



OPEN

## Detection of functional deterioration in glaucoma by trend analysis using comprehensive overlapping clusters of locations

Stuart K. Gardiner<sup>✉</sup> & Steven L. Mansberger

Detecting rapid visual field deterioration is crucial for individuals with glaucoma. Cluster trend analysis detects visual field deterioration with higher sensitivity than global analyses by using predefined non-overlapping subsets of visual field locations. However, it may miss small defects that straddle cluster borders. This study introduces a comprehensive set of overlapping clusters, and assesses whether this further improves progression detection. Clusters were defined as locations from where ganglion cell axons enter the optic nerve head within a  $\theta^\circ$  wide sector, centered at  $1^\circ$  intervals, for various  $\theta$ . Deterioration in eyes with or at risk of glaucomatous visual field loss was “detected” if  $\geq N_\theta$  clusters had deteriorated with  $p < p_{\text{Cluster}}$  chosen empirically to give 95% specificity based on permuting the series.  $N_\theta$  was chosen to minimize the time to detect subsequently-confirmed deterioration in  $\geq 1/3$ rd of eyes. Times to detect deterioration were compared using Cox survival models. Biannual series were available for 422 eyes of 214 participants. Predefined non-overlapping clusters detected subsequently-confirmed change in  $\geq 1/3$ rd of eyes in 3.41 years (95% confidence interval 2.75–5.48 years). After equalizing specificity, no criteria based on comprehensive overlapping clusters detected deterioration significantly sooner. The quickest was 3.13 years (2.69–4.65) for  $\theta^\circ = 20^\circ$  and  $N_\theta = 25$ , but the comparison with non-overlapping clusters had  $p = 0.672$ . Any improvement in sensitivity for detecting deterioration when using a comprehensive set of overlapping clusters was negated by the need to maintain equal specificity. The existing cluster trend analysis using predefined non-overlapping clusters provides a useful tool for monitoring visual field progression.

Patients with glaucoma commonly undergo both functional and structural diagnostic testing as part of their standard clinical care. Functional testing usually takes the form of standard automated perimetry, where the test subject is presented with a series of visual stimuli of different contrast and location, and is asked to respond by pressing a button whenever stimuli are seen<sup>1,2</sup>. Since damage caused by glaucoma is currently irreversible, the aim of this testing is to determine how quickly damage is progressing, so that the clinician can decide whether the current management of the patient’s disease needs to be adjusted to prevent blindness within their lifetime<sup>3,4</sup>. However, glaucoma typically results in localized functional defects, rather than affecting all areas of the visual field equally. Thus, a global average of defect depth is insensitive to changes, with the true progression signal being swamped by considerable test–retest variability<sup>5</sup>. Yet, that same variability also means that estimates of contrast sensitivity at a single location are so variable that it takes several years to determine whether apparent change is real, or even statistically significant<sup>6</sup>. The ideal way to analyze results from clinical perimetry needs to find a balance between the two competing priorities of identifying small localized defects, while retaining the lower variability achieved from averaging information from several locations.

The key to optimizing this balance lies in the fact that the localized functional defects tend to follow specific patterns. Glaucomatous damage to retinal ganglion cells is generally assumed to occur at or near the optic nerve head (ONH)<sup>7</sup>; and so cells whose axons enter the ONH in close proximity to each other are more likely to become damaged at the same time<sup>8,9</sup>. Thus, averaging contrast sensitivity within arcuate clusters based on topographic structure–function mapping has been hypothesized as maximizing the signal-to-noise ratio and thus optimizing monitoring of the rate of disease progression. Specifically, visual field testing is most commonly performed at a regular grid of locations known as the 24–2 test pattern. Axons emanating from retinal ganglion cells whose

Devers Eye Institute, Legacy Research Institute, 1225 NE 2nd Ave, Portland, OR 97232, USA. ✉email: sgardiner@deverseye.org

	Mean	Standard deviation	Range	Interquartile range
Series length (number of visits)	14.0	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 at start of series (years)	63.8	10.9	33.7 to 86.7	56.7 to 71.3
Initial mean deviation (dB)	-0.84	2.8	-17.7 to +3.0	-1.4 to +0.8
Final mean deviation (dB)	-1.71	4.0	-26.3 to +3.0	-2.7 to +0.8
Initial pattern standard deviation (dB)	2.45	2.5	0.9 to 15.7	1.4 to 2.1
Final pattern standard deviation (dB)	3.27	3.3	1.0 to 15.5	1.5 to 3.1
Rate of change of mean deviation (dB/year)	-0.13	0.3	-1.7 to +0.8	-0.2 to +0.0

**Table 1.** Characteristics of the dataset used.

soma (and consequent receptive field) are found at each of these retinal locations have been traced to determine their angle of entry into the ONH<sup>10–12</sup>. If two or more locations correspond with axons that enter the ONH in close proximity, then sensitivities at those locations are more likely to deteriorate at approximately the same rate<sup>9</sup>.

This principle has been applied clinically in the cluster trend analysis within the EyeSuite software developed for the Octopus perimeter (Haag-Streit Inc., Bern, Switzerland)<sup>13</sup>. Pointwise total deviation values (the difference in contrast sensitivity from age-matched normal) are averaged within each of ten predefined non-overlapping clusters of locations. If this average significantly deteriorates over time, then the cluster is flagged as deteriorating<sup>13,14</sup>. We previously showed that this approach detects glaucomatous progression sooner than global analyses, and has a higher probability than pointwise analyses that any detected deterioration will be subsequently confirmed upon further testing<sup>15</sup>.

However, there are two potentially sub-optimal aspects of using ten non-overlapping clusters in this way. Firstly, localized defects are not anatomically constrained to these predefined clusters, but could straddle the border between two of the clusters<sup>15</sup>. Secondly, the topographic structure–function relation varies between eyes due to the individualized anatomy, including physiologic variations in the exact position of the ONH relative to the visual field test locations<sup>16,17</sup>. Thus, some localized defects may be underestimated if they do not correspond perfectly with any single cluster, especially if they affect only half or fewer of the cluster's locations.

Recently we took the first step towards assessing the clinical impact of these sub-optimal aspects of cluster analysis, by adding an additional set of eleven overlapping clusters to the original ten non-overlapping clusters from the EyeSuite software. However, using these 21 overlapping clusters did not significantly reduce the time to detect deterioration (for equal specificity) when compared to only using the original ten predefined non-overlapping clusters<sup>18</sup>. While that study represented important progress, it still relied on using the same relatively small set of fixed clusters for every eye. We therefore wanted to explore whether a more comprehensive set of overlapping clusters would provide optimal performance when compared to the original ten-cluster approach.

In this study, we describe a method to detect deterioration using comprehensive overlapping clusters of visual field locations. This method looks for deterioration among a greater number of clusters, and could therefore reduce specificity. We therefore use a permutation analysis technique<sup>19</sup> to determine whether the hypothesized increase in sensitivity outweighs the potential reduction in specificity. If successful, use of a comprehensive set of overlapping clusters could improve the ability of both clinicians and researchers to rapidly and reliably detect visual field deterioration.

## Results

Series of at least five reliable visual fields were available for 422 eyes of 214 participants. 146 of those eyes (35%) had an abnormal result on the first visual field in their series, defined as either abnormal Pattern Standard Deviation ( $p < 5\%$ ) or a Glaucoma Hemifield Test result of either “Abnormal” or “Borderline”. 163 eyes (39%) were recorded as having been prescribed IOP-lowering medications. Table 1 summarizes other characteristics of the cohort.

Table 2 shows the time taken to detect deterioration, and the time to detect confirmed deterioration (i.e. the series was still “deteriorating” after the inclusion of the next test date in the analysis), in at least one third of the cohort (i.e. the lower tertile of survival times) using a Kaplan–Meier survival model, based on a range of different criteria. Since testing was conducted as close as possible to biannually, and only series of length  $\geq 5$  visits were analyzed, the first possible date at which deterioration could be detected was approximately 2 years. All the criteria based on comprehensive overlapping clusters detected confirmed deterioration significantly sooner than MD. However, none of them detected confirmed deterioration significantly sooner using the ten predefined non-overlapping clusters from the EyeSuite software.

The cluster criterion for cluster width  $\theta^\circ$  could be considered optimal for detecting defects that cover exactly  $\theta^\circ$  at the optic nerve head, but sub-optimal for defects of other widths. To test this, the above analysis was repeated using all available clusters of width either  $10^\circ$  or  $30^\circ$ , rather than just one width. However, this still did not detect subsequently-confirmed deterioration significantly sooner than using the ten predefined non-overlapping clusters ( $p = 0.497$ ) after equalizing specificity. Neither did a similar analysis using clusters of width either  $10^\circ$  or  $60^\circ$  ( $p = 0.464$ ). Using clusters of width either  $30^\circ$  or  $60^\circ$  actually detected subsequently-confirmed deterioration slightly slower than the predefined non-overlapping clusters, with  $p = 0.009$ .

Criterion definition	Optimal number of clusters, $N_\theta$	Years to detect deterioration in $\geq 33\%$ of eyes (95% confidence interval)	Years to detect confirmed deterioration in $\geq 33\%$ of eyes (95% confidence interval)	Probability deterioration is confirmed on next visit	Comparison against mean deviation	Comparison against predefined clusters
Mean deviation	1	4.51 (3.28– $\infty$ )	7.22 (3.41– $\infty$ )	92.2%		
Predefined clusters	1	3.28 (2.73–5.06)	3.41 (2.75–5.48)	94.4%	< 0.001	
$\theta = 10^\circ$ clusters	11	3.07 (2.72–4.17)	3.30 (2.74–5.90)	92.4%	< 0.001	0.806
$\theta = 20^\circ$ clusters	25	3.03 (2.69–4.16)	3.13 (2.69–4.65)	93.0%	< 0.001	0.672
$\theta = 30^\circ$ clusters	25	3.12 (2.72–4.17)	3.23 (2.73–4.51)	95.1%	< 0.001	0.654
$\theta = 45^\circ$ clusters	37	3.24 (2.74–4.48)	3.32 (2.75–9.10)	93.3%	< 0.001	0.341
$\theta = 60^\circ$ clusters	47	3.07 (2.72–4.45)	3.30 (2.74–9.10)	91.8%	< 0.001	0.356
$\theta = 90^\circ$ clusters	34	3.28 (2.74–4.45)	3.41 (2.77– $\infty$ )	91.1%	0.001	0.194
$\theta = 10^\circ$ unique clusters	4	3.07 (2.72–4.20)	3.41 (2.76– $\infty$ )	86.7%	0.002	0.265
$\theta = 20^\circ$ unique clusters	5	3.28 (2.75–4.20)	3.37 (2.77–5.35)	93.4%	< 0.001	0.685
$\theta = 30^\circ$ unique clusters	6	3.14 (2.75–4.48)	3.41 (2.81– $\infty$ )	89.0%	0.002	0.067
$\theta = 45^\circ$ unique clusters	8	3.14 (2.73–4.47)	3.58 (2.92– $\infty$ )	87.6%	0.002	0.076
$\theta = 60^\circ$ unique clusters	10	3.14 (2.74–4.50)	3.30 (2.76–6.98)	91.8%	< 0.001	0.482
$\theta = 90^\circ$ unique clusters	17	3.28 (2.75–5.34)	3.43 (2.77– $\infty$ )	90.3%	0.003	0.076

**Table 2.** Comparison of the criteria for detection deterioration. For a given cluster width  $\theta^\circ$  (left column), deterioration was “detected” on the first visit at which  $N_\theta$  clusters were deteriorating with  $p < p_{\text{Overall}}$  with this cutoff chosen to give exactly 95% specificity as detailed in the Methods. For the first set of results (labelled “ $\theta = 10^\circ$  clusters” etc.), one cluster was defined centered at each angle around the optic nerve head in  $1^\circ$  increments, subject to the cluster containing at least two visual field locations. For the second set of results (labelled “unique clusters”), any clusters that contained exactly the same set of field locations were only counted once. The number of clusters  $N_\theta$  was chosen to minimize time to detect confirmed deterioration, and is shown in the second column. The next two columns give the time to detect deterioration, or subsequently-confirmed deterioration, respectively, with 95% confidence intervals. A value of  $\infty$  indicates that too few eyes met the criterion by the end of their available series of data for this to be calculated. The next column shows the probability that deterioration was subsequently confirmed, i.e. an eye that met the criterion for deterioration still met the criterion after the addition of the next visual field in the series. The final columns show p-values for whether the criterion detected confirmed deterioration significantly sooner than using Mean Deviation, or than using the ten predefined non-overlapping clusters of locations.

Table 3 shows the time to detect confirmed deterioration, and comparison against the ten predefined non-overlapping clusters, for two subsets of the cohort: 276 eyes without and 146 eyes with an abnormal visual field at the start of their series. Only 32% of eyes without an initially abnormal field showed confirmed deterioration by MD before the end of their series, hence the estimated time for 1/3rd of eyes to show confirmed deterioration in Table 3 is infinite. Unsurprisingly, deterioration was detected sooner in eyes that were abnormal at baseline, as they are more likely to progress rapidly<sup>20</sup>. However, while the time to detect deterioration in 1/3rd of eyes appeared shorter using some of the comprehensive overlapping cluster criteria than using predefined non-overlapping clusters, these differences were never statistically significant. Figure 1 shows Kaplan–Meier survival plots for a selection of criteria, using the optimal number of clusters for detecting confirmed deterioration in each case.

## Discussion

Previously, we have shown that cluster trend analysis, assessing the significance of the rate of change within each of ten predefined non-overlapping clusters of test locations, provided a useful clinical tool for assessing glaucomatous progression. It detected deterioration sooner than using global metrics such as MD; and any changes that were detected were more likely to be confirmed upon subsequent retesting than when using pointwise analyses. It thus provides a useful compromise between these two competing priorities. However, we hypothesized that it could be further improved, since it could ‘miss’ defects that straddle cluster borders<sup>9</sup>, and/or are in eyes whose anatomy dictates a different expected topographic relation between ONH damage and visual field defect<sup>17</sup>. In this study, we describe a novel method to avoid those problems using an objectively-chosen set of comprehensive overlapping clusters of visual field locations. Our new technique aims to ensure that if there is a nerve fiber layer defect of width  $\theta^\circ$  at the ONH, then there will be at least one of the visual field clusters based on  $\theta^\circ$  width that exactly corresponds with this defect. However, increasing the number of possible clusters in this way causes problems with multiple comparisons, and so the criterion for deterioration has to be adjusted accordingly to achieve the same overall specificity. After making these adjustments to equalize specificity, the new technique did not detect deterioration significantly sooner than the simpler alternative of using the ten predefined non-overlapping clusters as implemented in the current clinical software<sup>13–15</sup>.

The strength of cluster trend analysis, compared with global or pointwise analyses, is that it better reflects the typical glaucomatous disease process which causes arcuate field loss due to the spatial arrangement of nerve fiber axon bundles in the retina<sup>9</sup>. Generalized loss, affecting the entire visual field approximately equally, would presumably be better detected by global indices. Such loss does appear in glaucoma<sup>21–24</sup>, but is of relatively

Criterion definition	Normal visual field at baseline		Abnormal visual field at baseline	
	Years to detect confirmed deterioration in $\geq 33\%$ of eyes (95% confidence interval)	Comparison against predefined clusters	Years to detect confirmed deterioration in $\geq 33\%$ of eyes (95% confidence interval)	Comparison against predefined clusters
Mean deviation	$\infty$ (3.85– $\infty$ )		3.55 (2.65– $\infty$ )	
Predefined clusters	3.30 (2.72–9.10)		3.28 (2.65–6.98)	
$\theta = 10^\circ$ clusters	3.41 (2.72– $\infty$ )	0.442	3.07 (2.63–5.90)	0.740
$\theta = 20^\circ$ clusters	3.32 (2.69– $\infty$ )	0.993	2.77 (2.61–6.98)	0.731
$\theta = 30^\circ$ clusters	3.31 (2.73– $\infty$ )	0.144	2.77 (2.61–6.98)	0.085
$\theta = 45^\circ$ clusters	3.76 (2.81– $\infty$ )	0.985	2.75 (2.60–5.53)	0.237
$\theta = 60^\circ$ clusters	3.58 (2.73– $\infty$ )	0.771	2.77 (2.63– $\infty$ )	0.338
$\theta = 90^\circ$ clusters	4.01 (2.75– $\infty$ )	0.968	3.07 (2.65–6.98)	0.811
$\theta = 10^\circ$ unique clusters	3.58 (2.76– $\infty$ )	0.700	3.28 (2.65– $\infty$ )	0.329
$\theta = 20^\circ$ unique clusters	3.61 (2.76– $\infty$ )	0.907	3.07 (2.65–6.98)	0.668
$\theta = 30^\circ$ unique clusters	3.76 (2.81– $\infty$ )	0.907	3.28 (2.65– $\infty$ )	0.536
$\theta = 45^\circ$ unique clusters	4.17 (2.96– $\infty$ )	0.478	3.28 (2.65– $\infty$ )	0.825
$\theta = 60^\circ$ unique clusters	3.99 (2.81– $\infty$ )	0.609	2.77 (2.61–4.64)	0.907
$\theta = 90^\circ$ unique clusters	4.16 (2.76– $\infty$ )	0.937	3.07 (2.63–8.64)	0.295

**Table 3.** Comparison of the criteria for detection deterioration as in Table 2, but differentiated into eyes with and without an abnormal visual field at baseline. For each subset, the two columns give the time to detect subsequently-confirmed deterioration, with 95% confidence interval in parentheses; and the p-value for whether the criterion detected confirmed deterioration significantly sooner than when using ten predefined non-overlapping clusters of locations. A value of  $\infty$  indicates that too few eyes met the criterion by the end of their available series of data for this to be calculated.

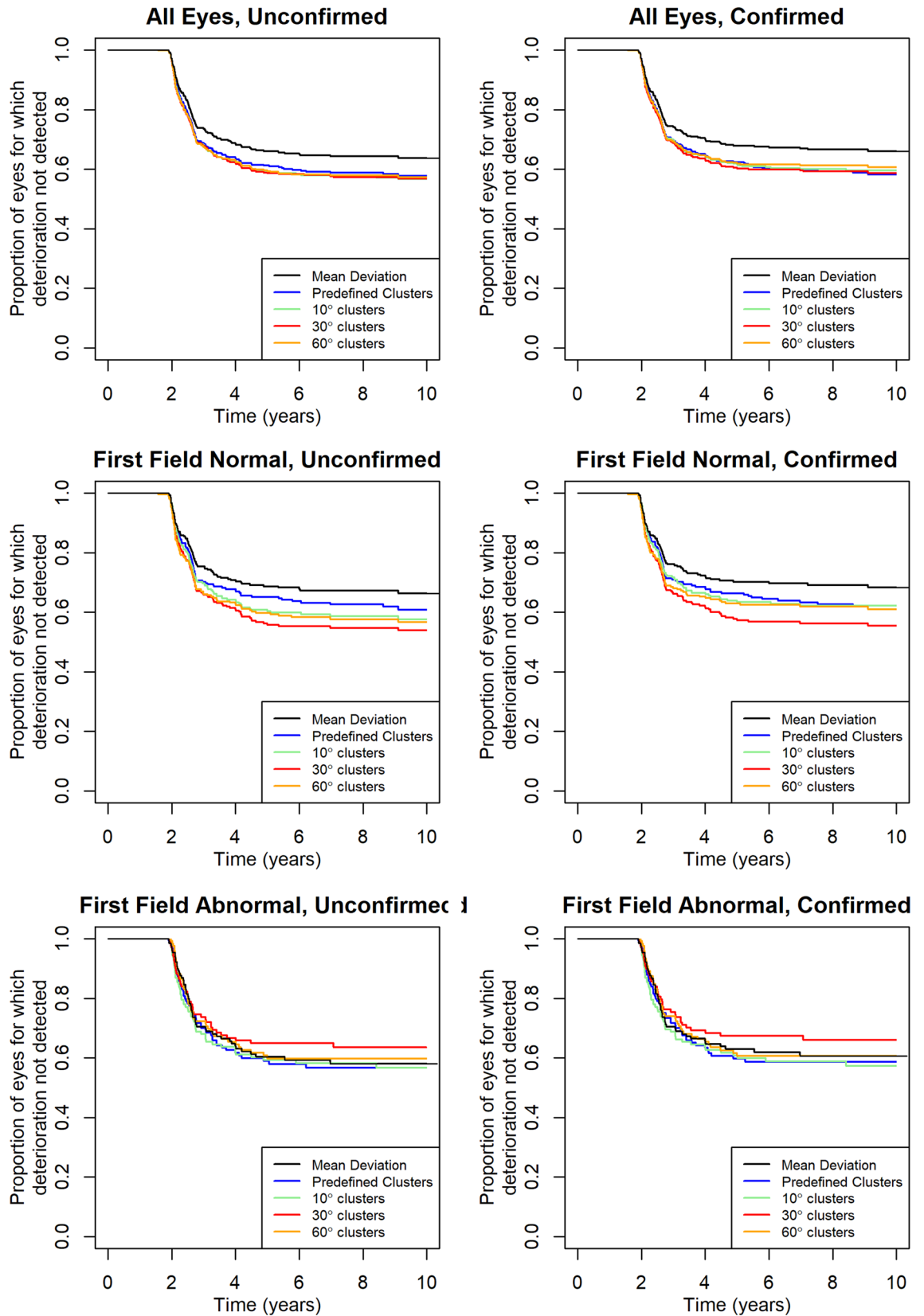
small magnitude compared with localized scotoma; thus deterioration can be detected sooner when focusing on localized changes<sup>15</sup>. Defects that cover only a single visual field location in the 24–2 test pattern can also occur, especially in the central region of the field<sup>25</sup>; but apparent loss at a single location is often just due to variability and so may not be subsequently confirmed<sup>26–28</sup>. The ideal balance between these issues will depend on the clinical situation. In particular, a small localized defect may be ‘trusted’ more if there is a structural defect at the corresponding location in the nerve fiber layer. No single analytic tool is optimal in all clinical scenarios. However, there is now compelling evidence that cluster-based analyses should be one of the set of tools made available to clinicians; and that perceived flaws such as its reliance on a limited predefined set of clusters do not significantly reduce its usefulness.

Although testing was performed using clinical instruments and protocols, there are still potentially important differences between this study and typical clinical practice. Eyes were tested once every six months, or as close as could be scheduled, whereas it would be more common clinically to test people more frequently if rapid progression were suspected<sup>29</sup>. Study participants are also highly experienced with automated perimetry, and there may be greater benefits of averaging information from larger numbers of locations in less-experienced patients who typically have higher test–retest variability<sup>30,31</sup>.

The majority of the cohort had either no visual field loss or early loss; even at the end of the series the average MD was  $-1.7$  dB. Both eyes were tested, even if clinically only one eye would be considered glaucomatous. Eyes with glaucoma were being managed clinically to slow their rates of progression, and given that these participants were motivated enough to participate in a long-term study, they could be hypothesized to have greater compliance with their prescribed medications than a more general population. For all these reasons, we would not expect many eyes to show rapid progression<sup>20,32</sup>. There are no obvious reasons to suppose that comparisons between criteria would be different in a cohort of eyes undergoing more rapid progression, and indeed an eye that is progressing sufficiently rapidly would be detected by any of the tested criteria; but such a differential effect cannot be ruled out.

The analysis technique used in this study relies on ordinary least squares regression for each series. However, the overall p-values are derived empirically by comparison against the permutation distribution. This reduces caveats concerning the validity of the underlying assumptions of the analysis, in particular with regard to normality of the residuals. Some eyes could be more variable than others, both due to individual factors<sup>33</sup> and due to the increase in variability with disease severity<sup>34</sup>. A mixed effects model would typically assume homogeneity of the residuals, but permutation analysis only requires the much weaker assumption of homogeneity within the series for an individual eye. Another caveat with the analysis is that no adjustment was made for multiple comparisons using different sector widths  $\theta^\circ$ , and so the probability of finding a statistically significant difference was inflated; yet even without such adjustment, no such significant differences were found.

In conclusion, we found that although cluster trend analysis as implemented in the clinical EyeSuite software uses only ten predefined non-overlapping clusters of locations instead of a more comprehensive evaluation of possible clusters, this does not significantly delay the detection of visual field deterioration. Any benefit from identifying more defects using a more comprehensive set of clusters was negated by the adjustments to the criteria needed to maintain specificity. Cluster trend analysis provides a useful tool for monitoring deterioration in glaucomatous visual fields.



**Figure 1.** Kaplan–Meier survival plots showing the time until deterioration (left) and subsequently-confirmed deterioration (right) was detected by a selection of criteria. Plots in the top row include all 422 eyes; plots in the second row include only the 276 eyes with normal visual field at baseline; and plots in the third row include only the 146 eyes with abnormal visual field at baseline.



## Methods

**Participants.** The same data were used as in our previous study<sup>18</sup>. This was a retrospective analysis of data taken from the ongoing Portland Progression Project (P3), conducted at Legacy Devers Eye Institute, in which participants undergo a range of structural and functional testing once every six months. Inclusion criteria were a diagnosis of primary open-angle glaucoma and/or likelihood of developing glaucomatous damage, as determined subjectively by each participant's physician, in order 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, significant cataract or media opacities likely to significantly increase light scatter, or other conditions or medications that affect the visual field. Standard automated perimetry was performed using a Humphrey Field Analyzer HFAIIi perimeter, with the 24–2 test pattern, a size III white-on-white stimulus, and the SITA Standard algorithm<sup>35</sup>. Eyes were included in the analyses for this study if they had at least five reliable tests, defined as  $\leq 15\%$  false positives and  $\leq 33\%$  fixation losses. All protocols 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.

**Analysis—definition of comprehensive overlapping clusters.** The comprehensive clustering system used in this study relies on the structure–function map published by Garway-Heath et al.<sup>10</sup> In that study, axon bundles were manually traced on nerve fiber layer photographs from each test location to the point at which they entered the ONH. This gave an average angle of entry for axons corresponding with each location in the 24–2 test pattern. A “cluster” of visual field locations in this study was defined as the set of locations whose corresponding axons enter the ONH between  $\alpha - (\theta/2)^\circ$  and  $\alpha + (\theta/2)^\circ$ , for a chosen center  $\alpha^\circ$  and width  $\theta^\circ$ .

The exact topographic relation varies; the reported standard deviation of these traced angles between eyes was  $7.2^\circ$ <sup>10</sup>. A structural defect in the nerve fiber layer could also be centered at any angle around the ONH. Therefore, one cluster was formed for every possible center  $\alpha^\circ$ , at  $1^\circ$  intervals. Thus, any nerve fiber layer defect of width  $\theta^\circ$  can be assumed to exactly correspond with one of these clusters. This remains true even if the rotational error from the average eye is not constant around the disc, for example due to inter-individual differences in axial length<sup>16</sup>. To avoid excessive variability, only clusters containing at least two visual field locations were considered. Hence, there were fewer than 360 clusters for any given width  $\theta^\circ$ , with the actual number increasing with  $\theta^\circ$ . Figure 2 shows the average angle of entry to the ONH for axons corresponding to each visual field location in a right eye, together with three of the clusters based on a width of  $\theta^\circ = 30^\circ$ . Note that several clusters centered at neighboring angles  $\alpha^\circ$  could consist of the same set of visual field locations. For the primary analysis, these were considered as separate clusters. A secondary analysis was performed using only “unique clusters”, in which the same set of visual field locations is never included more than once.

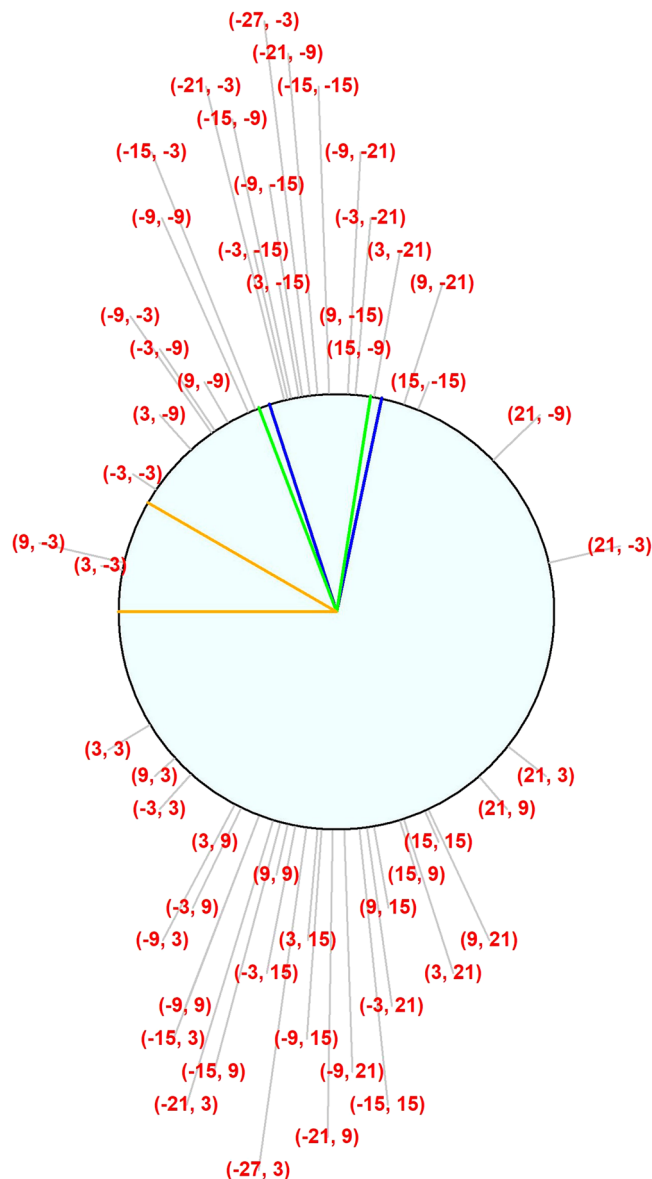
**Analysis—detecting deterioration.** The total deviation (on the native decibel scale) was averaged across locations within each eligible cluster of width  $\theta^\circ$ <sup>15</sup>. Total deviation values for each location were used instead of raw sensitivities to account for the effect of normal 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. Thus, total deviation values for any such locations were set to equal the total deviation value for a sensitivity of 15 dB<sup>36,37</sup>. All analyses were performed using the R statistical programming language (Version 4.0.0)<sup>38</sup>.

Seeking deterioration in any one of multiple clusters would be expected to increase sensitivity, but at the cost of reduced specificity. In order to fairly compare criteria given the competing demands of sensitivity vs. specificity, we sought to establish criteria for whether an eye is deteriorating with exactly 95% specificity, based on the previously published Permutation Analyses of Pointwise Linear Regression (PoPLR) approach<sup>19</sup>.

For each cluster of locations,  $p_{Cluster}$  was defined as the p-value from an ordinary least squares regression of the average total deviation against time. For a given number of clusters  $N$ ,  $p_{Overall}$  was defined as the  $N$ th smallest of these  $p_{Cluster}$  values. A permutation distribution for  $p_{Overall}$  was derived by repeatedly reordering visual fields  $1 - V$ , but retaining the original test dates. For  $V = 5$ , this was done for all 120 possible reorderings of the 5 visits; for  $V > 5$ , 475 randomly-chosen reorderings were used, to avoid excessive computation time (475 reorderings allows a specificity of 95% to be calculated with a confidence interval of  $\pm 1\%$  based on a binomial distribution). “Deterioration” was detected on the first visit  $V$  for which  $p_{Overall}$  was below the 5<sup>th</sup> percentile of the permutation distribution. Thus, a group of criteria are derived each with specificity exactly equal to 95%, but with different numbers of clusters  $N$ , and different cluster widths  $\theta^\circ$ .

**Analysis—confirmation of deterioration.** “Confirmed deterioration” was detected on the first visit  $V$  at which the  $N$ th smallest observed p-value was below the 5th percentile of the  $N$ th smallest p-values from all reorderings, for both 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  $N$ th 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<sup>26,39</sup>.

**Analysis—comparison of criteria.** For each width  $\theta^\circ$ , the optimal number of clusters  $N_\theta$  was chosen as the criteria that detected subsequently-confirmed deterioration in  $\geq 33\%$  of eyes soonest, based on a Kaplan–Meier survival analysis. In the event of a tie, the smallest such number  $N$  was used as  $N_\theta$  thereafter. The optimal criteria using  $N_\theta$  clusters of width  $\theta^\circ$  were compared against each other; and also against using  $N$  of the ten



**Figure 2.** Illustration of the clustering system used. Numbers in red give the coordinates (in degrees) of each visual field location for a right eye. Gray lines show the angle at which the corresponding axons enter the optic nerve head. The 30° sector shown by the green lines contains 12 visual field locations. The 30° sector shown by the blue lines contains the same 12 visual field locations, plus one more location (3, -21). The 30° sector shown by the orange lines only contains 2 visual field locations, (9, -3) and (3, -3).

predefined non-overlapping clusters from the EyeSuite software in the same manner, or using Mean Deviation (effectively the same analysis as above but with  $N=1$  and  $\theta=360^\circ$ ).

As in our previous study<sup>18</sup>, for each criterion, Kaplan–Meier survival analysis was used to determine the time taken until  $\geq 33\%$  of eyes had shown “deterioration” or “confirmed deterioration”. 95% confidence intervals for these times were found using standard errors based on Greenwood’s formula<sup>40</sup>. Survival curves were compared using a stratified Cox proportional hazards model<sup>41</sup>, with strata identifying fellow eyes of the same individual<sup>42</sup>. Sub-analyses 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.

### Data availability

The datasets analyzed during the current study are available from the corresponding author on reasonable request.

Received: 22 August 2020; Accepted: 13 October 2020

Published online: 28 October 2020

## References

1. Flammer, J., Drance, S. M., Augustiny, L. & Funkhouser, A. Quantification of glaucomatous visual field defects with automated perimetry. *Invest. Ophthalmol. Vis. Sci.* **26**, 176–181 (1985).
2. Anderson, D. & Patella, V. *Automated Static Perimetry*. 2 edn (Mosby, 1999).
3. Caprioli, J. The importance of rates in glaucoma. *Am. J. Ophthalmol.* **145**, 191–192 (2008).
4. Saunders, L. J., Russell, R. A., Kirwan, J. F., McNaught, A. I. & Crabb, D. P. Examining visual field loss in patients in glaucoma clinics during their predicted remaining lifetime. *Invest. Ophthalmol. Vis. Sci.* **55**, 102–109 (2014).
5. Smith, S. D., Katz, J. & Quigley, H. A. Analysis of progressive change in automated visual fields in glaucoma. *Invest. Ophthalmol. Vis. Sci.* **37**, 1419–1428 (1996).
6. Chauhan, B. C. *et al.* Practical recommendations for measuring rates of visual field change in glaucoma. *Br. J. Ophthalmol.* **92**, 569–573 (2008).
7. Quigley, H. A., Addicks, E. M., Green, W. & Maumenee, A. E. Optic nerve damage in human glaucoma: II. The site of injury and susceptibility to damage. *Arch. Ophthalmol.* **99**, 635–649 (1981).
8. Gardiner, S. K., Crabb, D. P., Fitzke, W. & Hitchings, R. A. Reducing noise in suspected glaucomatous visual fields by using a new spatial filter. *Vis. Res.* **44**, 839–848 (2004).
9. Gardiner, S. K., Johnson, C. A. & Cioffi, G. A. Evaluation of the structure-function relationship in glaucoma. *Invest. Ophthalmol. Vis. Sci.* **46**, 3712–3717 (2005).
10. Garway-Heath, D. F., Poinoosawmy, D., Fitzke, F. W. & Hitchings, R. A. Mapping the visual field to the optic disc in normal tension glaucoma eyes. *Ophthalmology* **107**, 1809–1815 (2000).
11. Jansonius, N. M. *et al.* A mathematical description of nerve fiber bundle trajectories and their variability in the human retina. *Vis. Res.* **49**, 2157–2163 (2009).
12. Denniss, J., McKendrick, A. M. & Turpin, A. An anatomically customizable computational model relating the visual field to the optic nerve head in individual eyes. *Invest. Ophthalmol. Vis. Sci.* **53**, 6981–6990 (2012).
13. Racette, L. *et al.* in *Visual Field Digest: A Guide to Perimetry and the Octopus Perimeter* (8th edn.) (ed Haag-Streit AG) 165–192 (2019).
14. Naghizadeh, F. & Hollo, G. Detection of early glaucomatous progression with octopus cluster trend analysis. *J. Glaucoma* **23**, 269–275 (2014).
15. Gardiner, S. K., Mansberger, S. L. & Demirel, S. Detection of functional change using cluster trend analysis in glaucoma. *Invest. Ophthalmol. Vis. Sci.* **58**, Bio180–Bio190 (2017).
16. Ballae Ganeshrao, S., Turpin, A., Denniss, J. & McKendrick, A. M. Enhancing structure–function correlations in glaucoma with customized spatial mapping. *Ophthalmology* **122**, 1695–1705 (2015).
17. McKendrick, A. M., Denniss, J., Wang, Y. X., Jonas, J. B. & Turpin, A. The proportion of individuals likely to benefit from customized optic nerve head structure-function mapping. *Ophthalmology* **124**, 554–561 (2017).
18. Gardiner, S. K. & Mansberger, S. L. Detection of functional deterioration in glaucoma by trend analysis using overlapping clusters of locations. *Transl. Vis. Sci. Technol.* **9**, 12–12 (2020).
19. O’Leary, N., Chauhan, B. C. & Artes, P. H. Visual field progression in glaucoma: estimating the overall significance of deterioration with permutation analyses of pointwise linear regression (PoPLR). *Invest. Ophthalmol. Vis. Sci.* **53**, 6776–6784 (2012).
20. Gardiner, S. K., Demirel, S. & Johnson, C. A. Perimetric indices as predictors of future glaucomatous functional change. *Optom. Vis. Sci.* **88**, 56–62 (2011).
21. Drance, S. M. Diffuse visual field loss in open-angle glaucoma. *Ophthalmology* **98**, 1533–1538 (1991).
22. Chauhan, B. C., LeBlanc, R. P., Shaw, A. M., Chan, A. B. & McCormick, T. A. Repeatable diffuse visual field loss in open-angle glaucoma. *Ophthalmology* **104**, 532–538 (1997).
23. Henson, D. B., Artes, P. H. & Chauhan, B. C. Diffuse loss of sensitivity in early glaucoma. *Invest. Ophthalmol. Vis. Sci.* **40**, 3147–3151 (1999).
24. Grewal, D. S., Sehi, M. & Greenfield, D. S. Diffuse glaucomatous structural and functional damage in the hemifield without significant pattern loss. *Arch. Ophthalmol.* **127**, 1442–1448 (2009).
25. Hood, D. C., Raza, A. S., de Moraes, C. G., Liebmann, J. M. & Ritch, R. Glaucomatous damage of the macula. *Prog. Retin. Eye Res.* **32**, 1–21 (2013).
26. Keltner, J. L. *et al.* Confirmation of visual field abnormalities in the Ocular Hypertension Treatment Study. Ocular Hypertension Treatment Study Group. *Arch. Ophthalmol.* **118**, 1187–1194 (2000).
27. Gardiner, S. K. & Crabb, D. P. Examination of different pointwise linear regression methods for determining visual field progression. *Invest. Ophthalmol. Vis. Sci.* **43**, 1400–1407 (2002).
28. Lee, A. C. *et al.* Infrequent confirmation of visual field progression. *Ophthalmology* **109**, 1059–1065 (2002).
29. Prum, B. E. *et al.* Primary open-angle glaucoma preferred practice pattern\* guidelines. *Ophthalmology* **123**, P41–P111 (2016).
30. Heijl, A. & Bengtsson, B. The effect of perimetric experience in patients with glaucoma. *Arch. Ophthalmol.* **114**, 19–22 (1996).
31. Gardiner, S. K., Demirel, S. & Johnson, C. A. Is there evidence for continued learning over multiple years in perimetry?. *Optom. Vis. Sci.* **85**, 1043–1048 (2008).
32. Gardiner, S. K., Johnson, C. A. & Demirel, S. Factors predicting the rate of functional progression in early and suspected glaucoma. *Invest. Ophthalmol. Vis. Sci.* **53**, 3598–3604 (2012).
33. Junoy Montolio, F. G., Wesselink, C., Gordijn, M. & Jansonius, N. M. Factors that influence standard automated perimetry test results in glaucoma: Test reliability, technician experience, time of day, and season. *Invest. Ophthalmol. Vis. Sci.* **53**, 7010–7017 (2012).
34. Henson, D. B., Chaudry, S., Artes, P. H., Faragher, E. B. & Ansons, A. Response variability in the visual field: Comparison of optic neuritis, glaucoma, ocular hypertension, and normal eyes. *Invest. Ophthalmol. Vis. Sci.* **41**, 417–421 (2000).
35. Bengtsson, B., Olsson, J., Heijl, A. & Rootzen, H. A new generation of algorithms for computerized threshold perimetry, SITA. *Acta Ophthalmol. Scand.* **75**, 368–375 (1997).
36. Gardiner, S. K., Swanson, W. H. & Demirel, S. The effect of limiting the range of perimetric sensitivities on pointwise assessment of visual field progression in glaucoma. *Invest. Ophthalmol. Vis. Sci.* **57**, 288–294 (2016).
37. Gardiner, S. K., Swanson, W. H., Goren, D., Mansberger, S. L. & Demirel, S. Assessment of the reliability of standard automated perimetry in regions of glaucomatous damage. *Ophthalmology* **121**, 1359–1369 (2014).
38. *R: A Language and Environment for Statistical Computing v. 1.9.1* (R Foundation for Statistical Computing, Vienna, Austria, 2013).
39. Keltner, J. L. *et al.* Normal visual field test results following glaucomatous visual field end points in the Ocular Hypertension Treatment Study. *Arch. Ophthalmol.* **123**, 1201–1206 (2005).
40. Greenwood, M. The errors of sampling of the survivorship tables. *Rep. Public Health Stat. Subj.* **33**, 26 (1926).
41. Cox, D. R. Regression models and life-tables. *J. R. Stat. Soc. Ser. B (Methodological)* **34**, 187–220 (1972).
42. O’Quigley, J. & Stare, J. Proportional hazards models with frailties and random effects. *Stat. Med.* **21**, 3219–3233 (2002).



### Author contributions

S.K.G.: Study conception, study design, data acquisition, data analysis, data interpretation, drafting manuscript, approval of final version. S.L.M.: Study design, data interpretation, revising manuscript, approval of final version.

### Funding

NIH NEI grant R01-EY020922 (SKG); unrestricted research support from The Legacy Good Samaritan Foundation, Portland, OR, USA. The sponsors and funding organizations had no role in the design or conduct of this research.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to S.K.G.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://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 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

© The Author(s) 2020