tvst

Detection of Functional Deterioration in Glaucoma by Trend Analysis Using Overlapping Clusters of Locations

Stuart K. Gardiner[1] and Steven L. Mansberger[1]

[1] Devers Eye Institute, Legacy Research Institute, Portland, OR, USA

Correspondence: Stuart Gardiner, Devers Eye Institute, Legacy Research Institute, 1225 NE 2nd Ave, Portland, OR 97232, USA. e-mail: sgardiner@deverseye.org

Purpose: Cluster trend analysis detects glaucomatous deterioration within predefined subsets (clusters) of visual field locations. However, it may miss small defects straddling boundaries between the clusters. This study assesses whether simultaneously using a second set of clusters, overlapping the first, could improve progression detection.

Methods: Deterioration in eyes with or at risk of glaucomatous visual field loss was "detected" by mean deviation (MD) on the first visit at which the P value from linear regression over time was below the fifth percentile of its permutation distribution. Similarly, P values were calculated for each of 10 predefined nonoverlapping clusters of locations, or 21 overlapping clusters; deterioration was "detected" when the Nth-smallest P value was below the fifth percentile of its permutation distribution, for different N. Times to detect deterioration were compared using survival models.

Results: Biannual series of ≥ 5 visual fields (mean = 14) were available for 420 eyes of 213 participants. Deterioration of 33% of eyes was detected earliest using $N = 1$ overlapping cluster in 3.3 years (95% confidence interval 2.7–4.6 years); or $N = 2$ nonoverlapping clusters in 3.3 years (2.7–5.0) (comparison $P = 0.654$). There was also no significant difference in the probability that deterioration would be confirmed (92.8% vs. 94.4%, $P = 0.289$). Both overlapping and nonoverlapping clusters detected deterioration significantly sooner than MD (4.5 years, $P \leq 0.001$).

Conclusions: After equalizing specificity, overlapping clusters of locations did not significantly reduce the time to detect deterioration compared with nonoverlapping clusters.

Translational Relevance: Cluster trend analyses detected deterioration sooner than global analyses even when defects straddled cluster borders.

Introduction

Standard automated perimetry (SAP) remains the most commonly used technique to measure longitudinal changes in a glaucoma patient's remaining visual function, both by clinicians and researchers. Although technically no eye is truly "stable," in as much as there is a gradual reduction in function even in healthy eyes, it is important to know how quickly that progression is occurring.[1] Treatment decisions can then be made on the basis of whether functional loss is progressing rapidly enough to be a severe threat to the patient's quality of life.[2] However, the best way to use visual field data to assess progression remains unclear. In particular, there is disagreement on the extent to which global indices should be relied on, such as mean deviation (MD)[3] or the visual field index (VFI)[4,5] when using the Humphrey Field Analyzer (HFA; Carl Zeiss Meditec Inc., Dublin, CA, USA). A recent survey of UK ophthalmologists reported that "Strong agreement consensus was achieved that visual field stability should be assessed by trend analysis or by summary measures of VFI/MD progression."[6] By contrast, the World Glaucoma Association's consensus on Progression of Glaucoma states that "Both local and global metrics are needed for assessment of progression."[7]

We recently reported that pointwise analyses (assessing change at each individual location in the visual field) detected deterioration earlier than
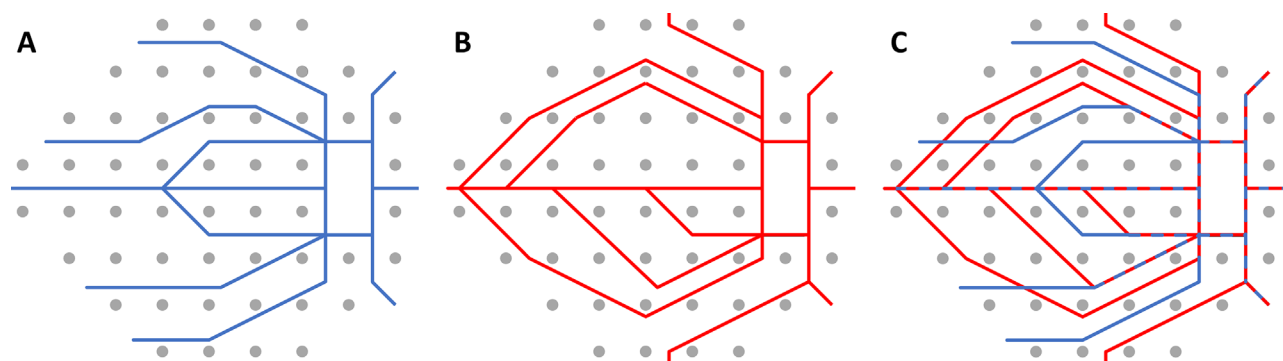
translational vision science & technology

Figure 1. The clusters of visual field locations from the 24-2 test pattern that were used for analyses. (A) The ten sets of locations used for the nonoverlapping clusters analysis. (B) The 11 additional sets of locations used for the overlapping clusters analysis. (C) All 21 sets of locations used for the overlapping clusters analysis.

global analyses (relying on a summary measure such as MD). However, results from global analyses were more likely to be confirmed when using information from subsequent visits, whereas pointwise analyses tended to "overcall" progression.[8] As a reasonable middle ground between those two conflicting priorities, we suggested looking at the rate of change within 10 predefined nonoverlapping subsets ("clusters") of visual field locations, as implemented in the cluster trend analysis that forms part of the EyeSuite software developed for the Octopus perimeter (Haag-Streit Inc., Bern, Switzerland).[9]

However, on the retina, localized defects are not constrained to these predefined clusters. A defect covering several test locations that straddles the border between two of the clusters may not be detected by this technique, especially if it affects only half or fewer of the locations within any single cluster. In the worst case scenario, a deep localized defect could cover as many as six locations, i.e., an area of approximately 18° × 12°, without covering more than 50% of the locations in any single cluster, potentially making the existing cluster trend analysis suboptimal for its detection and monitoring.

There is therefore compelling logic suggesting that the existing cluster trend analysis may fail to detect some visual field defects, especially smaller defects that occur in early glaucoma. In this study, we hypothesized that such defects would be better detected by supplementing the existing 10 predefined clusters of test locations (Fig. 1A) with a second overlapping set of clusters (Fig. 1B), chosen based on the known topographic structure-function relationship. We compare the ability of three techniques (global analyses, nonoverlapping predefined clusters, and the new overlapping clusters) to detect significant deterioration of the visual field in a large cohort with a long

period of longitudinal follow-up, for exactly the same specificity. This technique could improve the ability of both clinicians and researchers (especially in clinical trial scenarios) to rapidly and reliably detect visual field progression.

Methods

Participants

This was a retrospective analysis from an ongoing cohort study, including participants from the Portland Progression Project (P3) located at Legacy Devers Eye Institute. Inclusion criteria were a diagnosis of primary open-angle glaucoma or likelihood of developing glaucomatous damage, as determined subjectively by each participant's physician, to reflect current clinical practice. Exclusion criteria were an inability to perform reliable visual field testing, best-corrected visual acuity at baseline 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 are tested every six months with a variety of structural and functional tests. Many of the participants were recruited from cohorts involved in previous studies that used annual visual field testing, but only visits occurring since the switch to biannual testing were used, to maximize consistency of the

test-retest intervals. Standard automated perimetry was performed using a Humphrey Field Analyzer HFAII perimeter, with the 24-2 test pattern, a size III white-on-white stimulus, and the SITA Standard algorithm.[10] Only tests with ≤15% false positives and ≤33% fixation losses were used for the primary analysis. Because fixation losses are frequently the result of inaccurate mapping of the blind spot at the start of the test rather than true unreliability due to unstable fixation, a secondary analysis was performed using tests with ≤15% false positives but without any criterion for fixation losses. For this study, only eyes with series of at least five reliable tests by these criteria were included in the analysis.

Analysis—Definition of Clusters

Three types of analysis of the visual field testing data were considered for this study: Global (based on MD), nonoverlapping clusters, and overlapping clusters. For the nonoverlapping clusters analysis, the visual field was split into 10 clusters, using cluster definitions from the EyeSuite software (Haag-Streit Inc, Bern, Switzerland), as shown by the blue lines in Figure 1A; and the total deviation on the native decibel scale was averaged within each cluster.[8] Total deviation values for each location were used instead of raw sensitivities to account for the effect of normal aging; therefore in the absence of disease progression each cluster would be expected to show zero change over time. We have previously shown that "censoring" sensitivities below 15 dB and setting them equal to 15 dB improves reliability and, hence, the ability to detect change; therefore the total deviation values for any such locations were set to equal the total deviation value for a sensitivity of 15dB.[11,12] For the overlapping clusters analysis, these 10 clusters were supplemented by an additional eleven clusters, as shown by the red lines in Figure 1B. These additional clusters were selected based on the average angle at which nerve fiber bundles from each location enter the optic nerve head, according to the map of Garway-Heath et al.[13]; and such that each cluster contained at least two locations. The clusters include locations whose average angle of entry is 20° to 60°; 60° to 80°; 80° to 100°; 100° to 120°; 120° to 180°; 180° to 240°; 240° to 260°; 260° to 280°; 280° to 290°; 290° to 340°; and 340° to 20°. Therefore there are a total of 21 overlapping clusters, shown in Figure 1C. Note that the two clusters temporal of the blind spot are counted twice in the overlapping clusters analysis to ensure that any defects in this underrepresented region of the visual field will still be detected.

Analysis—Detecting Deterioration by Global Trend Analysis

Looking for deterioration in any one of multiple clusters would be expected to increase sensitivity, but this would inevitably come at the cost of reduced specificity. To avoid this confound, we equalize specificity by using a permutation technique to determine whether an eye is "deteriorating,"[14] rather than using the same P value cutoff for both overlapping and nonoverlapping cluster analyses.[8] The method for detecting deterioration was based on the published permutation analyses of pointwise linear regression approach[14] and implemented in R statistical programming language (version 3.5.0).[15] For a series of V visual fields, where $V \geq 5$, the "observed" significance of the rate of change was defined as the P value from an ordinary least squares regression over time. Next, a permutation distribution for this P value was derived. The values of MD for visits 1-V were reordered, and these reordered MD values were regressed against the original test dates. For $V = 5$, this was done for all 120 possible reorderings of the five visits; for $V > 5$, 475 randomly chosen reorderings were used to avoid excessive computation time. *Deterioration* in MD was "detected" at the first visit V for which the observed significance was below the fifth percentile of the permutation distribution. Therefore this criterion has a specificity of 95% and, based on a binomial distribution 475 reorderings, gives a confidence interval for this specificity of ±1%.

Note that this procedure gives very similar results to just "detecting" deterioration on the first visit at which the observed one-sided P value for the rate of change is less than 5%. However, it makes fewer distributional assumptions, particularly concerning homoscedasticity. More importantly, it can more easily be extended to cluster analyses, as detailed in the next section, to ensure consistency between the analysis types.[14]

Analysis—Detecting Deterioration by Cluster Trend Analysis

Within each predefined cluster of visual field locations (see Fig. 1), the mean total deviation value (i.e., the difference from age-matched normal) was regressed against visit date for visits 1-V, and the associated P value for the rate of change was recorded. This gives a set of 10 P values when using nonoverlapping clusters and 21 P values when using overlapping clusters. These were sorted from smallest to largest. As before, the visits in the series were then reordered, either using all possible reorderings for $V = 5$ or 475 random reorderings for $V > 5$. For a given number of clusters

N, deterioration was "detected" on the first visit V for which the Nth-smallest observed P value was below the fifth percentile of the Nth-smallest P values from all reorderings. Note that the Nth-smallest P value will not always be from the same cluster in each reordering; instead the P values are sorted from smallest to largest for each individual reordering before deriving the fifth percentile for each N. This analysis ensures that the specificity is exactly equal to 95% for each technique, allowing a fair comparison between the times to detect deterioration.

Analysis—Confirmation of Deterioration

"Confirmed deterioration" was detected on the first visit V at which the Nth-smallest observed P value was below the fifth percentile of the Nth-smallest P values from all reorderings, for both visits the series 1 through V and the series 1 through $(V+1)$. The date of detection is defined as being visit V, not visit $V+1$. It was not necessary for the Nth-smallest P value to come from the same sector for both time points. The probability that "confirmed deterioration" was detected on the same date that "deterioration" was detected can then be taken as a metric of the robustness of a particular analysis.[16,17]

Analysis – Comparison of Criteria

Kaplan-Meier survival curves were plotted to show how soon each criterion detected "deterioration" or "confirmed deterioration." The lower tertile survival times were found (the first date by which deterioration had been detected in $\geq$33% of eyes), together with 95% confidence intervals using standard errors based on Greenwood's formula.[18] Differences between the survival curves were assessed using a stratified Cox proportional hazards model,[19] with strata identifying fellow eyes of the same individual. Including strata is equivalent to using generalized estimating equations (GEE) in a linear model to adjust for intereye correlations.[20] Subanalyses were performed within the subset

of eyes that were abnormal at the start of their series, defined as either abnormal pattern standard deviation ($P < 5$%) or a glaucoma hemifield test result of either "abnormal" or "borderline"; and within the subset of eyes that did not meet those criteria and so would be considered normal at the start of their series. GEE logistic regressions were used to determine whether the probability that deterioration would be subsequently confirmed differed between criteria.

Results

Series of at least five reliable visual fields were available for 420 eyes of 213 participants. One hundred forty-six of those eyes had an abnormal result on the first visual field in their series, as defined above. Table 1 summarizes other characteristics of the cohort.

Table 2 shows the time taken to detect deterioration, and confirmed deterioration (i.e., the series was still "deteriorating" after the inclusion of the next test date in the analysis), in at least 33% of the cohort (i.e., the lower tertile of survival times); based on mean deviation, and based on up to three clusters. More than three clusters took longer to detect deterioration (results not shown). Because testing was conducted biannually (or as close as possible), and only series of length $\geq$5 visits were analyzed, the first possible date at which deterioration could be detected was approximately two years. Table 3 shows P values comparing the time to detect deterioration between each pair of these criteria based on stratified Cox proportional hazards models; and in italics the equivalent P values comparing the time to detect confirmed deterioration. Cluster trend analysis, using either one or two nonoverlapping clusters, or using one to three overlapping clusters, detected deterioration significantly sooner than MD. The criteria that detected deterioration soonest were using one nonoverlapping cluster, or one or two overlapping clusters; these three criteria were not significantly different from each other for detecting either deterioration or

Table 1. Characteristics of the Dataset Used

	Mean	Standard Deviation	Range	Interquartile Range
Series length (number of visits)	13.9	4.7	5 to 23	10 to 18
Series length (years)	7.8	2.2	1.9 to 10.4	6.0 to 9.6
Age (years)	72.1	10.8	39.9 to 94.0	65.5 to 79.4
Initial mean deviation (dB)	−0.81	2.8	−17.7 to +3.0	−1.4 to +0.9
Final mean deviation (dB)	−1.72	4.0	−26.3 to +3.0	−2.6 to +0.8
Rate of change of mean deviation (dB/year)	−0.14	0.3	−2.5 to +0.8	−0.2 to +0.0

Table 2.　Time (in Years) to Detect Deterioration in One Third of Eyes in the Cohort, By Various Different Criteria

Criteria for Deterioration	Years to Detect Deterioration in ≥33% of Eyes	95% Confidence Interval	Years to Detect Confirmed Deterioration in ≥33% of Eyes	95% Confidence Interval	Probability of Confirmation
Mean Deviation	4.51	3.28–∞	7.22	3.41–∞	92.2%
1 Non-overlapping cluster	3.28	2.73–5.06	3.41	2.75–5.48	94.4%
2 Non-overlapping clusters	3.59	2.96–9.10	4.01	3.03–∞	94.0%
3 Non-overlapping clusters	4.20	3.23–∞	6.98	3.36–∞	87.7%
1 Overlapping cluster	3.28	2.75–4.56	3.70	2.92–9.10	87.8%
2 Overlapping clusters	3.28	2.74–4.60	3.41	2.76–6.98	92.8%
3 Overlapping clusters	3.41	2.76–4.64	3.73	3.03–∞	91.5%

Deterioration was "detected" on the first visit date at which the significance of the rate of change was in the lower fifth percentile of its permutation distribution, either based on mean deviation, or based on the cluster with the Nth-smallest P value. Confirmed deterioration was "detected" if this criterion was still met after the addition of the next visit to the series. The final column shows the probability that if deterioration was detected, then it would be confirmed on the next visit; i.e., the probability that the time to detect deterioration equals the time to detect confirmed deterioration.

Table 3.　Significance of Pairwise Comparisons Between the Times to Detect Deterioration By Different Criteria

	Mean Deviation	1 Non-Overlapping Cluster	2 Non-Overlapping Clusters	3 Non-Overlapping Clusters	1 Overlapping Cluster	2 Overlapping Clusters
1 Nonoverlapping cluster	0.001					
	<0.001					
2 Nonoverlapping clusters	0.049	0.062				
	0.038	*0.007*				
3 Nonoverlapping clusters	0.199	0.025	0.324			
	0.937	*<0.001*	*0.017*			
1 Overlapping cluster	0.001	0.718	0.053	0.018		
	0.010	*0.055*	*0.338*	*0.006*		
2 Overlapping clusters	<0.001	0.654	0.038	0.008	0.910	
	<0.001	*0.371*	*0.045*	*<0.001*	*0.305*	
3 Overlapping clusters	0.002	0.683	0.090	0.020	0.534	0.399
	0.006	*0.067*	*0.363*	*0.002*	*0.792*	*0.191*

Values in *italics* represent between the times to detect confirmed deterioration.

confirmed deterioration. There were also no significant differences among these three criteria in the probability that deterioration would be confirmed ($P = 0.910$ for one nonoverlapping vs. one overlapping, by logistic GEE regression; $P = 0.229$ for one nonoverlapping vs. two overlapping; and $P = 0.197$ for one nonoverlapping vs. two nonoverlapping). Figure 2 shows Kaplan-Meier survival curves comparing the time to detect deterioration by these three criteria, and by MD.

Figure 3 shows the time to detect deterioration by these criteria, with the cohort split into 146 eyes for

which the visual field was abnormal (either by PSD, or GHT, or both) at the start of the series, versus 274 eyes for which the initial visual field was normal. Among eyes with initially normal fields, the criterion based on two overlapping clusters detected confirmed deterioration in one third of eyes slightly sooner than the criterion based on one overlapping cluster ($P = 0.043$), but none of the other comparisons between the three cluster criteria were statistically significant (all $P > 0.05$).

The secondary analysis included visual fields regardless of the reported proportion of fixation losses, so
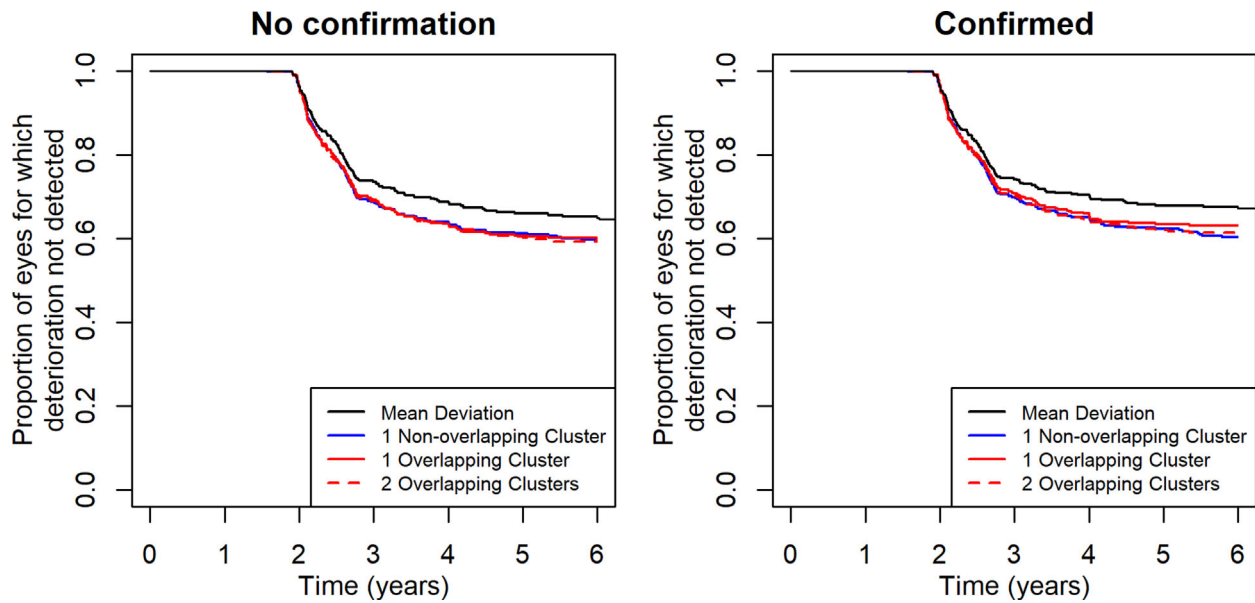
Figure 2. Time to detect deterioration (*left*), or confirmed deterioration (*right*), for the three best-performing cluster trend analysis criteria, together with mean deviation for comparison. Deterioration was "detected" on the first visit date at which the significance of the rate of change was in the lower fifth percentile of its permutation distribution, either based on mean deviation, or based on the cluster with the *N*th-smallest *P* value. Confirmed deterioration was "detected" if this criterion was still met after the addition of the next visit to the series.

long as they had ≤15% false positives. This increased the total number of available visual fields from 8628 to 10,220. It also increased the number of eyes for which at least five reliable fields were available (and, hence, could be included in the analysis for this study) from 439 to 485. Because more frequent visual fields were available for many eyes, the time to detect deterioration in one third of eyes (as in Table 2) was reduced to 2.96 years for MD (95% confidence interval 2.34–∞); 2.37 years for one nonoverlapping cluster (2.21–3.00); 2.43 years for one overlapping cluster (2.22–3.07); and 2.32 years for two overlapping clusters (2.20–2.75). All three of these cluster-based criteria detected deterioration significantly sooner than MD (all $P \leq 0.001$), but as in the primary analysis the overlapping cluster criteria were not significantly sooner than the nonoverlapping cluster criterion ($P = 0.394$ for one overlapping cluster vs. one nonoverlapping cluster; $P = 0.095$ for two overlapping clusters vs. one nonoverlapping cluster).

Figure 4 shows how frequently each of the predefined sectors was the one whose rate of change had the smallest *P* value based on a series of seven visits and, hence, was the sector that would trigger "detection" by visit seven when $N = 1$, among the 107 eyes for which deterioration would be detected by this visit using nonoverlapping clusters (Fig. 4A) and among the 116 eyes for which deterioration would be detected using overlapping clusters (Figs. 4B and 4C).

Discussion

This study extends our previous results showing that cluster trend analysis, a technique which assesses the rate of change within each of a set of 10 predefined clusters of test locations, seems to provide a useful compromise between the relative advantages and disadvantages of global versus pointwise methods for detecting visual field change. We tested the hypothesis that visual defects that straddle a cluster may decrease sensitivity or increase the time to detect deterioration with cluster analysis, using permutation analysis to equalize specificity and found that using a second, overlapping set of clusters did not significantly reduce the time to detect deterioration compared with nonoverlapping clusters. We also confirmed the previous finding that nonoverlapping clusters detected deterioration earlier than MD, with the additional advantage of preserving spatial information lost when using global indexes. Therefore we conclude that cluster analysis has utility even when visual field defects straddle a predetermined cluster border and can aid diagnostic decision making in both research and clinical care.

Nonoverlapping clusters could in theory miss defects that straddle the borders between clusters. In this study, we aimed to reduce this problem by using a second, overlapping set of clusters, as shown
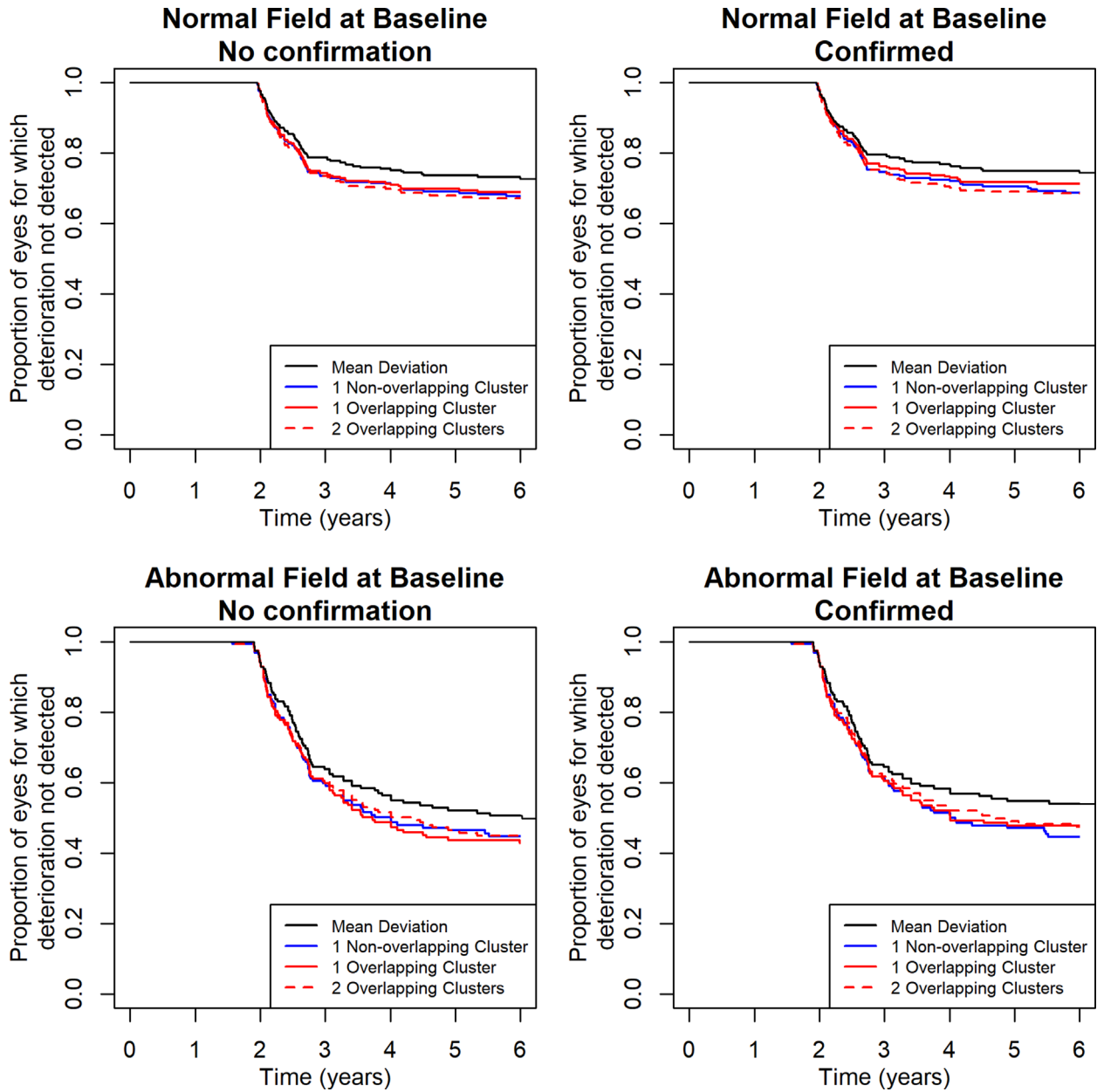
Figure 3. Time to detect deterioration (*left*), or confirmed deterioration (*right*), for the three best-performing cluster trend analysis criteria, together with mean deviation for comparison; for the 287 eyes whose visual field was within normal limits for both pattern standard deviation and glaucoma hemifield Test on their initial visit (*top*), and for the 154 eyes whose visual field was outside normal limits on their initial visit (*bottom*). Deterioration was "detected" on the first visit date at which the significance of the rate of change was in the lower fifth percentile of its permutation distribution, either based on mean deviation, or based on the cluster with the *N*th-smallest *P* value. Confirmed deterioration was "detected" if this criterion was still met after the addition of the next visit to the series.

in Figure 1. This should improve sensitivity for detecting deterioration. However, simply changing the criteria for "deterioration" from "at least one of ten nonoverlapping clusters has $P < 5\%$" to "at least one of 21 overlapping clusters has $P < 5\%$" would naturally increase sensitivity while decreasing specificity for detecting deterioration, especially when the ten non-overlapping clusters form a subset of the

21 overlapping clusters. To avoid this problem, we used permutation analysis to equalize specificity between criteria and between series lengths. After equalizing specificity, we found that using overlapping clusters of locations did not significantly reduce the time to detect deterioration compared with nonoverlapping clusters. We thus conclude that the identified caveat with the previous results, concerning defects straddling cluster
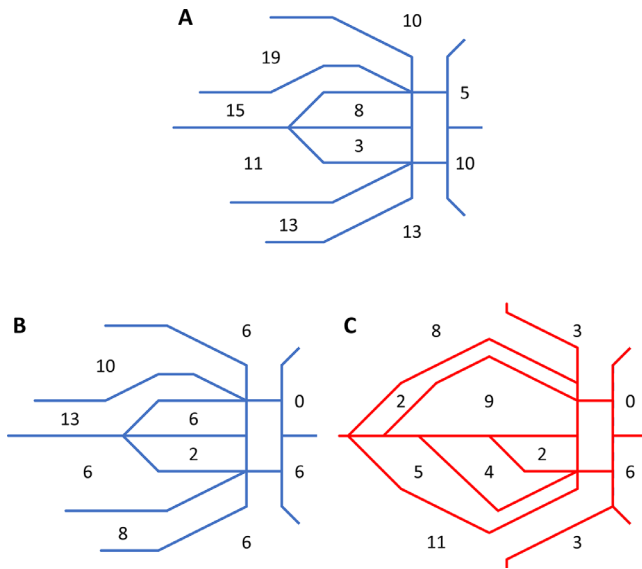
Figure 4. How frequently (number of eyes) each cluster of locations was the cluster with the most significant rate of change, among eyes showing significant deterioration over series of seven visual fields. (A) Using the 10 nonoverlapping predefined clusters. (B and C) Using the 21 overlapping predefined clusters (split into two sets for ease of viewing).

borders, does not in fact significantly reduce the utility of the cluster trend analysis technique.

Functional loss due to glaucoma most commonly takes the form, at least initially, of localized scotoma. Generalized loss is also associated with glaucoma,[21–25] but it is less disease specific because it can also be caused by other factors such as cataract.[4,26,27] Therefore it is natural to suppose that diagnostic measures of glaucomatous progression would be better based on localized rather than global analyses. However, global analyses also have advantages. Because they are based on information from multiple locations, global indexes are considerably less variable than pointwise sensitivities, and so any deterioration that is detected is more likely to be confirmed on subsequent retesting,[8] reducing the need for as many confirmatory follow-up visits.[28] The likelihood of confirmation is also enhanced by using trend analyses rather than event analyses, for which as many as half of clinical trial endpoints may not be confirmed upon retest.[16] Cluster trend analysis[9] appears to represent a good compromise between these two priorities, providing rapid detection of subsequently confirmed progression without excessive false-positive determinations.[8]

Our results should not be taken as implying that cluster trend analysis using ten predefined nonoverlapping clusters will not miss defects that straddle the cluster boundaries. Instead, the results imply that any improvement in earlier detection using overlapping

clusters is short enough that it does not outweigh the accompanying reduction in specificity. The nonoverlapping cluster analysis as currently implemented within the EyeSuite software[9] is simpler to visually represent and, hence, quicker to understand and so may be preferred for clinical use.

By contrast, both versions of cluster trend analysis detected deterioration significantly sooner than MD. The time until deterioration had been detected in one third of eyes was over a year later for MD, equivalent to two more visits. This is in agreement with the conclusions of our earlier study (although the average time to detect deterioration was longer in that previous study, because it included data from eyes that had undergone annual testing, whereas this study used biannual testing).[8] The poorer performance of MD could be related to the lower number of test locations in central versus peripheral regions, which is only partly compensated for by the increased weighting given to those central locations; but it seems more likely that it is because glaucomatous progression consists of not just generalized but also localized deterioration which a global average is poorly suited to detecting. Cluster trend analysis does appear to provide a useful tool for clinical care; indeed it could be a useful addition to the analysis software available for other commercial perimeters.

It is a truism that all eyes are technically progressing, because the visual system deteriorates with age.[29] Participants in this study had either clinically diagnosed or suspected glaucoma, and so it would be expected that the average rate of deterioration would be faster than in a healthy population. Despite this, deterioration was detected in fewer than half of the cohort even after an average of eight years. This partly reflects the relatively early functional loss, if any, of many eyes in the study (not least because both eyes were included even if only one eye would be considered glaucomatous clinically). It also reflects the fact that the participants were not only being clinically managed, but were also interested and dedicated enough to participate in a long-term study, and as such could be expected to have higher medication adherence than a more general population of glaucoma patients.[30,31] However it also reflects the fact that intertest variability makes it difficult to detect disease progression by functional testing alone.[32–36]

Analysis tools can improve our ability to make best use of the data obtained from perimetry, but ultimately it remains important to develop methods to reduce the variability in the first place; for example by altering stimulus characteristics or testing algorithms.[37–39] Among eyes for which deterioration was detected by the end of their series using one non-overlapping

cluster, the average time until detection was 2.8 years. Given that a minimum of five visual fields was required, which would cover a period of approximately two years, analysis tools may be approaching the point at which any further improvements in the time to detect progression will become less clinically significant. There is still room for improvement in data analysis, but there may be greater benefits to be obtained by developing new testing methods that can reduce test variability to the point where a reliable measurement of the rate of change can be obtained using only four or fewer visits.

Participants in this study were tested twice annually, aiming to get as close as possible to six-monthly intertest intervals. Although inevitably some visits were missed, or excluded due to unreliability, the mean time between visits was 0.67 years, or 0.54 years for the secondary analysis that did not exclude fields based on the proportion of fixation losses. Although this consistency of the testing schedule across disease stages improves the robustness of the results, it would be more common clinically for patients with more severe glaucoma or experiencing more rapid disease progression to be tested more frequently, whereas those with early and stable glaucoma may only perform visual field testing annually (or even less).[40] With a longer intertest interval, the amount of deterioration that has occurred relative to the test-retest variability would be higher, and so estimates of the rate of change should become more reliable.[36] This would be expected to have greater impact on cluster analyses than on global analyses such as mean deviation, where the large number of pointwise deviation values being included in the weighted average means that measurement noise is already greatly reduced. We would therefore conjecture that the benefits of cluster analyses over global indexes would be greater with less frequent testing. However, this has not yet been tested. A further caveat is that the participants in this study are highly experienced with automated perimetry, because of the time they have been in the study; averaging information from a large number of locations may be more beneficial in less-experienced patients who might be expected to give more variable test results.[41,42]

The clusters used aim to reflect the shapes of defects related to localized structural loss, using the known structure-function relation.[13] This topographic map is known to vary between individuals.[43] This variability should not greatly affect the clusters, which are based on the relative proximity of axons as they enter the optic nerve head rather than the exact position at which this happens. The possible exception to this would be along the temporal raphe, because the most temporal visual field locations could map to either hemifield of the disc.[44] Although we cannot discount the possibility that customized structure-function mapping could improve the utility of cluster analyses, it is notable in Figure 4 that the clusters bordering (and potentially encompassing) the raphe were no less likely to detect deterioration than more superior or inferior clusters.

As seen in Figure 4, deterioration was most commonly detected quickest in the mid-peripheral region of either hemifield, with a higher likelihood of appearing in the superior hemifield, as would be expected.[45,46] However, it should be noted that there were eyes for which the strongest signal of deterioration came from central locations, particularly in the superior hemifield. This is in agreement with the recent work of Hood and colleagues,[47,48] showing that central visual field loss commonly occurs in glaucoma, even though it is often missed due to the low number of field locations within this region in the 24-2 test grid relative to the density of retinal ganglion cells.

In conclusion, we found that cluster trend analysis detected progression of the visual field significantly sooner than Mean Deviation. However, there was no significant additional benefit from supplementing the ten predefined non-overlapping clusters with a second set of overlapping clusters, since the increase in sensitivity was counteracted by a decrease in specificity. Cluster-based endpoints may be preferable for determining glaucomatous functional progression, for both clinical trials and clinical care.

Acknowledgments

References

1. Caprioli J. The importance of rates in glaucoma. *Am J Ophthalmol*. 2008;145:191–192.
2. Saunders LJ, Russell RA, Kirwan JF, McNaught AI, Crabb DP. Examining Visual Field Loss in Patients in Glaucoma Clinics During Their Predicted Remaining Lifetime. *Investigative Ophthalmology & Visual Science*. 2014;55:102–109.
3. Heijl A, Lindgren G, Lindgren A, et al. *Extended empirical statistical package for evaluation of single and multiple fields in glaucoma: Statpac 2*. In: Mills

R, Heijl A (Eds.), *Perimetry Update 1990/1991*. Amsterdam: Kugler & Ghedini;1991:303–315.

4. Bengtsson B, Heijl A. A visual field index for calculation of glaucoma rate of progression. *Am J Ophthalmol*. 2008;145:343–353.

5. Artes PH, O'Leary N, Hutchison DM, et al. Properties of the statpac visual field index. *Invest Ophthalmol Vis Sci*. 2011;52:4030–4038.

6. Lakhani BK, Giannouladis K, Leighton P, King AJ. Seeking a practical definition of stable glaucoma: a Delphi consensus survey of UK glaucoma consultants. *Eye (London, England)*. 2020;34:335–343.

7. Weinreb RN, Garway-Heath DF, Leung C, Crowston JG, Medeiros F. Progression of glaucoma. Kugler Publications;2011.

8. Gardiner SK, Mansberger SL, Demirel S. Detection of Functional Change Using Cluster Trend Analysis in Glaucoma. *Invest Ophthalmol Vis Sci*. 2017;58:Bio180–bio190.

9. Racette L, Fischer M, Bebie H, Hollo G, Johnson CA, Matsumoto C. Interpretation of visual field progression. In: Haag-Streit AG (Ed.), *Visual Field Digest: A guide to perimetry and the Octopus perimeter* (8th edition). Koniz, Switzerland;2019:165–192.

10. Bengtsson B, Olsson J, Heijl A, Rootzen H. A new generation of algorithms for computerized threshold perimetry, SITA. *Acta Ophthalmol Scand*. 1997;75:368–375.

11. Gardiner SK, Swanson WH, Demirel S. The Effect of Limiting the Range of Perimetric Sensitivities on Pointwise Assessment of Visual Field Progression in Glaucoma. *Invest Ophthalmol Vis Sci*. 2016;57:288–294.

12. 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;121:1359–1369.

13. Garway-Heath DF, Poinoosawmy D, Fitzke FW, Hitchings RA. Mapping the visual field to the optic disc in normal tension glaucoma eyes. *Ophthalmology*. 2000;107:1809–1815.

14. O'Leary N, Chauhan BC, Artes PH. Visual field progression in glaucoma: estimating the overall significance of deterioration with permutation analyses of pointwise linear regression (PoPLR). *Invest Ophthalmol Vis Sci*. 2012;53:6776–6784.

15. R Development Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing;2013.

16. Keltner JL, Johnson CA, Quigg JM, Cello KE, Kass MA, Gordon MO. Confirmation of visual field abnormalities in the Ocular Hypertension Treatment Study. Ocular Hypertension Treatment Study Group. *Arch Ophthalmol*. 2000;118:1187–1194.

17. Keltner JL, Johnson CA, Levine RA, et al. Normal visual field test results following glaucomatous visual field end points in the Ocular Hypertension Treatment Study. *Archives of Ophthalmology*. 2005;123:1201–1206.

18. Greenwood M. The errors of sampling of the survivorship tables. *Reports on Public Health and Statistical Subjects*. 1926;33:26.

19. Cox DR. Regression Models and Life-Tables. *Journal of the Royal Statistical Society Series B (Methodological)*. 1972;34:187–220.

20. O'Quigley J, Stare J. Proportional hazards models with frailties and random effects. *Statistics in Medicine*. 2002;21:3219–3233.

21. Airaksinen PJ, Drance SM, Douglas GR, Mawson DK, Nieminen H. Diffuse and localized nerve fiber loss in glaucoma. *Am J Ophthalmol*. 1984;98:566–571.

22. Drance SM. Diffuse visual field loss in open-angle glaucoma. *Ophthalmology*. 1991;98:1533–1538.

23. Chauhan BC, LeBlanc RP, Shaw AM, Chan AB, McCormick TA. Repeatable diffuse visual field loss in open-angle glaucoma. *Ophthalmology*. 1997;104:532–538.

24. Henson DB, Artes PH, Chauhan BC. Diffuse loss of sensitivity in early glaucoma. *Invest Ophthalmol Vis Sci*. 1999;40:3147–3151.

25. Grewal DS, Sehi M, Greenfield DS. Diffuse glaucomatous structural and functional damage in the hemifield without significant pattern loss. *Arch Ophthalmol*. 2009;127:1442–1448.

26. Smith SD, Katz J, Quigley HA. Effect of cataract extraction on the results of automated perimetry in glaucoma. *Arch Ophthalmol*. 1997;115:1515–1519.

27. Koucheki B, Nouri-Mahdavi K, Patel G, Gaasterland D, Caprioli J. Visual field changes after cataract extraction: the AGIS experience. *Am J Ophthalmol*. 2004;138:1022–1028.

28. Gardiner SK, Crabb DP. Examination of different pointwise linear regression methods for determining visual field progression. *Invest Ophthalmol Vis Sci*. 2002;43:1400–1407.

29. Johnson CA, Adams AJ, Lewis RA. Evidence for a neural basis of age-related visual field loss in normal observers. *Invest Ophthalmol Vis Sci*. 1989;30:2056–2064.

30. Tsai JC. A comprehensive perspective on patient adherence to topical glaucoma therapy. *Ophthalmology*. 2009;116:S30–S36.

31. Kinast RM, Mansberger SL. Glaucoma Adherence-From Theriac to the Future. *American Journal of Ophthalmology*. 2018;191:xiii–xv.

32. Gardiner SK, Johnson CA, Demirel S. The effect of test variability on the structure-function relationship in early glaucoma. *Graefes Arch Clin Exp Ophthalmol*. 2012;250:1851–1861.

33. Smith SD, Katz J, Quigley HA. Analysis of progressive change in automated visual fields in glaucoma. *Invest Ophthalmol Vis Sci*. 1996;37: 1419–1428.

34. Spry PG, Johnson CA, McKendrick AM, Turpin A. Variability components of standard automated perimetry and frequency-doubling technology perimetry. *Invest Ophthalmol Vis Sci*. 2001;42: 1404–1410.

35. Spry PG, Johnson CA. Identification of progressive glaucomatous visual field loss. *Surv Ophthalmol*. 2002;47:158–173.

36. Gardiner SK, Fortune B, Demirel S. Signal-to-Noise Ratios for Structural and Functional Tests in Glaucoma. *Transl Vis Sci Technol*. 2013;2:3.

37. Wall M, Kutzko KE, Chauhan BC. Variability in patients with glaucomatous visual field damage is reduced using size V stimuli. *Invest Ophthalmol Vis Sci*. 1997;38:426–435.

38. 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;4:10.

39. Gardiner SK, Mansberger SL. Effect of Restricting Perimetry Testing Algorithms to Reliable Sensitivities on Test-Retest Variability. *Invest Ophthalmol Vis Sci*. 2016;57:5631–5636.

40. Prum BE, Rosenberg LF, Gedde SJ, et al. Primary open-angle glaucoma preferred practice pattern-guidelines. *Ophthalmology*. 2016;123:P41–P111.

41. Heijl A, Bengtsson B. The effect of perimetric experience in patients with glaucoma. *Arch Ophthalmol*. 1996;114:19–22.

42. Gardiner SK, Demirel S, Johnson CA. Is there evidence for continued learning over multiple years in perimetry? *Optom Vis Sci*. 2008;85:1043–1048.

43. Ballae Ganeshrao S, Turpin A, Denniss J, McKendrick Allison M. Enhancing Structure–Function Correlations in Glaucoma with Customized Spatial Mapping. *Ophthalmology*. 2015;122:1695–1705.

44. McKendrick AM, Denniss J, Wang YX, Jonas JB, Turpin A. The Proportion of Individuals Likely to Benefit from Customized Optic Nerve Head Structure-Function Mapping. *Ophthalmology*. 2017;124:554–561.

45. Gazzard G, Foster PJ, Viswanathan AC, et al. The Severity and Spatial Distribution of Visual Field Defects in Primary Glaucoma: A Comparison of Primary Open-Angle Glaucoma and Primary Angle-Closure Glaucoma. *Archives of Ophthalmology*. 2002;120:1636–1643.

46. Yousefi S, Sakai H, Murata H, et al. Asymmetric Patterns of Visual Field Defect in Primary Open-Angle and Primary Angle-Closure Glaucoma. *Investigative Ophthalmology & Visual Science*. 2018;59:1279–1287.

47. Hood DC, Raza AS, de Moraes CG, Liebmann JM, Ritch R. Glaucomatous damage of the macula. *Prog Retin Eye Res*. 2013;32:1–21.

48. De Moraes CG, Hood DC, Thenappan A, et al. 24-2 Visual Fields Miss Central Defects Shown on 10-2 Tests in Glaucoma Suspects, Ocular Hypertensives, and Early Glaucoma. *Ophthalmology*. 2017;124:1449–1456.

translational vision science & technology