaches -20dB,¹⁴ it is likely that some individual locations have reached their measurement floor by this point.

Overall then, of the three global indices we tested, we would recommend MD for monitoring glaucomatous progression, since it consistently detected deterioration at least as early as VFI or PSD across the spectrum of disease severities in our cohort, and detected deterioration significantly sooner than VFI or PSD at some stages of the disease.

Even though our study was designed to exclude eyes with severe cataract, it is still possible that a proportion of the eyes will have had mild cataracts that were not removed. If these worsened over time, then this could be theorized to cause a deterioration in MD, but not in VFI or PSD. However, even among eyes that showed deterioration by PSD, and hence can be assumed to have a worsening that is unrelated to cataracts, MD still detected deterioration significantly sooner than VFI. This is particularly impressive considering that requiring deterioration in PSD will bias the sample towards eyes with more localized loss, and so would be expected to favor VFI over MD. A second sub-analysis excluded eyes from the date that they had either lost two lines of visual acuity or undergone cataract surgery, and again MD detected deterioration sooner than VFI or PSD.

Although MD appears the most useful of the three indices examined here, no single index would be expected to detect every case of progression soonest. There were cases in which VFI or PSD detected significant deterioration sooner than MD. This appeared to be caused more by the way the criteria were defined than any characteristics of those eyes. Notably, linear regression could yield a slope with p-value very slightly above $Crit_{MD}$ for one eye and very slightly below $Crit_{MD}$ for another, causing potentially large differences in the date at which significant deterioration was detected even though the series were essentially the same. It is always recommended that clinicians examine the visual fields that make up a

sequence and take all available information into account when making decisions, rather than relying solely on one index.

While our results show that MD is the best global index for detecting functional deterioration out of the three currently available in the HFA perimeter, this certainly does not preclude the development of other indices that may perform better. Glaucoma typically produces paracentral and arcuate scotomas, affecting different subsets of visual field locations in different eyes. Other indices are needed that are more sensitive than MD for detecting deterioration that may be consistent with the development of smaller scotomas, without being limited by the high variability that reduces the specificity of pointwise analyses.³⁰ It has been suggested that increasing the stimulus size used in clinical perimetry may also reduce variability³¹ and extend the dynamic range.^{32, 33} We would expect our conclusions to be preserved using larger stimuli, but this has not yet been tested.

While the majority of the cohort tested in this study undoubtedly had developed extant glaucoma (with the proportion being highly subjective due to its reliance on the criterion chosen³⁴), it contained mostly eyes with early glaucomatous damage, with median MD -0.505dB . As such, we can only speculate about the relative performance of the global indices in severe glaucoma. However, it is known that PSD ceases to increase monotonically with damage later in the disease,¹⁴ reducing its utility for monitoring deterioration. VFI also exhibits a discontinuity once MD reaches around -20dB , with highly variable effects.²⁰ We specifically excluded such visual fields from the present study. It seems likely that in eyes with more severe functional loss, MD would still be preferable to either VFI or PSD.

It would be interesting to see similar analyses in a cohort containing eyes with moderate or severe cataract. It is possible that in such eyes, VFI may have advantages over MD. In particular, if VFI remains more stable than MD in the face of cataract development and/or extraction, then it may be better for monitoring glaucomatous changes over that time period. In the meantime, the clinician should be aware of the potential effect of media opacities on MD, and take this into consideration when making glaucoma treatment decisions in eyes with cataract. It should also be remembered that cataract removal may lower IOP,³⁵ and hence protect against glaucomatous deterioration.

The criteria for “significant deterioration” used in this study were based solely on the p-values from tests of significance of whether the rate of change was significantly different from zero, and were chosen to give equal specificity of 95% in the test-retest cohort. In patients that are younger, have risk factors for rapid progression, or already have significant visual field loss in the other eye, a more sensitive criterion may be desirable at the expense of worse specificity. In other analyses (not shown), the lower specificity criteria that were tested allowed earlier detection of deterioration, as would be expected, but comparisons between the three global indices were unaffected. It should also be noted that “95% specificity” in this case was based on series of five visual fields, and the criteria for deterioration were not altered when longer series were being considered. The criteria used here will likely have less than 95% specificity in longer series, since the resultant reduction in the uncertainty of the slope estimate brought about by using a longer sequence will reduce the associated p-values.

All three of the global indices considered here rely upon age-referenced normative values. So long as these normative values are accurate, any change in the global index is interpreted as being due to pathology. However, it is likely that the true rate of age-related decline varies between individuals. Partly in order to alleviate this issue, some criteria for defining deterioration also impose a minimum rate of change. For example, pointwise linear regression (which has typically been based on non-age-corrected sensitivity values) may require the rate of change to be worse than -1dB/yr as well as having $p < 0.01$.³⁰ Such a criterion could also be useful for global indices, but this was not done in the current study. This is because it would be impossible to fairly derive criteria with equal specificity for the three indices, because one index's criterion could emphasize the magnitude of change while another index's criterion emphasizes the variability, confounding direct comparisons between the indices.

An alternative approach to this question would be to determine which of the global indices best agrees with expert assessment of whether glaucomatous progression has occurred. However, the main problem with using expert assessment of the visual fields as the "gold standard" for deterioration is that such an assessment would include looking at the global indices, and it is not feasible to objectively determine how much effect each of the indices has on the assessment. If experts were asked to assess the visual fields without them having access to the indices in order to remove this potential bias, then this would no longer be replicating the clinical gold standard. Relying solely on stereophotos for the expert agreement would remove this confound, but has been shown to have poor reproducibility both between clinicians (kappa 0.34 – 0.68) and for the same clinician on different days (kappa 0.55 – 0.78).³⁶

In conclusion, we found that Mean Deviation detected significant deterioration of the visual field sooner than the Visual Field Index or Pattern Standard Deviation in patients with early or moderate glaucoma. The magnitude and statistical significance of these differences varied with disease stage, but Mean Deviation performed consistently well in all comparisons. We would recommend using change in Mean Deviation over the other currently-available global indices for monitoring glaucomatous visual field progression.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

- a. Funding/Support:
NIH NEI grant R01-EY020922 (SKG)
NIH NEI grant R01-EY019674 (SD)
Unrestricted research support from The Legacy Good Samaritan Foundation, Portland, OR, USA.
- b. Financial Disclosures:
SKG is a consultant for and has received fees/reimbursements from Haag-Streit Inc, and has been a consultant for Carl Zeiss Meditec Inc.
SD is a consultant for Pfizer Inc.

- c. No other acknowledgments.

All authors attest that they meet the current ICMJE criteria for authorship.

References

1. Hood D, Kardon R. A framework for comparing structural and functional measures of glaucomatous damage. *Prog Retin Eye Res.* 2007; 26(6):688–710. [PubMed: 17889587]
2. Hood DC, Anderson SC, Wall M, Raza AS, Kardon RH. A Test of a Linear Model of Glaucomatous Structure-Function Loss Reveals Sources of Variability in Retinal Nerve Fiber and Visual Field Measurements. *Invest Ophthalmol Vis Sci.* 2009; 50(9):4254–4266. [PubMed: 19443710]
3. Gardiner SK, Demirel S, Reynaud J, Fortune B. Changes in Retinal Nerve Fiber Layer Reflectance Intensity as a Predictor of Functional Progression in Glaucoma. *Invest Ophthalmol Vis Sci.* 2016; 57(3):1221–1227. [PubMed: 26978028]
4. Na JH, Sung KR, Baek S, et al. Detection of glaucoma progression by assessment of segmented macular thickness data obtained using spectral domain optical coherence tomography. *Invest Ophthalmol Vis Sci.* 2012; 53(7):3817–26. [PubMed: 22562510]
5. Leung CK, Ye C, Weinreb RN, Yu M, Lai G, Lam DS. Impact of age-related change of retinal nerve fiber layer and macular thicknesses on evaluation of glaucoma progression. *Ophthalmology.* 2013; 120(12):2485–92. [PubMed: 23993360]
6. Öhnell H, Heijl A, Brenner L, Anderson H, Bengtsson B. Structural and Functional Progression in the Early Manifest Glaucoma Trial. *Ophthalmology.* 2016; 123(6):1173–1180. [PubMed: 26949119]
7. Medeiros FA, Tatham AJ. Structure versus Function in Glaucoma: The Debate That Doesn't Need to Be. *Ophthalmology.* 2016; 123(6):1170–1172. [PubMed: 27210598]
8. Demirel S, de Moraes CGV, Gardiner SK, et al. The rate of visual field change in the ocular hypertension treatment study. *Invest Ophthalmol Vis Sci.* 2012; 53(1):224–227. [PubMed: 22159015]
9. 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(2):1252–1259. [PubMed: 23349433]
10. Bleisch D, Furrer S, Funk J. Rates of glaucomatous visual field change before and after transscleral cyclophotocoagulation: a retrospective case series. *BMC Ophthalmol.* 2015; 15(1):1–7. [PubMed: 25571769]
11. Heijl A, Buchholz P, Norrgren G, Bengtsson B. Rates of visual field progression in clinical glaucoma care. *Acta Ophthalmol.* 2013; 91(5):406–12. [PubMed: 23066646]
12. Song MK, Sung KR, Han S, et al. Progression of primary open angle glaucoma in asymmetrically myopic eyes. *Graefes Arch Clin Exp Ophthalmol.* 2016; 254(7):1331–1337. [PubMed: 27063279]
13. Anderson, D., Patella, V. *Automated Static Perimetry. 2.* St. Louis, MO: Mosby; 1999. p. 147-159.
14. Bengtsson B, Heijl A. A Visual Field Index for Calculation of Glaucoma Rate of Progression. *Am J Ophthalmol.* 2008; 145(2):343–353. [PubMed: 18078852]
15. Chen P, Budenz D. The effect of cataract on the visual field of eyes with chronic open-angle glaucoma. *Am J Ophthalmol.* 1998; 125(3):325–33. [PubMed: 9512149]
16. Carrillo MM, Artes PH, Nicoleta MT, LeBlanc RP, Chauhan BC. Effect of cataract extraction on the visual fields of patients with glaucoma. *Arch Ophthalmol.* 2005; 123(7):929–932. [PubMed: 16009833]
17. Artes PH, O'Leary N, Hutchison D, et al. Properties of the Statpac Visual Field Index. *Invest Ophthalmol Vis Sci.* 2011; 52(7):4030–8. [PubMed: 21467169]
18. Sousa MC, Biteli LG, Dorairaj S, Maslin JS, Leite MT, Prata TS. Suitability of the Visual Field Index according to Glaucoma Severity. *J Curr Glaucoma Pract.* 2015; 9(3):65–8. [PubMed: 26997838]
19. Gardiner SK, Swanson WH, Goren D, Mansberger SL, Demirel S. Assessment of the reliability of standard automated perimetry in regions of glaucomatous damage. *Ophthalmology.* 2014; 121(7):1359–69. [PubMed: 24629617]

20. Rao HL, Senthil S, Choudhari NS, Mandal AK, Garudadri CS. Behavior of Visual Field Index in Advanced Glaucoma. *Invest Ophthalmol Vis Sci*. 2013; 54(1):307–312. [PubMed: 23233259]
21. Aptel F, Aryal-Charles N, Giraud JM, et al. Progression of visual field in patients with primary open-angle glaucoma - ProgF study 1. *Acta Ophthalmol*. 2015; 93(8):e615–20. [PubMed: 26095771]
22. Yousefi S, Balasubramanian M, Goldbaum MH, et al. Unsupervised Gaussian Mixture-Model With Expectation Maximization for Detecting Glaucomatous Progression in Standard Automated Perimetry Visual Fields. *Transl Vis Sci Technol*. 2016; 5(3):2.
23. Ernest PJ, Schouten JS, Beckers HJ, Hendrikse F, Prins MH, Webers CA. Prediction of Glaucomatous Visual Field Progression Using Baseline Clinical Data. *J Glaucoma*. 2016; 25(2): 228–35. [PubMed: 25265001]
24. Bengtsson B, Olsson J, Heijl A, Rootzen H. A new generation of algorithms for computerized threshold perimetry, SITA. *Acta Ophthalmol Scand*. 1997; 75(4):368–75. [PubMed: 9374242]
25. Greenwood M. The errors of sampling of the survivorship tables. *Reports on public health and statistical subjects*. 1926; 33(1):26.
26. Cox DR. Regression Models and Life-Tables. *Journal of the Royal Statistical Society Series B (Methodological)*. 1972; 34(2):187–220.
27. O'Quigley J, Stare J. Proportional hazards models with frailties and random effects. *Stat Med*. 2002; 21(21):3219–33. [PubMed: 12375300]
28. R Development Core Team. R: A language and environment for statistical computing, 1.9.1 ed. Vienna, Austria: R Foundation for Statistical Computing; 2013.
29. Chauhan BC, Malik R, Shuba LM, Rafuse PE, Nicolela MT, Artes PH. Rates of Glaucomatous Visual Field Change in a Large Clinical Population. *Invest Ophthalmol Vis Sci*. 2014; 55(7):4135–4143. [PubMed: 24917147]
30. 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. [PubMed: 11980853]
31. Wall M, Doyle CK, Zamba KD, Artes P, Johnson CA. The Repeatability of Mean Defect with Size III and Size V Standard Automated Perimetry. *Invest Ophthalmol Vis Sci*. 2013; 54(2):1345–1351. [PubMed: 23341012]
32. Wall M, Woodward KR, Doyle CK, Zamba GJ. The Effective Dynamic Ranges of Standard Automated Perimetry Sizes III and V and Motion and Matrix Perimetry. *Arch Ophthalmol*. 2010; 128(5):570–576. [PubMed: 20457977]
33. Gardiner SK, Demirel S, Goren D, Mansberger SL, Swanson WH. The Effect of Stimulus Size on the Reliable Stimulus Range of Perimetry. *Transl Vis Sci Technol*. 2015; 4(2):10.
34. Ng M, Sample PA, Pascual JP, et al. Comparison of visual field severity classification systems for glaucoma. *J Glaucoma*. 2012; 21(8):551–561. [PubMed: 21878817]
35. Mansberger SL, Gordon MO, Jampel H, et al. Reduction in intraocular pressure after cataract extraction: the Ocular Hypertension Treatment Study. *Ophthalmology*. 2012; 119(9):1826–1831. [PubMed: 22608478]
36. Azuara-Blanco A, Katz LJ, Spaeth GL, Vernon SA, Spencer F, Lanzl IM. Clinical agreement among glaucoma experts in the detection of glaucomatous changes of the optic disk using simultaneous stereoscopic photographs. *Am J Ophthalmol*. 2003; 136(5):949–950. [PubMed: 14597063]

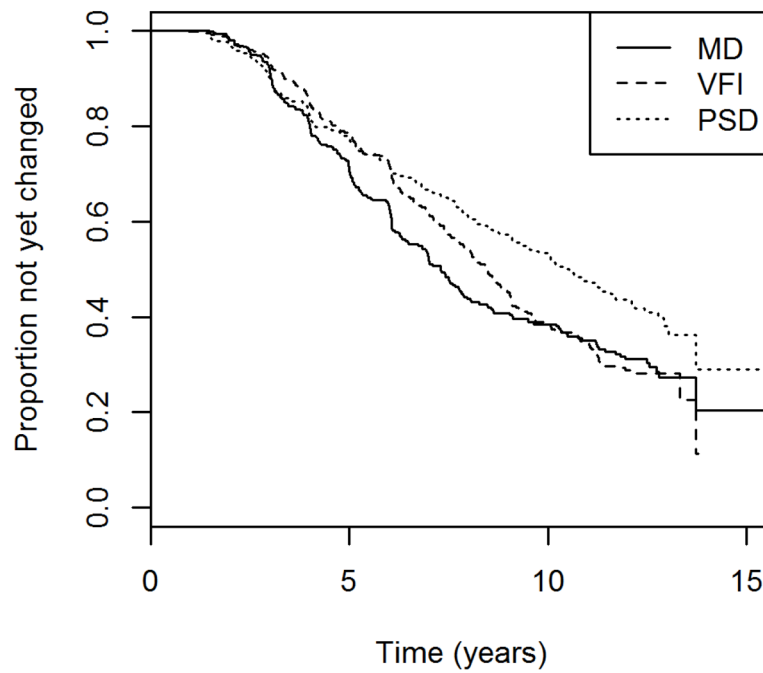


Figure 1.

The time to detect “significant deterioration” using each of three global indices in a longitudinal cohort of participants with early or suspected glaucoma.

Mean Deviation (MD), Visual Field Index (VFI) and Pattern Standard Deviation (PSD).

Criteria were chosen to give 95% specificity in a separate test-retest cohort.

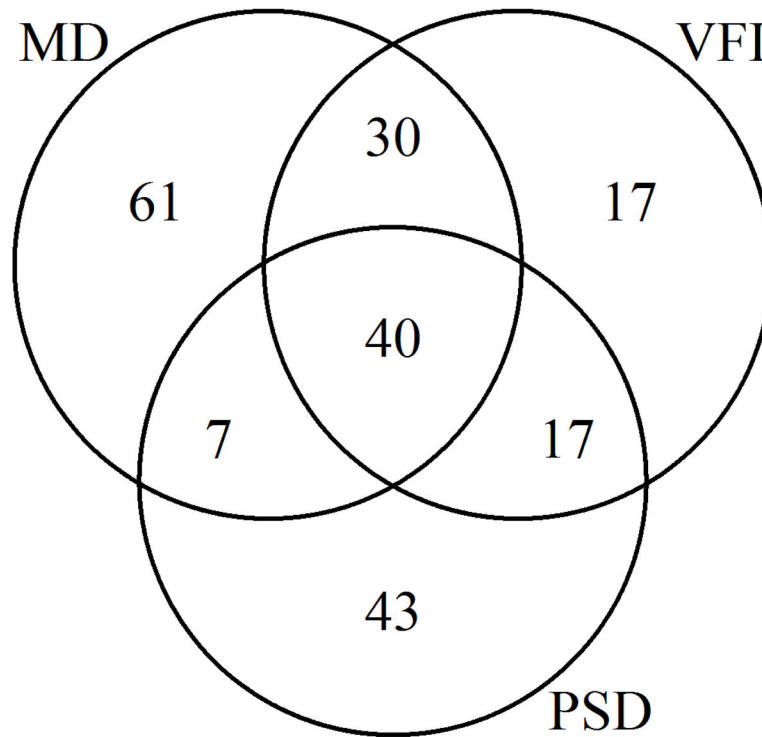


Figure 2. Venn diagram showing the number of eyes in which “significant deterioration” was detected within the first five years of the series, for each of three global indices, in a longitudinal cohort of participants with early or suspected glaucoma. MD = Mean Deviation; VFI = Visual Field Index; PSD = Pattern Standard Deviation.

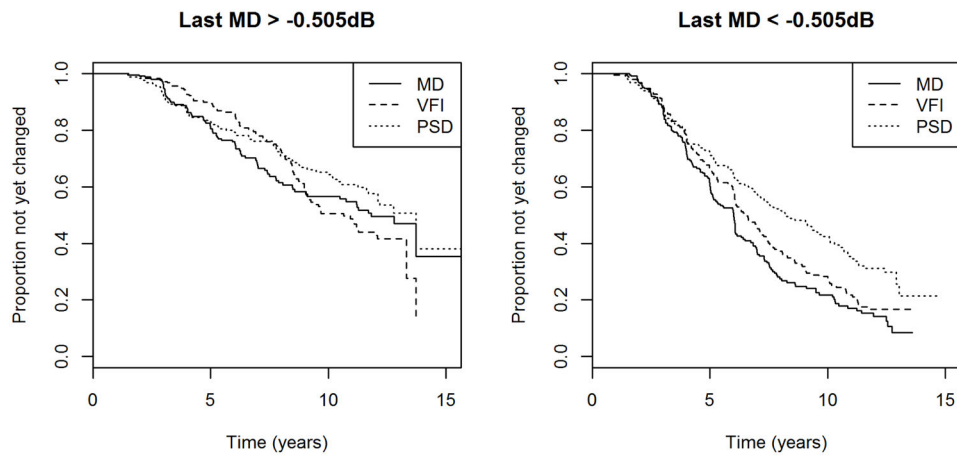


Figure 3.

The time to detect “significant deterioration” using each of three global indices in two halves of a longitudinal cohort of participants with early or suspected glaucoma, split according to whether the most recent Mean Deviation (MD) in the series for an eye was above or below the median value in the cohort.

VFI = Visual Field Index; PSD = Pattern Standard Deviation. Criteria were chosen to give 95% specificity in a separate test-retest cohort.

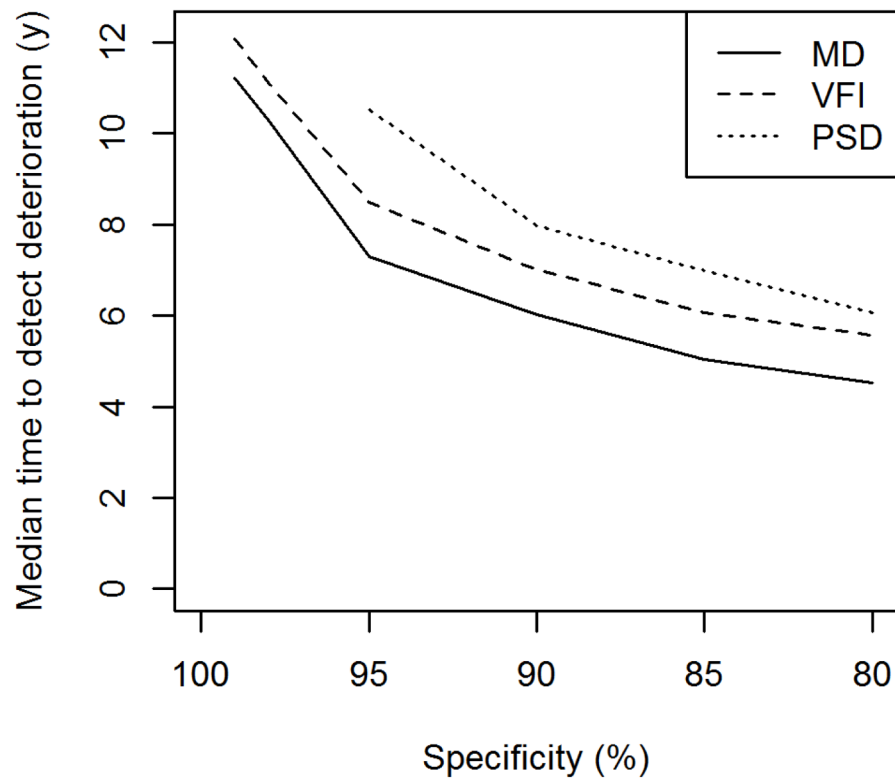


Figure 4. The median time to detect “significant deterioration” using each of three global indices, for criteria with different specificities, in a longitudinal cohort of participants with early or suspected glaucoma. MD = Mean Deviation; VFI = Visual Field Index; PSD = Pattern Standard Deviation.

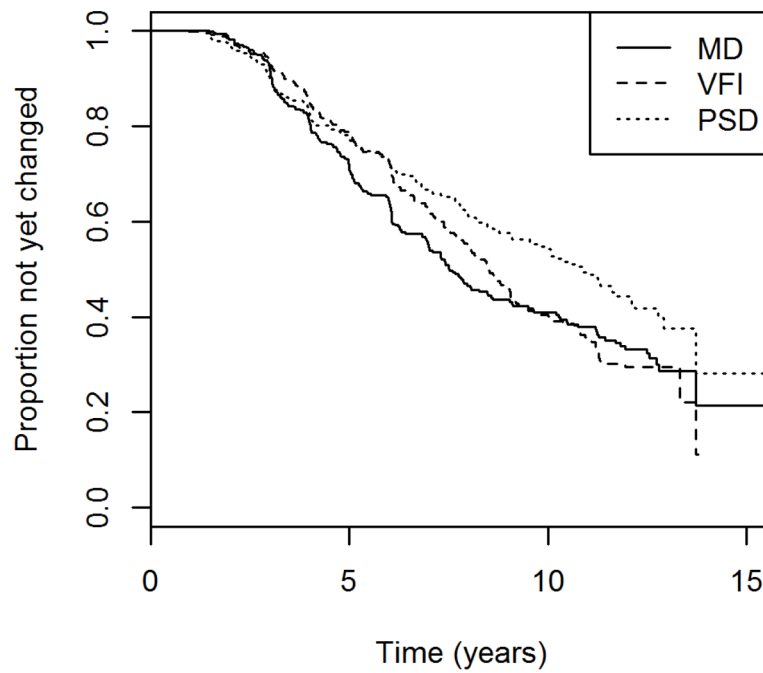


Figure 5.

The time to detect “significant deterioration” using each of three global indices, in a longitudinal cohort of participants with early or suspected glaucoma, censoring series if they either lost two lines of visual acuity or underwent cataract surgery.

Mean Deviation (MD), Visual Field Index (VFI) and Pattern Standard Deviation (PSD).

Criteria were chosen to give 95% specificity in a separate test-retest cohort.

Table 1

Clinical characteristics of the participants in the test-retest and longitudinal cohorts.

		Mean	Interquartile Range	Full Range
Test-retest cohort	Series Length (Days)	39	28 – 48	23 – 69
	Age (years)	66	59 – 69	51 – 93
	Most recent MD (dB)	-2.2	-2.7 – +0.8	-12.6 – +2.5
Longitudinal cohort	Series Length (# Visits)	13	9 – 17	4 – 23
	Age (years)	69	62 – 77	37 – 90
	Most recent MD (dB)	-1.5	-2.8 – +0.8	-19.5 – +2.9
	Rate of MD change (dB/yr)	-0.12	-0.31 – +0.12	-3.05 – +2.03

Age and Mean Deviation (MD) are at the date of their most recent visual field test. The rate of MD change in the longitudinal cohort is over their most recent 6 visual field tests, and so excludes eyes with only 4 or 5 visits. Both cohorts consist of participants with early or suspected glaucoma.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

The time to detect “significant deterioration” using each global index for eyes in the longitudinal cohort.

Global Index	Median Time (years)	95% Confidence Interval
Mean Deviation (MD)	7.3	6.8 – 7.9
Visual Field Index (VFI)	8.5	7.9 – 9.0
Pattern Standard Deviation (PSD)	10.5	9.3 – 11.7

The criteria for “significant deterioration” are based on achieving 95% specificity in the test-retest cohort. The cohort consists of participants with early or suspected glaucoma.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript