tvst

DOI: 10.1167/tvst.5.3.2

Article

Unsupervised Gaussian Mixture-Model With Expectation Maximization for Detecting Glaucomatous Progression in Standard Automated Perimetry Visual Fields

Siamak Yousefi[1], Madhusudhanan Balasubramanian[2], Michael H. Goldbaum[1], Felipe A. Medeiros[1], Linda M. Zangwill[1], Robert N. Weinreb[1], Jeffrey M. Liebmann[3], Christopher A. Girkin[4], and Christopher Bowd[1]

[1] Hamilton Glaucoma Center and the Department of Ophthalmology, University of California San Diego, La Jolla, CA, USA
[2] Department of Electrical and Computer Engineering; Department of Biomedical Engineering, University of Memphis, Memphis, TN, USA
[3] Department of Ophthalmology, New York University, New York, NY, USA
[4] Department of Ophthalmology, University of Alabama, Birmingham, AL, USA

Correspondence: Christopher Bowd, Hamilton Glaucoma Center – 178, Department of Ophthalmology, University of California, San Diego, La Jolla, CA 92037-0946, USA. email: cbowd@ucsd.edu

Received: 27 March 2015
Accepted: 6 March 2016
Published: 3 May 2016

Keywords: glaucoma; standard automated perimetry; visual field; unsupervised machine learning; Gaussian mixture-model; progression detection.

Citation: 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. Trans Vis Sci Tech. 2016; 5(3):2, doi:10.1167/tvst.5.3.2

Purpose. To validate Gaussian mixture-model with expectation maximization (GEM) and variational Bayesian independent component analysis mixture-models (VIM) for detecting glaucomatous progression along visual field (VF) defect patterns (GEM–progression of patterns (POP) and VIM-POP). To compare GEM-POP and VIM-POP with other methods.

Methods. GEM and VIM models separated cross-sectional abnormal VFs from 859 eyes and normal VFs from 1117 eyes into abnormal and normal clusters. Clusters were decomposed into independent axes. The confidence limit (CL) of stability was established for each axis with a set of 84 stable eyes. Sensitivity for detecting progression was assessed in a sample of 83 eyes with known progressive glaucomatous optic neuropathy (PGON). Eyes were classified as progressed if any defect pattern progressed beyond the CL of stability. Performance of GEM-POP and VIM-POP was compared to point-wise linear regression (PLR), permutation analysis of PLR (PoPLR), and linear regression (LR) of mean deviation (MD), and visual field index (VFI).

Results. Sensitivity and specificity for detecting glaucomatous VFs were 89.9% and 93.8%, respectively, for GEM and 93.0% and 97.0%, respectively, for VIM. Receiver operating characteristic (ROC) curve areas for classifying progressed eyes were 0.82 for VIM-POP, 0.86 for GEM-POP, 0.81 for PoPLR, 0.69 for LR of MD, and 0.76 for LR of VFI.

Conclusions. GEM-POP was significantly more sensitive to PGON than PoPLR and linear regression of MD and VFI in our sample, while providing localized progression information.

Translational Relevance. Detection of glaucomatous progression can be improved by assessing longitudinal changes in localized patterns of glaucomatous defect identified by unsupervised machine learning.

Introduction

Glaucoma is a blinding optic neuropathy that may cause significant visual impairment when left untreated. It is the second leading cause of blindness worldwide.[1–4] Detection of glaucomatous visual function defects and detection of their progression are critical for management of the disease. Identifying patterns of visual function defects and tracking their change over time likely is a promising approach for clinical management of glaucoma.[5]

Perimetric visual fields (VF) are used routinely in clinical practice to assess visual function defects attributable to glaucoma. In standard automated perimetry (SAP), the status (e.g., within normal limits or outside of normal limits) of 52 test locations (of the 24-2 test pattern) is determined by statistical comparison of the test measurements with a normative

database composed of age normalized SAP exams. Linear regression of the commercially available SAP software parameters mean deviation (MD) and visual field index (VFI) often are used to assess progression of visual function deficits, over time. Also commercially available, and designed specifically for progression detection, is the Guided Progression Analysis algorithm (GPA).[6] Because MD and VFI are global indices, these methods may not be ideal for progression detection, in that they include visual field locations that have little impact on the VF progression.

Since as early as 1990 (Goldbaum MH, et al. *IOVS* 1990;31;ARVO Abstract 503), studies have used supervised machine-learning classifiers successfully to separate healthy from glaucomatous eyes based on VF and optical imaging measurements and to predict conversion to glaucoma in glaucoma suspect eyes.[7–23] More recently, we have effectively employed unsupervised machine-learning techniques to discern how VF data are organized into patterns. We found it useful to represent the structure of VFs by clusters of healthy eyes, early glaucoma eyes, and advanced glaucoma eyes, and to represent the structure within each cluster by axes obtained by independent component analysis. The estimation of the best structure representation was accomplished with post hoc assessment of the MD of the clusters, and visual inspection of the patterns of defect within the observed clusters.[24–28] We aimed to diminish the effects of human bias by designing a process for detecting change over time along mathematically determined glaucomatous patterns obtained by unsupervised learning techniques without human intervention, and we aimed to improve effectiveness by eliminating noncontributing data and concentrating on the data that are changing.[29–31]

Our initially described method, the variational Bayesian independent component analysis mixture-model (VIM), is a semiautomatic, unsupervised machine learning approach that has been shown to cluster VFs in a meaningful way and to generate nearly maximally independent, clinically recognizable patterns of glaucomatous VF defects.[24,27,32] The Gaussian mixture-model with expectation maximization (GEM) produces a similar output, but learns 50 times faster and is a fully automated unsupervised learning approach.[28,31] The application of progression of patterns (POP) to VIM and GEM make the progression detectors VIM-POP and GEM-POP.[29–31]

Other approaches also exist for the specific task of progression detection, including point-wise linear regression (PLR)[33] that evaluates change at each individual test location over the entire follow-up duration based on a fixed number of changing test locations; permutation analysis of PLR (PoPLR),[34] an individualized analysis that uses a P value combination function and permutation analysis to detect glaucomatous change; combined binomial tests with PLR,[35] and methods based on variational Bayesian analysis.[36]

In this paper, we assess the clinical effectiveness of VIM-POP and GEM-POP methods for detecting glaucomatous progression along VF defect patterns. We also compare VIM-POP and GEM-POP with other methods for detecting progression. Finally, we validate the specificity of all methods using independent datasets.

Methods

In the current study, we use the same data set to compare GEM-POP and VIM-POP with other progression-detection methods. We first evaluate the ability of GEM to cluster healthy and glaucomatous VFs and to generate patterns of visual field defects within each cluster.[28,31] We then compare the clustering performance of GEM with VIM, based on specificity and sensitivity for clustering VFs as healthy and glaucomatous. Next, we detect glaucomatous progression in study eyes based on significant change of longitudinal VF measurements (exams) along the previously generated GEM and VIM defect patterns, using POP. Finally, we compare the accuracy of GEM-POP and VIM-POP, to PLR,[33] PoPLR,[34] and linear regression of MD and VFI, with detect progression in VFs from known progressing eyes.

Participant Selection and Testing

Study participants were selected from two prospective longitudinal studies designed to evaluate visual function and optic nerve structure in glaucoma: The University of California at San Diego (UC San Diego; San Diego, CA)-based Diagnostic Innovations in Glaucoma Study (DIGS) and the UC San Diego–based African Descent and Glaucoma Evaluation Study (ADAGES). ADAGES is a three-site collaborative study among the Hamilton Glaucoma Center of the Department of Ophthalmology at UC San Diego, the New York Eye and Ear Infirmary (NYEEI; New York, NY), and the Department of Ophthalmology, University of Alabama, Birmingham

Table 1. Demographic Information of Study Eyes in the Classification Study Group for Assessing Diagnostic Classification Accuracy

Parameter	Abnormal Visual Field	Normal Visual Field	*P* Value
Number of eyes	859	1117	-
Number of subjects	617	699	-
Age at exam in years (SD)	58.0 (13.9)	46.6 (14.3)	< 0.01
Percent male	42.3	33.6	< 0.01
MD in dB (SD)	−4.15 (4.78)	−0.43 (1.25)	< 0.01
PSD in dB (SD)	4.32 (3.10)	1.50 (0.24)	< 0.01

(UAB; Birmingham, AL). Both studies follow identical protocols and the methodological details have been described previously.[37] The institutional review boards of UC San Diego, NYEEI, and UAB approved all DIGS and ADAGES methods. All methods adhered to the tenets of the Declaration of Helsinki and to the Health Insurance Portability and Accountability Act. DIGS and ADAGES are registered as cohort clinical trials with www.clinicaltrial.gov (NCT00221897 and NCT00221923, respectively; September 14, 2005).

Participants underwent a comprehensive ophthalmologic examination, including medical history review, best-corrected visual acuity, slit-lamp biomicroscopy, IOP measurement with Goldmann applanation tonometry, gonioscopy, dilated examination of fundus by indirect ophthalmoscopy, stereoscopic optic disc photography, and SAP. SAP testing was performed using the 24-2 Swedish Interactive Thresholding Algorithm (SITA). Only reliable tests ($\leq$20% fixation losses, $\leq$33% false-negative results, and $\leq$15% false-positive results) were included. Trained reviewers from the UC San Diego-based Visual Field Assessment Center (VisFACt) ensured that all VF tests studied were free of apparent artifacts (e.g., lid or rim artifacts, signs of fatigue).

For analysis of unsupervised learning techniques (described below), POAG patients were defined as those with repeatable (two consecutive) abnormal SAP VFs in one or both eyes. A designation of abnormal VF required a pattern standard deviation (PSD) less than or equal to 0.05 or a glaucoma hemifield test (GHT) outside of normal limits.[38] In this study, participants with normal VFs were defined as those with no evidence of repeatable abnormal VFs (as defined above) in each eye.

Because the overall goals of the current study are to assess the clinical effectiveness of VIM and GEM methods in generating optimal visual defect patterns and detecting glaucomatous progression using these defect patterns, four independent groups of study eyes were used: (1) the Classification Study Group, (2) the Stability Definition Group, (3) the Progression Study Group, and the (4) Validation Group.

Classification Study Group

The Classification Study Group included 1976 eyes (of 1316 study participants). Abnormal SAP VFs were found in 859 eyes (of 617 study participants) and 1117 eyes (of 699 study participants) had SAP VFs within normal limits. Study eyes in this group were classified as those with VF defects and those without VF defects, regardless of their optic disc assessment. We used the visual field status alone as an indicator of glaucoma because this study group was used primarily to generate the optimal VIM and GEM defect patterns of the VF, and not to detect glaucoma. If classification groups had been defined based on the presence or absence of apparent glaucomatous optic neuropathy (GON), it is likely that a significant number of GON eyes would have provided VFs within normal limits. Table 1 shows the demographic information of participants in the abnormal and normal groups and their mean MD and PSD values.

Stability Definition Group

The Stability Definition Group included 84 eyes from 45 study participants with repeatable (i.e., $\geq$ 2 consecutive) glaucomatous SAP defects at baseline. Each study eye was tested once a week with one baseline exam and a mean follow-up of 4.8 exams over a mean follow-up duration of 4.3 weeks. A total of 403 SAP VF measurements were collected. We considered the VF measurements in this group as stable glaucoma because disease-related progression in adequately treated glaucoma eyes generally occurs over years and not weeks. Any changes during this short follow-up duration likely would be due to variability in the function of diseased ganglion cells or in the attentiveness of the patient and not due to disease-related progression. Table 2 shows the demographic information of the participants in the Stability Definition Group.

Table 2. Demographic Information of Study Eyes in the Stability Definition Group and Progression Study Group for Assessing the Accuracy of Detecting Glaucomatous Progression

Parameter	Stable	Progressed by Photo	*P* Value
Number of eyes	84	83	-
Number of subjects	45	74	-
Mean number of follow-ups (SD)	4.8 (0.8)	14 (4.8)	< 0.01
Months of follow-up (SD)	1.2 (0.2)	109.2 (26.4)	< 0.01
Age at baseline in years (SD)	70.8 (9.4)	62.5 (12.4)	< 0.01
Percent male	53.3	48.2	0.19
Baseline MD in dB (SD)	−7.3 (8.1)	−4.4 (5.8)	< 0.01
Baseline PSD in dB (SD)	6.1 (4.1)	5.0 (4.2)	0.13

Progression Study Group

The Progression Study Group included eyes with known progressive glaucomatous optic neuropathy (PGON). We defined PGON based on structural evidence of progression independent of VF measurements so as not to bias the assessment of the methods that use VF measurements to detect progression. The baseline and each follow-up photograph of the eyes in the progression study group were assessed for PGON by two expert-trained observers viewing digitized stereoscopic image pairs on a 21-inch or larger computer monitor. Progressive glaucomatous optic neuropathy was defined as a decrease in the neuroretinal rim width or the appearance of a new or enlarged retinal nerve fiber layer (RNFL) defect evident in paired stereoscopic images. Observers were masked to patient identification and all clinically relevant results. A third observer adjudicated any disagreement in assessment between the first two observers. From 74 participants, 83 eyes were identified as progressed by PGON. A total of 1161 SAP VF measurement visits were available from this group. The mean number of follow-up visits was 14.0 (SD = 4.8), and the mean follow-up duration was 9.1 (SD = 2.2) years, yielding a mean interval between exams of 7.8 months. Demographic information for this group is presented in Table 2.

Validation Groups

In order to validate the 95% limits of stability used for all progression detection methods described, we investigated the specificity of each method in two external datasets. The first datatset was analogous to our Stability Definition Group and was composed of 115 glaucoma eyes (consecutive abnormal VFs prior to baseline with GON on ophthalmic examination) tested five times over 4 weeks and provided by Douglas Anderson, MD from Bascom Palmer Eye Institute, University of Miami Health Systems (Miami, FL; the institutional review board of Bascom Palmer Eye Institute approved all testing and methods adhered to the tenets of the Declaration of Helsinki and to the Health Insurance Portability and Accountability Act). The second dataset was composed of 54 healthy eyes from UC San Diego DIGS study tested every 6 months over 2 years (approximately 5 visits). Research protocol was part of DIGS and therefore, adheres to all ethics and regulatory guidelines. All VFs used for validation were reliable, as described previously.

Generating Vim and Gem Clusters and Axes Used for Detecting Progression

VIM and GEM methods assigned each of the 1976 VF exams in the Classification Study Group to several clusters and further generated several VF defect patterns (axes) within these clusters. Each cluster centroid contains information about disease severity and each axis contains information about the pattern (shape) of VF defect that can be investigated further for progression by looking at points projected along that axis. During classification, no prior knowledge was provided to VIM or GEM as to whether the input VFs were abnormal or normal (i.e., the learning was unsupervised).

For VIM, all of the visual fields in each cluster were decomposed into different axes using independent component analysis (ICA).[39] Independence of axes was forced within each cluster, not between different clusters. The generated visual field at ±2 SD from the cluster mean on each axis, and the VFs associated with each cluster, characterized the patterns of visual defect. To avoid working with a large number of axes, only axes with significant contributions (described later) were retained in each cluster. VIM-POP was equipped with a sliding window because it is expected that glaucoma-associated change in visual function and structural imaging is nonlinear in some eyes. For example, a 5-visit spurt of

progression in a trend of 10 visits might be missed by a 10-visit linear regression while regression in a sliding 5-visit window along the 10-visit range would be more likely to capture the progression spurt.

For GEM, we investigated a modular use of ICA to generate defect patterns (axes) within each of the two classes of disease severity (within each cluster). The visual field patterns were represented as axes through each cluster centroid and each cluster axis had the property of representing the visual field loss patterns from mild to advanced disease. Like VIM-POP, GEM-POP also was equipped with a nonlinear progression detection algorithm, which uses a sliding-window of size 5 (approximately 3 years of follow-up).

For cross-sectional VF assessment, clinicians typically rely on the total deviation (TD) or pattern deviation (PD) plots supplied by the Humphrey Field Analyzer (HFA; Carl Zeiss Meditec, Inc., Dublin, CA) StatPac analysis. Because clinicians are accustomed to the TD display, we used simulated TD plots to display the patterns of VF defects identified by VIM and GEM, relative to healthy eyes. The simulated TD plot is a vector obtained by subtracting absolute sensitivities at the centroid of the healthy cluster from individual VF points in absolute sensitivity space. Total deviation plots were generated at ±2 SD along each of the axes and displayed. The simulated TD plots were displayed in color to help visualization, with green denoting positive values (more sensitive than normal) and red denoting negative values (less sensitive than normal). The ±26 dB simulated VF sensitivities were displayed in equal steps of color from red to green, with −26 as pure red and +26 as pure green, representing the entire range of normalized measures of deviation from normal in each location. With this representation, it is easier to visualize increasing VF defect severity as deviations from the healthy mean along each glaucoma axis.

Variational Bayesian Independent Components Analysis Mixture Model (VIM)

VIM has been described in detail in previous publications.[29,32] Briefly, VIM is a combination of multiple ICA models weighted in a probabilistic manner. This combination allows the unsupervised identification of independent clusters of data, each containing statistically independent axes of information. Clustering and axis development are done simultaneously in VIM. VIM is a semiautomatic clustering method because the user selects the model with the highest average of sensitivity and specificity among a very large number of VIM models and the optimal model is retrained to further improve its diagnostic accuracy. In the current study, the VIM training feature set had 53 features; the absolute sensitivity values from 52 of the 54 VF test points (2 blind spot points excluded) and participant age, for each of the 1976 SAP tests. VIM varied the maximum number of clusters, the maximum number of axes within each cluster, and the number of Gaussians to create 720 models. Each model was iterated 500 times, employing a different number of axes and several random seeds at initialization. It was assumed that the single best model (highest average of sensitivity and specificity for identifying abnormal and normal VFs) would provide the best environment for finding glaucomatous patterns and for detecting progression. The optimal number of axes within each cluster was chosen based on the contribution of each axis in cluster decomposition. This number was called a "knee point" and was chosen by ranking the axes in each cluster based on their lengths or magnitudes and including the number of axes with the largest magnitudes and excluding axes with smaller magnitudes.[32] On a graph of axis contribution versus number of axes, the chosen number of axes occurred at the point where the slope of the graph changed from steep to nearly horizontal; hence, the term knee point. The optimal model was retrained 500 times to determine the final best specificity and sensitivity.

Gaussian Mixture Model with Expectation Maximization (GEM)

GEM has been described in detail in a previous publication.[28] Briefly, GEM combines multivariate Gaussian components to model the VF data points and uses the expectation maximization (EM) procedure to estimate the parameters of the model, iteratively. Similar to VIM, we used absolute sensitivity at 52 SAP locations and participant age as inputs to GEM. Clusters were created by selecting the component that maximized the maximum a posteriori probability, based on the EM-estimated parameters. GEM is a probabilistic approach with a hierarchical modular framework that allows identification of clusters first, followed by identification of axes within each cluster using ICA. Unlike VIM, where clustering and axis development are done simultaneously, GEM is sequential. To select an optimal GEM model that represents glaucoma categories and VF defect patterns, we generated 600 GEM models (to be roughly comparable to VIM for assessing the computational

complexity, however we have observed that 50 models generally are sufficient for GEM), and selected the model that provided the highest average of sensitivity and specificity for discriminating between abnormal and normal VFs. The optimal number of axes within each cluster was chosen based on the previously described knee point.

Procedure for Detecting Glaucomatous Progression

Using the VIM and GEM environments composed of the VF defect patterns or axes; we first defined the limits of stability using the Stability Definition Group. Progression by VIM and GEM was identified based on the progression of each study eye along each VIM and GEM axis. Details of identifying the limits of stability and detecting progression using the VIM and GEM environments are described next. More detailed technical descriptions are available elsewhere.[28,31]

Stability Definition

Glaucoma boundary limits were detected by projecting the longitudinal sequence of VFs (each field in a 52-dimensional space) of each stable eye (i.e., eye from the Stability Definition Group) onto each of the predefined VIM or GEM glaucoma axes. Because eyes in the stable group presumably showed no disease related progression, the variability in this group was used to define the limits of stability or the progression limits. The VF exams of each stable eye were randomly resampled with replacement (i.e., bootstrapped) to simulate an eye with seven visits. One thousand resampled VF series were used to determine the progression limit, or limit of stability, on each axis. The slope of each visual field sequence was determined for each window (of size 5) using least squares linear regression for each VIM and GEM axis. The 95th percentiles (one-tailed toward the direction of deterioration) for each of the multiple axes and windows (described in the Results section) for detecting glaucoma progression were calculated and compensated to accommodate an overall 95% specificity based on each axis and the sliding window (i.e., specificity along each axis was adjusted upward to result in a total 95% specificity for all axes and windows combined for both VIM and for GEM). For each axis, the rate of progression of an eye along an axis was normalized to the adjusted 95th percentile threshold value for that axis. An eye with a normalized rate of progression (of VF threshold) greater than or equal to 1 was classified as progressed, and an eye with a normalized rate of progression less than 1 was classified as nonprogressed. This method of defining progression along each VIM and GEM axis is called POP.

Glaucoma Progression Detection

Each visit of a longitudinal VF series of each study eye in the Glaucoma Study Group was projected onto each of the VIM and GEM axes. The mean rate of progression (slope) of the longitudinal sequence of windows along each axis was determined for each eye using the least squares linear regression model. A study eye was classified as progressed when the normalized rate of progression on any one of the axes equaled or exceeded 1 (i.e., when the regression line fell below the lower limit of stability); otherwise, the eye was classified as having no evidence of progression. The axis demonstrating the most change was identified as the progressing axis. The pattern of that axis was then considered as the glaucoma pattern of progression.

The 95th percentile progression limits for MD and VFI were computed separately using the Stability Definition Group (there was no need to adjust the progression limits because MD and VFI each have only one axis). Then, for each eye in the Progression Study Group, the MD and VFI slopes were computed over the follow-up period. If the slope of MD or VFI exceeded the corresponding progression limit, the eye was classified as progressed; otherwise the eye was classified as nonprogressed.

Point-wise linear regression was conducted using the methods described by Fitzke et al.[33] To detect progression for PLR, for each of 52 VF points, two different parameters; the slope and the significance of the slope (P value) were considered. Progression of the VF was defined based on one to three deteriorating points with significant P value (smaller than 0.01) of the slope exceeding the specified threshold, allowing six different criteria for VF progression (two different slopes and three possible locations). We applied these six criteria to our stability definition group to compute specificity and to our progression study group to compute sensitivity to identify sensitivity/specificity trade-offs at several discrete points along the receiver operating characteristic (ROC) curve similar to O'Leary et al.[34] We also implemented PoPLR according to method described by O'Leary et al.[34] We combined the significance values across the VF and compared it with the null distribution of P values of all combinations of the VF sequences obtained by

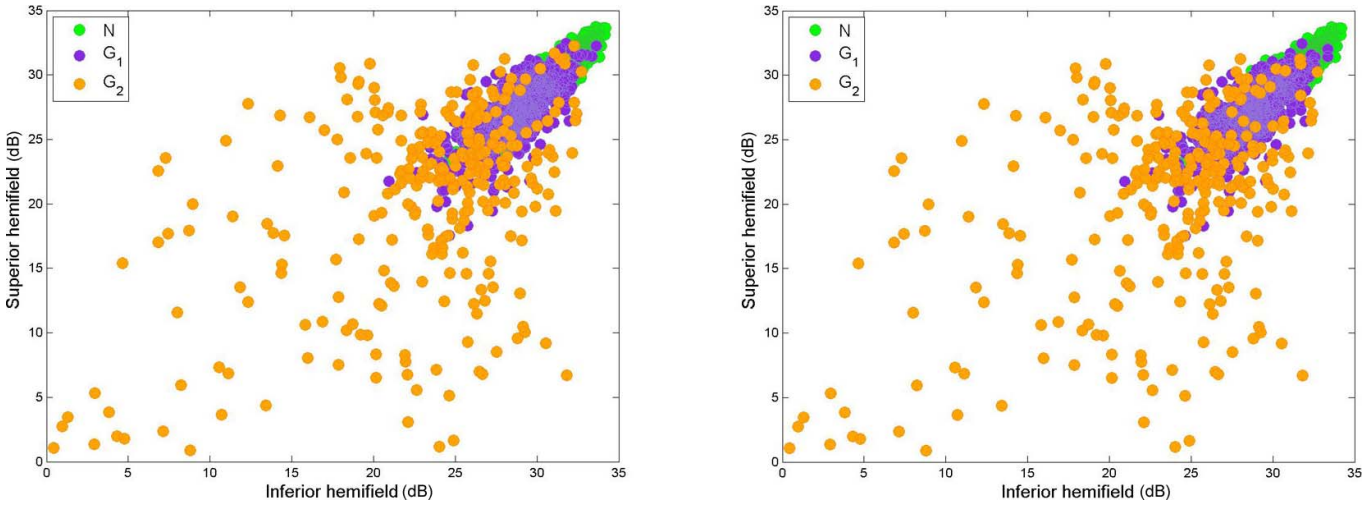

Figure 1. Scatter plot of mean threshold of superior hemifield versus mean threshold of inferior hemifield to show 2-dimensional margin displayin each VIM (left panel) and GEM (right panel) cluster. Each cluster is color coded for identification according to the Figure legend. *N* indicates the primarily normal cluster and G_1 and G_2 indicate the early to moderate and moderate to advanced glaucoma clusters, respectively.

permutation of visits. We considered progression if the combined significance value of actual visits was greater than 95th percentile of the null distribution of all permuted visits and computed the full ROC curve.

Full and partial (0.85–1.0 specificity) ROC curves were generated for VIM-POP, GEM-POP, PoPLR, linear regression of MD, and linear regression of VFI, for comparison.

Results

For both VIM and GEM, environments were automatically built that separated SAP VFs into normal and abnormal clusters and identified patterns of field defects (axes) within each cluster. Subsequently, these environments were used for progression analysis (i.e., POP).

SAP MD ± SD was −0.43 ± 1.25 dB and −4.14 ± 4.78 dB for normal and abnormal VFs, respectively, within the Classification Study Group. For both VIM and GEM, three clusters were identified: clusters N (representing normal VFs), G_1 (representing early to moderate glaucoma, based on post hoc assessment of MD), and G_2 (moderate to advanced glaucoma, based on post hoc assessment of MD). The best VIM model was composed of nine axes: two axes for each of the first two clusters (N and G_1) and five axes for the third cluster (G_2). Similarly, the best GEM model was composed of nine axes: two axes for each of the first two clusters and five axes for the third cluster.

VIM and GEM Degrees of Disease Severity (Clusters)

For VIM, cluster N contained primarily VFs from healthy participants (normal VFs, 1084 of 1117 eyes; specificity = 97.0%). Clusters G_1 and G_2 combined contained primarily VFs from participants with glaucoma (abnormal VFs, 799 of 859, sensitivity = 93.0%). Cluster G_1 contained 503 eyes with abnormal VFs and 32 eyes with normal VFs. Cluster G_2 contained 296 eyes with abnormal VFs and one eye with a normal VF.

For GEM, cluster N contained primarily normal VFs (1048 of 1117 eyes; specificity = 93.8%). Clusters G_1, and G_2 combined contained primarily VFs from participants with glaucoma (772 out of 859, sensitivity = 89.9%). Cluster G_1 contained 478 eyes with abnormal VFs and 69 eyes with normal VFs. Cluster G_2 contained 294 eyes with abnormal VFs and no eyes with normal VFs. Figure 1 shows scatter plots of the mean threshold values in the superior hemifield plot versus the inferior hemifield plot, to demonstrate cluster variability and Figure 2 shows MD versus PSD of all eyes assigned to each of the VIM and GEM clusters, to demonstrate clustering by VF defect severity.

VIM Patterns of Glaucomatous Defect (Axes)

The centroid of a VIM cluster represents the mean defect severity in that cluster. The mean of cluster G_2

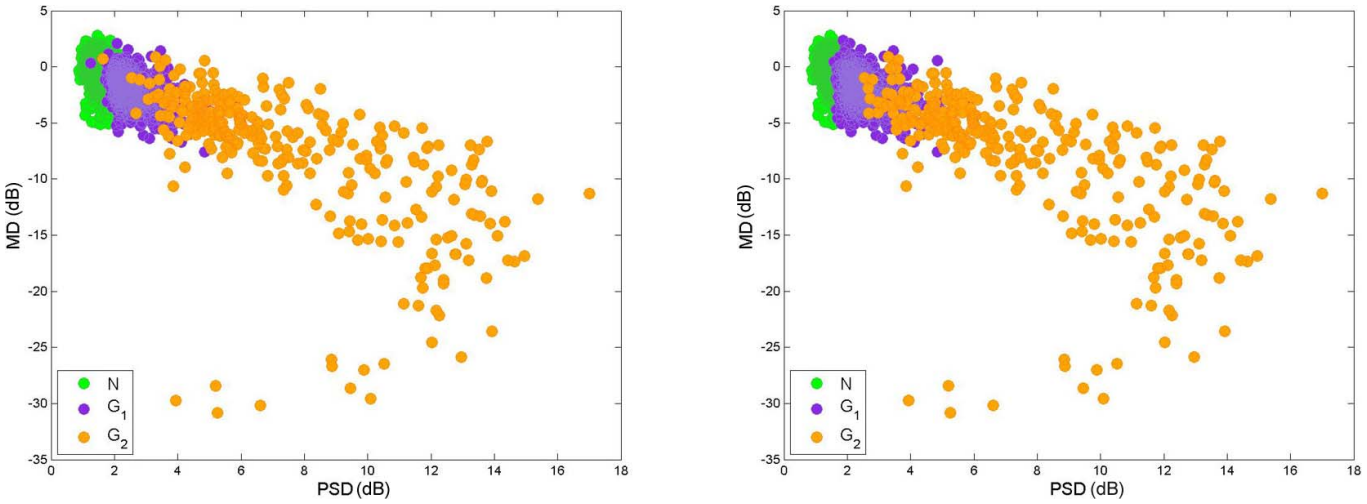

Figure 2. Scatter plot of MD versus PSD in each VIM (left panel) and GEM (right panel) cluster. Each cluster is color coded for identification according to the figure legend.

was farther from the mean of cluster N than the mean of cluster G_1, indicating that the overall defect was more severe in G_2. Within each cluster, axes represented the pattern (shape) of each VF defect. VIM efficiently identified the magnitude of the various patterns of defect present in each study cluster by projecting each VF along each axis. If a point moved along any axis away from the cluster mean, the direction of motion would be positive if the distance (vector) of the point from the normal mean increased with movement, and the direction would be negative if the vector from the normal mean to the point decreased with movement.

To discern whether there was progression in an eye, each visit for that eye was projected onto an axis. Change from the cluster centroid in a positive direction (toward +2 SD) along a given axis generally resulted in an increase in depth of defect or an increase in area of defect. Change in the negative direction (toward −2 SD) generally resulted in a decrease in depth of defect or a decrease in area of defect.

Figure 3 shows the various VIM defect patterns (simulated TD plots) that were generated at ±2 SD from the respective cluster centroids on each axis in clusters N, G_1, and G_2. Similar to the total deviation plot in SAP, these patterns represent the degree of defect severity and deviation from normal. The two defect patterns of cluster N shown are normal, without glaucomatous damage. Cluster G_1 includes eyes with early to moderate glaucomatous damage (average MD = −2.15, SD = 1.52 dB); consequently, the patterns shown illustrate primarily mild defects. Axis 1 of G_1 (at +2 SD) displays early superior and inferior nasal steps and arcuate reductions of sensitivity. Axis 2 of G_1 (at +2 SD) suggests an early superior arcuate defect. Cluster G_2 includes eyes with more moderate to advanced glaucomatous damage (average MD = −7.98, SD = 6.27 dB); hence, the axis patterns shown in Figure 3c are more advanced. Axis 1 of G_2 (at +2 SD) represents a diffuse depression with increased depression in the superior hemifield. Axis 2 of G_2 (at +2 SD) shows early diffuse reduction of sensitivity with an inferior hemifield defect. Axis 3 of G_2 (at +2 SD) represents a peripheral defect with increased defect at the nasal steps and the superior arcuate zone. Axis 4 of G_2 (at +2 SD) presents a nasal step defect with inferior nasal exaggeration and axis 5 of G_2 (at +2 SD) suggests a central-to-nasal defect, with inferior weighting.

GEM Patterns of Glaucomatous Defect (Axes)

Similar to VIM, points at ±2 SD on GEM-defined axes in clusters N, G_1, and G_2 were displayed as TD plots. These generated TD plots at points on the axes within each cluster represented distinct VF defect patterns.

Figure 4 shows the various GEM defect patterns (axes) at ±2 SD from the cluster centroids of axes in clusters N, G_1, and G_2. The defect patterns of cluster N are normal, without glaucomatous damage. Cluster G_1 includes eyes with early to moderate glaucomatous damage (average MD = −2.19, SD = 1.55 dB). Axis 1 of G_1 (at +2 SD) represents a mild peripheral superior arcuate depression, and axis 2 of G_1 (at +2 SD) represents early superior arcuate and nasal step defects. Cluster G_2 includes eyes with moderate to

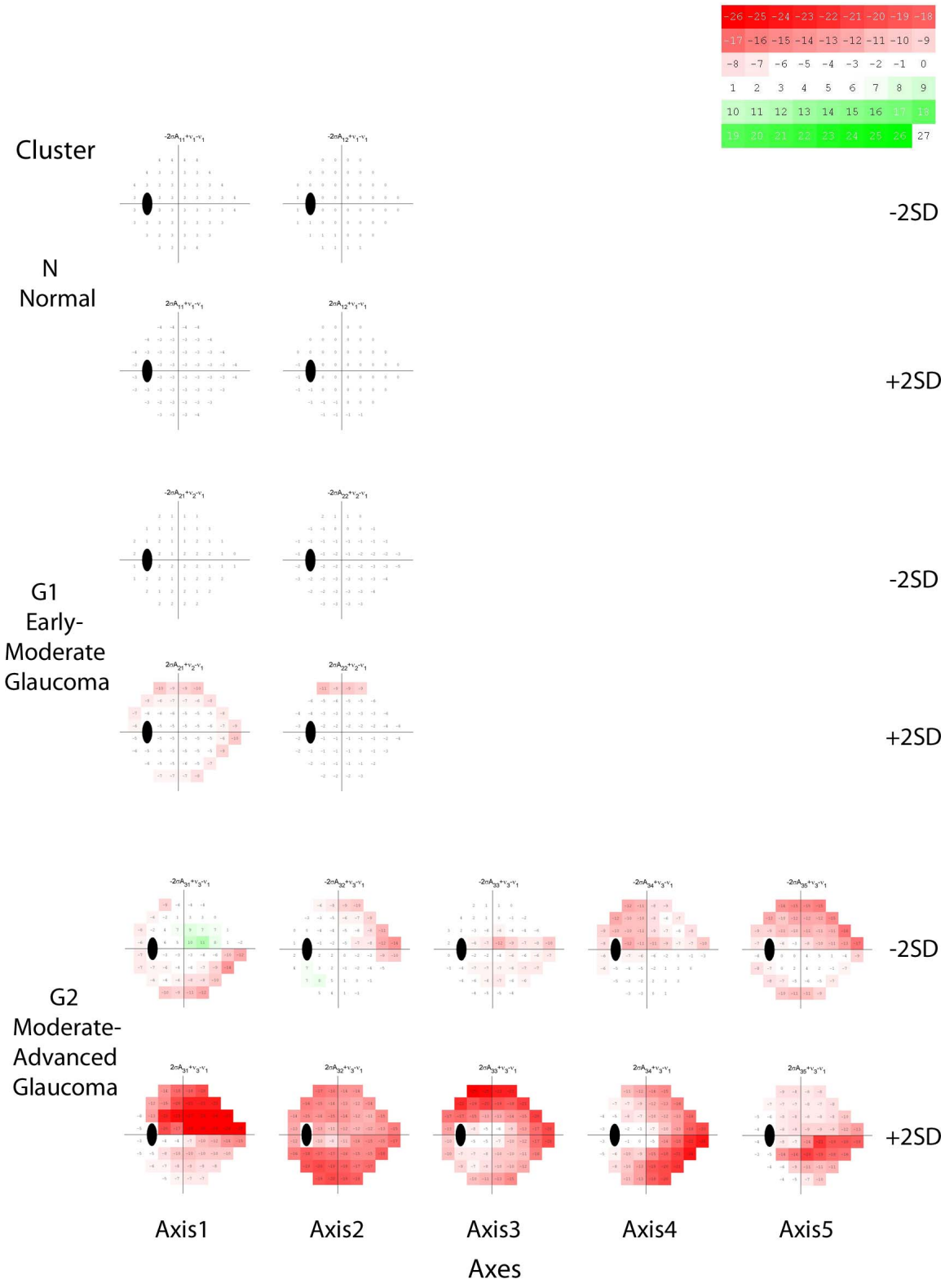

Figure 3. Visual field defect patterns that constituted the VIM environment. Cluster N shows patterns at −2 SD and +2 SD in the normal cluster. Cluster G1 shows defect patterns at −2 SD and +2 SD in the early to moderate glaucoma cluster. Cluster G2 shows defect patterns at −2 SD and +2 SD in the moderate to advanced glaucoma cluster.

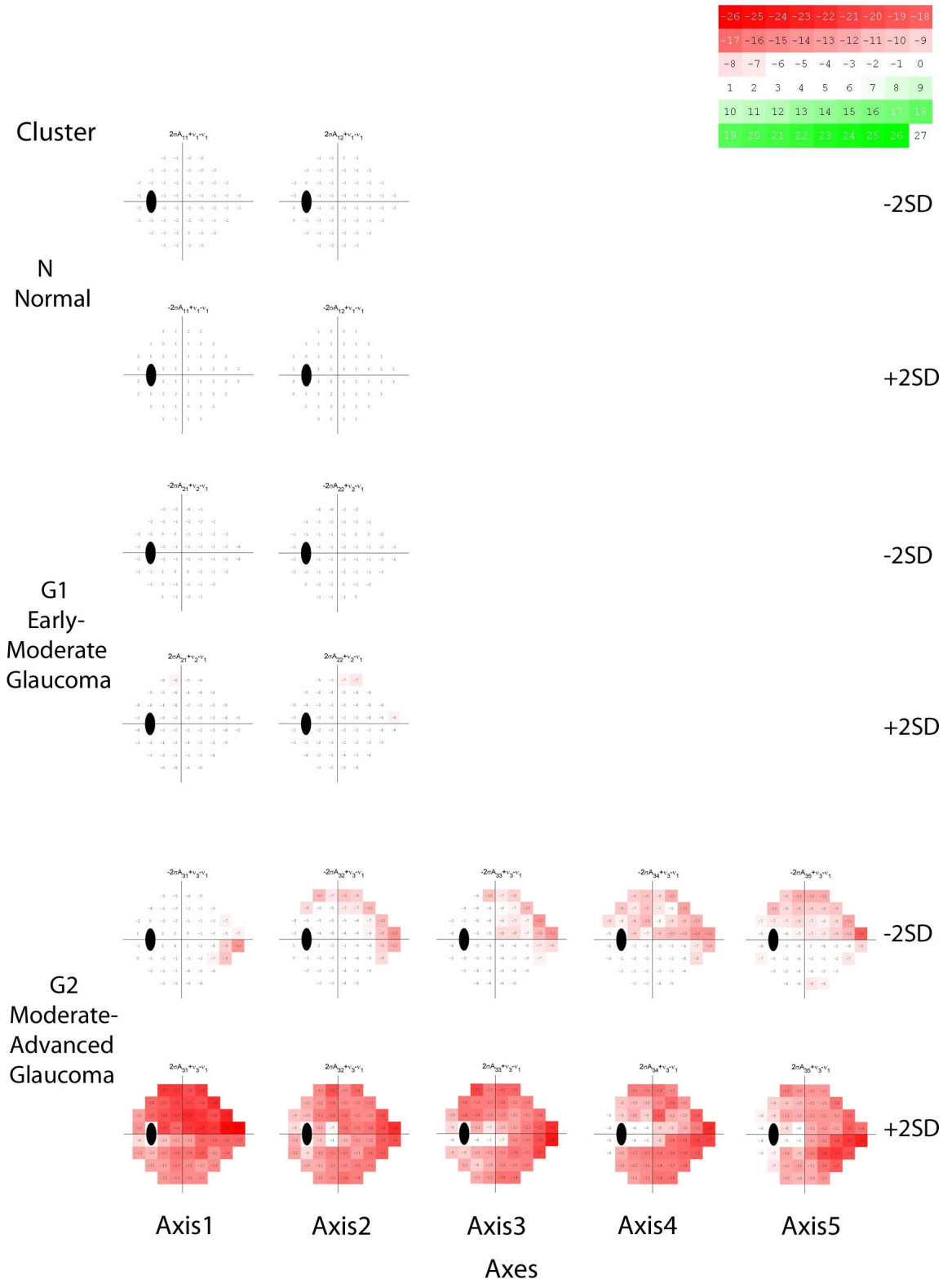

Figure 4. Visual field patterns that constituted the GEM environment. Cluster N shows patterns at −2 SD and +2 SD in the normal cluster. Cluster G1 shows defect patterns at −2 SD and +2 SD in the early to moderate glaucoma cluster. Cluster G2 shows defect patterns at −2 SD and +2 SD in the moderate to advanced glaucoma cluster.

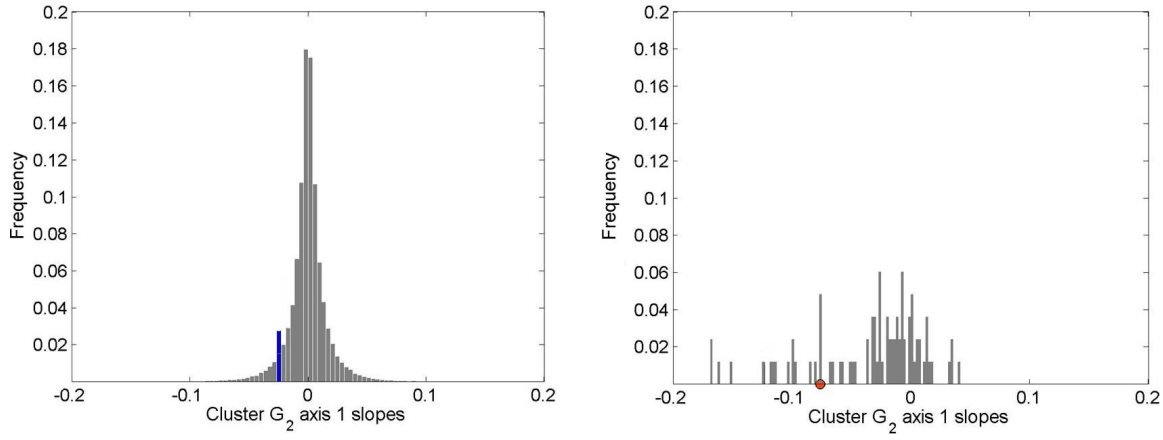

Figure 5. Limits of stability in the VIM environment. (Left panel) Histogram distribution of the slopes of the projected VFs (estimated using linear regression) of all eyes in the Stability Definition Group along the first axis in the G_2 cluster. The *blue line* indicates the left tail 95th percentile, or the stability limit. (Right panel) Actual observed rates of the projection coefficients of all Progression Study Group eyes along the first axis in the G_2 cluster. The *red circle* indicates the stability limit (obtained from the left panel with overall specificity shifted to adjust to 95%, to control for multiple axes). All eyes that exceed this limit (or fell to the left of the *red circle*) were classified as progressed along this axis.

advanced glaucomatous damage (average MD = −8.07, SD = 6.26 dB). Axis 1 of G_2 (at +2 SD) represents moderate diffuse reduction of sensitivity with a more pronounced superior hemifield defect and axis 2 of G_2 (at +2 SD) represents a diffuse reduction of sensitivity with increased defect in the nasal step positions. Axes 3, 4, and 5 (at +2 SD) primarily represent nasal step defects, with some variation in weighting among these three axes.

VIM and GEM Progression

SAP mean deviation (±SD) was −7.44 ± 8.20 dB and −4.41 ± 5.79 dB at baseline in the Stability Definition Group and the Progression Study Group, respectively.

The left panel of Figure 5 shows the distribution of the slopes of the projection magnitudes of all VFs in the Stability Definition Group along the first axis in G_2 cluster of VIM. The left tail of the 95th percentile limit is shown as a blue line that indicates the stability limit for this axis toward progression. The right panel of Figure 5 shows the distribution of the slopes of projected points on each axis from all eyes from the Progression Study Group on the first axis of the G_2 cluster of VIM. The red circle shows the stability limit computed from the distribution of the slopes of the Stability Definition Group on the first axis of cluster G_2 (Figure 5, left, blue line). For a test eye, if its slope (estimated by least square regression of projected VFs) exceeded this limit, the eye was classified as a progressed along this axis.

The left panel of Figure 6 shows the distribution of the slopes of the projection magnitudes of all VFs in the Stability Definition Group along the first axis in the G_2 cluster of GEM. The right panel of Figure 6 shows the distribution of the slopes of all Progression Study Group eyes on the first axis of the G_2 cluster of GEM. The red circle shows the stability limit that was computed from the distribution of the slopes of the Stability Definition Group on the first axis of cluster G_2 (Figure 6, left, blue line). As above, if the slope of a test eye exceeded this limit of stability, the eye was classified as a progressing eye along this axis.

Figure 7 shows progression detection using GEM-POP in two example eyes. The orange circles represent the magnitude of the defect pattern given by the first axis of cluster G_2 present in the 52-dimensional VF space. The blue line indicates the slope (linear regression of the orange circles). The gray line indicates the 95% progression limit for the slopes of the first axis of cluster G_2. If the linear model approximating the slope falls below the gray line (progression zone), then the eye is classified as progressed; otherwise, the eye is classified as non-progressed. Therefore, GEM-POP is detecting progression in the study eye in Figure 7 (left) and is detecting no evidence of progression in the study eye in Figure 7 (right) along the first axis of cluster G_2. In addition to event-related information, Figure 7 (left) provides information about the rate of progression in the example eye (blue circles).

Figure 8 shows ROC curve areas for VIM-POP, GEM-POP, PoPLR, PLR, linear regression of MD,

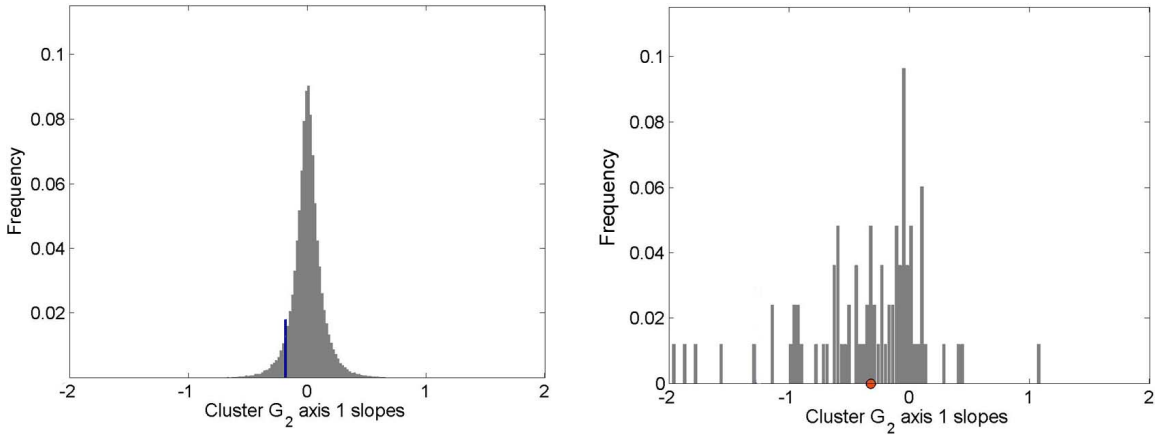

Figure 6. Limits of stability in the GEM environment. (Left panel) Histogram distribution of the slopes of the projected VFs (estimated using linear regression) of all eyes in the Stability Definition Group along the first axis in the G_2 cluster. The *blue line* indicates the left tail 95th percentile, or the stability limit. (Right panel) Actual observed rates of the projection coefficients of all Progression Study Group eyes along the first axis in the G_2 cluster. The *red circle* indicates the stability limit (obtained from the left panel with overall specificity shifted to adjust to 95%, to control for multiple axes). All eyes that exceed this limit (or fell to the *left* of the *red circle*) were classified as progressed along this axis.

and linear regression of VFI. For PLR, we applied six criteria to our stability definition group to compute specificity and to our progression study group to compute sensitivity, resulting in discrete points along the ROC curve. Therefore, the area under the ROC curve was not computed for PLR. The areas under the ROC curves for VIM-POP, GEM-POP, PoPLR, linear regression of MD, and linear regression of VFI are 0.82, 0.86, 0.81, 0.69, and 0.76, respectively. It can be observed that the ROC curve area of GEM-POP is similar to that of VIM-POP and is significantly greater than the ROC curve areas for PoPLR, linear regression of MD, and linear regression of VFI. Figure 9 shows the partial ROC curves for VIM-POP, GEM-POP, PoPLR, linear regression of MD, and linear regression at high specificities for better visualization. Table 3 shows the statistical difference among all methods.

Validation of Specificity of All Methods Using External Independent Datasets

Table 4 shows the specificities using all limits of stability (progression limits) in two validation groups (stable glaucoma and longitudinal healthy). These

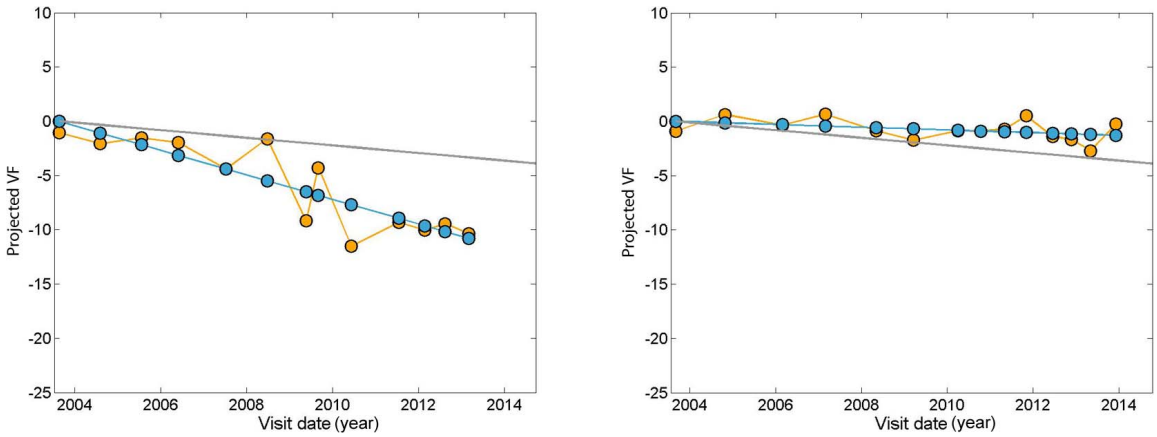

Figure 7. Progression detection by GEM in two example eyes. The *gray line* indicates the 95th percentile limit of stability for the rate of change of the projection coefficient parameter. The *orange circles* represent the actual projected VF values on the first axis of cluster G_2, and the *blue circles* are the projection coefficients estimated by a linear regression line approximating the projected visual field values on the first axis of cluster G_2. GEM is detecting progression along the first axis in the study eye in Figure 7 (left) and is detecting no progression along the first axis in the study eye in Figure 7 (right).

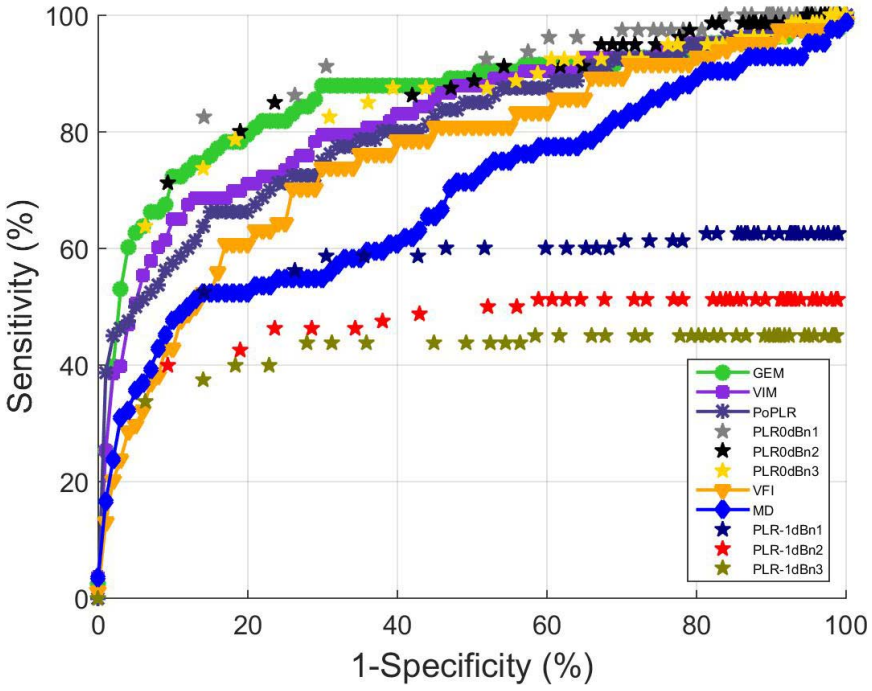

Figure 8. Full ROC curves for VIM-POP, GEM-POP, PLR, PoPLR, linear regression of MD, and linear regression VFI for detecting progressive glaucomatous optic neuropathy.

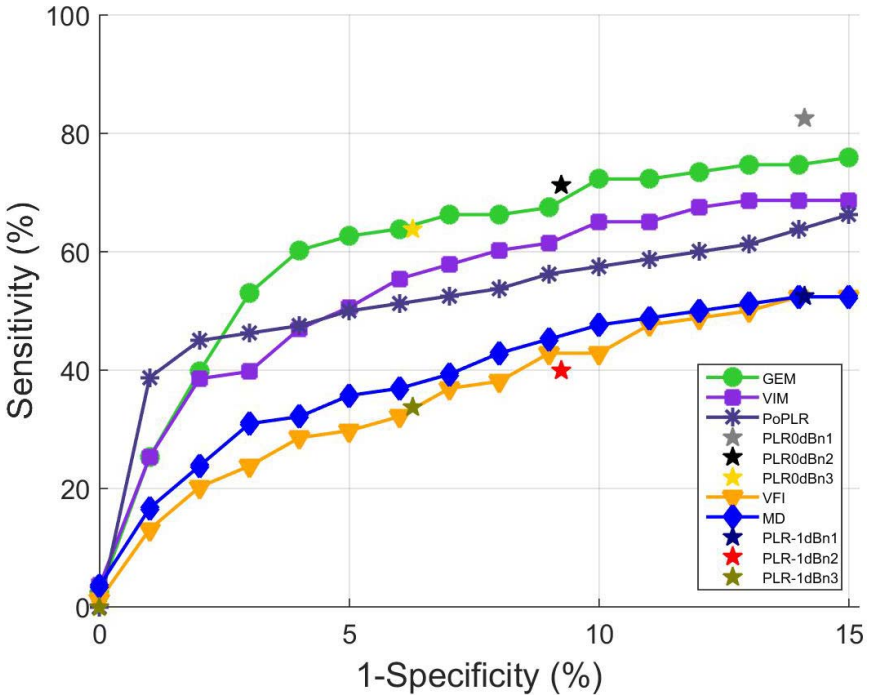

Figure 9. Partial ROC curves for VIM-POP, GEM-POP, PLR, PoPLR, linear regression of MD, and linear regression VFI for detecting progressive glaucomatous optic neuropathy.

Table 3. Statistical Significance in Difference Between Area Under the ROC Curves Among VIM-POP, GEM-POP, PoPLR, MD, and VFI

Method	GEM-POP	PoPLR	MD	VFI
VIM-POP	0.12	0.54	< 0.001	0.02
GEM-POP		0.03	< 0.001	< 0.001
PoPLR			< 0.001	0.09
MD				0.05

results indicate that the limits of stability used were applicable when applied to independent stable data sets, although specificities for one version of PLR and for linear regression of MD were somewhat low, suggesting dataset-dependence (i.e., reduced generalizability).

Discussion

VIM and GEM are novel unsupervised clustering approaches that identify intuitive (i.e., recognizable) patterns (axes) of VF defect that create an environment usable for detecting glaucomatous progression. The clusters identified using VIM and GEM in the current study were composed primarily of normal, early to moderate glaucoma, or moderate to advanced glaucoma VFs, based on MD. The diagnostic accuracy (average of sensitivity and specificity) of VIM to cluster VFs as abnormal and normal was not statistically different from that of GEM (95% and 92%, respectively). The diagnostic accuracy of VIM and GEM methods were similar with and without age as an input to the clustering step indicating that age had little effect (see also Bowd et al.,[24] for a similar VIM result using frequency doubling technology perimetry VFs). With VIM, the clusters and axes that constituted the VIM environment were determined simultaneously. In contrast, GEM used a modular approach to first generate the clusters of disease severity followed by identification of axes or defect patterns.

The area under the ROC curve of GEM-POP was significantly greater (for detecting PGON eyes) than areas under the curve of PoPLR (P value = 0.03), linear regression of MD (P value <0.001), and VFI (P value <0.001). The area under the ROC curve of GEM-POP was similar to the area under the curve of VIM-POP (P value = 0.12). Progression of patterns by VIM and GEM is based on progression along any one of seven (in the current environment) axes, whereas progression by linear regression of MD and VFI is based on a single metric, indicating that detecting localized change in defect patterns likely is a more sensitive technique than detecting global change. This superiority is expected, because in VIM-POP and GEM-POP, uncontributing VF locations are ignored; whereas for change in global indices the noncontributing VF locations are included. In fact, all progression detection methods that relied on local analysis (VIM-POP, GEM-POP, PLR [based on best performing parameter sets], and PoPLR) outperformed methods that relied on global analysis. It should be noted that PLR sensitivities can be significantly changed based on the selection of particular parameters (e.g., varying the required number of deteriorating locations or varying the slope).

From a machine learning perspective, both VIM and GEM first identified the hidden structures

Table 4. Validation of Specificity of all Methods Using Validation Groups

Method	Specificity (95% CI) on Miami Stable Dataset	Specificity (95% CI) on DIGS Normal Dataset
VIM-POP	97.5 (94.1, 100)	98.1 (93.6, 100.0)
GEM-POP	96.3 (92.2, 100.0)	96.3 (90.3, 100.0)
PoPLR	99.1 (96.7, 100.0)	96.3 (90.3, 100.0)
PLR (P value < 0.01; deterioration ≤ −1 dB/y, and at least 1 deteriorated point)	77.0 (69.0, 86.0)	72.0 (59.0, 85.0)
PLR (P value < 0.01; deterioration rate ≤ −1 dB/y, and at least 2 deteriorated points)	95.0 (91.0, 100.0)	89.0 (80.0, 98.0)
PLR (P value < 0.01; deterioration rate ≤ −1 dB/y, and at least 3 deteriorated points)	98.0 (95.0, 100.0)	96.0 (90.0, 100.0)
LR of MD	88.7 (82.2, 95.2)	81.5 (70.2, 92.8)
LR of VFI	90.6 (84.6, 96.6)	96.3 (90.3, 100.0)

(glaucoma defect patterns) and then created their respective environments for detecting progression. By extracting clinically meaningful patterns of VF defects in an unsupervised manner and studying their progression, VIM and GEM provide an optimal environment to detect progression. Because the extracted patterns are both visible to the health care provider and have shapes that are familiar and understood, these progression detection classifiers are not black boxes. Rather they provide the clinician with an understanding regarding change over time of specific patterns of defect. The modular approach developed in GEM has several advantages over the simultaneous convergence approach of VIM. Clustering of disease severity is a grouping process that is only weakly related to the discovery of independent axes or defect patterns. With a modular design, GEM allows the use of several classes of clustering algorithms and several classes of axes discovery approaches available for machine learning, by separating the clustering and axes discovery processes. As an example, the Gaussian mixture model used in GEM can be replaced with a simple k-means algorithm for clustering. In place of ICA used in GEM for axes discovery, principal component analysis (PCA) or other axes discovery approaches can be easily applied. The simultaneously converged VIM is far less adaptable.

We used GEM with ICA in the current study to make GEM consistent with VIM, because VIM also uses ICA. Post-hoc analyses on the current data revealed that sensitivity of GEM for detecting progression in PGON eyes improved somewhat when PCA was used instead of ICA (49.4% and 45.8%, respectively; a difference of 3 PGON eyes detected). PCA-generated axes are perpendicular to each other, which allows a more accurate calculation of the progression of defect pattern along each axis. Technically, because of the orthogonality of the PCA axes, calculation of the projection coefficients along each axis or the strength of a defect present in a VF is exact, unique and the calculations are simple. However with ICA, because of the lack of orthogonality of ICA axes, coefficient calculations are not exact, not unique and are computationally more complex. It is likely that lack of perpendicular axes resulted in a slight decrease in the PGON sensitivity of GEM with ICA. We observed, post hoc, that the PCA-based GEM defect patterns were not as visually representative of clinically observable defect patterns in glaucoma as ICA-based defect patterns. In other words, the addition of a measure of dissimilarity in ICA, which yields maximally different VF patterns, may increase the recognizability of the VF patterns at the expense of slightly reducing the sensitivity of progression detection compared with PCA; whereas, the orthogonality of PCA axes may increase the sensitivity of progression detection at the expense of finding and displaying maximally different VF patterns.

Another advantage of the modular design of GEM, compared with the simultaneous convergence of VIM, is the significant reduction in computational resources and time needed for training, due to a significant reduction in GEM's computational complexity. Reduced time for generating the final classifier means that there is more time to change variables and experiment with different classifiers. Creating the GEM environment took approximately 3 hours in a quad-core machine (8 gigabytes of memory). In contrast, VIM took approximately 168 hours (7 days) on the same machine to train all the classifiers and select the one used to generate the VIM progression environment.

The modular design of GEM allows the use of various classes of clustering and axes discovery techniques, which allow the building of an optimum GEM environment tailored to a specific modality or data source (for example, optical imaging data instead of visual function measurements, frequency doubling technology perimetry instead of SAP). Identifying an efficient clustering method and an axes discovery approach to build an optimal progression environment is difficult when each experimental run takes several days to complete. Therefore, the modular design in GEM also may facilitate building or improving optimal progression environments for various modalities.

In VIM, after the clustering step, the specific axes that constitute the VIM environment are manually chosen (based on the knee-point concept) and the initial VIM environment is further retrained. Hence, the VIM procedure is semiautomatic. In contrast, the modular nature of GEM separates the clustering and axes identification steps without the need for retraining the GEM environment. Therefore, GEM is fully automatic. A potential difficulty with semiautomatic methods is the need to have a both computational expertise to develop and modify algorithms and a clear understanding of the factors involved in glaucomatous progression to choose appropriate appearing axes, from a clinical standpoint.

Regarding methodological similarities with other studies, the axes discovery step in GEM is similar to

the Proper Orthogonal Decomposition (POD) framework for detecting progression from a baseline condition, described previously using confocal scanning laser ophthalmoscopy images.[40] In GEM, one set of axes that are discovered a priori describe the general patterns of glaucoma defects. Progression in a study eye is determined based on progression along these predetermined defect patterns. Whereas in POD, a set of axes is identified for each eye that describes the baseline conditions of the eye (known as the baseline subspace of the eye). Progression is determined based on the deviation of follow-up measurements from the baseline subspace of the eye.

VF progression detection methods recently proposed include Permutation Analysis of PLR (PoPLR)[34] and Analysis with Non-Stationary Weibull Error Regression and Spatial Enhancement (ANSWERS).[41] We compared GEM- and VIM-based progression detection methods with PoPLR directly. However, PoPLR[34] requires a minimum of one baseline and six follow-up VF exams (to provide at least 5000 unique permutations of the VF series for building null distributions for hypothesis testing) to generate a reliable and robust outcome.[34] Because the Stability Definition Group in the current study included an average of five VFs per eye, allowing 120 permutations, we generated sequences of seven visits for each eye to fulfill the PoPLR requirements and used the newly generated simulated test–retest dataset to compute the specificity of all methods. Therefore, the comparison is valid because all methods used the same test–retest dataset. ANSWERS relies on a mixture of Weibull distributions to model variability and a Bayesian method to aggregate spatial correlation of local measurements to confirm repeatable defects in the same or adjacent locations in follow-up examinations. The addition of spatial correlations of measurements improves this method, compared with ANSWERS' precursor, without spatial enhancement. We did not compare GEM- and VIM-based progression detection methods with ANSWERS because such a comparison is beyond the scope of the current manuscript.

Both PoPLR and ANSWERS were designed specifically to detect progression in SAP VFs and have not yet been shown to be successful when applied to other glaucoma-related measurements. Our GEM-POP and VIM-POP approaches were designed to be robust; in addition to working with SAP VFs, they can be applied to frequency doubling technology, to progressive retinal nerve fiber layer thinning, and to emerging data types, such as SDOCT mapping, to uncover patterns of defect and to detect progression of these defects using POP[24,43] (Bowd et al., *IOVS* 2014;55 ARVO Abstract 3008; Yousefi et al., *IOVS* 2015;56 ARVO Abstract 4564).

In this study, we modeled the bounds of stability for detecting progression using a simulated test–retest dataset specifically for comparison of VIM-POP and GEM-POP against PoPLR; 1000 stable pseudolongitudinal series generated by bootstrap, or resampling with replacement approach of each eye resulted in 84,000 sequences to allow us generate the full ROC curve for all methods. Longer longitudinal series (e.g., series with 7 visits) provide more confident bounds of stability. The distribution, or the region of stability, provided by the bootstrap resampling approach is asymptotically exact (i.e., distributions, or the region of stability, becomes more exact as the number of stable pseudolongitudinal series increases).[43] One of the limitations of this approach is that the effects of aging, glaucoma management, and long-term measurement variability cannot be modelled in a longer pseudoseries using only five exams. Nevertheless, this limitation does not affect the comparison of progression detection performance of VIM-POP and GEM-POP, PoPLR, PLR, MD, and VFI because all of the progression detection methods used the same simulated test–retest dataset.

The patterns of VF defects reported for the GEM algorithm is dependent on the disease (VF) status and demographic characteristics of the Classification Study Group (Table 1). The Classification Study Group used currently includes study eyes (with varying disease status) from three geographically separate clinical sites. Therefore, the GEM defect patterns reported in this study should be representative of general defect patterns in the United States and possibly internationally, which implies that the threshold between stability and progression that we derived can be used for detecting progression in glaucoma clinics that include patients with various clinical characteristics (i.e., some evidence exists that the method is generalizable).

In conclusion, GEM-POP for progression detection performs significantly better than PoPLR and linear regression of VFI and MD GEM-POP performs similarly to VIM-POP. However, GEM-POP provides a less complex environment than VIM-POP, is computationally more efficient than VIM-POP and is a fully automated technique. Although GEM-POP is more complex to develop than PLR and PoPLR, once the GEM environment for detecting progression

has been established, determining whether an eye has progressed by GEM-POP is simple and fast. Finally, GEM-POP and VIM-POP are designed to be applicable to other data types besides perimetry.

Acknowledgments

Supported by grants from National Institutes of Health (NIH) EY022039, NIH EY008208, NIH EY011008, NIH EY014267, NIH EY019869, NIH EY020518, P30EY022589 an unrestricted grant from Research to Prevent Blindness (New York), Eyesight Foundation of Alabama, Corinne Graber Research Fund of the New York Glaucoma Research Institute, and participant incentive grants in the form of glaucoma medication at no cost from Alcon Laboratories, Allergan, and Pfizer.

Disclosure: **S. Yousefi**, None; **M.H. Goldbaum**, None; **M. Balasubramanian**, None; **F.A. Medeiros**, Ametek (C), Alcon Laboratories Inc, (C), Allergan Inc. (F, C), Bausch + Lomb (F), Carl Zeiss Meditec (F, C), Heideleberg Engineering GmbH (F, C), Sensimed (F); **L.M. Zangwill**, Carl Zeiss Meditec Inc. (F), Heidelberg Engineering GmbH (F, R), Optovue Inc. (F), Topcon Medical Systems Inc. (F); **R.N. Weinreb**, Alcon Laboratories Inc. (C), Allergan Inc. (C), Carl Zeiss Meditec Inc. (F, C), Heidelberg Engineering GmbH (F), Optovue Inc. (F), Topcon Medical Systems Inc. (F, C); **C.A. Girkin**, None; **J.M. Liebmann**, Alcon Laboratories Inc. (C), Allergan Inc. (C), Carl Zeiss Meditec Inc. (F), Diopsys Corp. (F, C), Heidelberg Engineering GmbH (F), Optovue Inc. (F, C), Pfizer Inc. (C), Topcon Medical Systems Inc. (F, C); **C. Bowd**, None

References

1. Weinreb RN, Aung T, Medeiros FA. The pathophysiology and treatment of glaucoma: a review. *JAMA*. 2014;311:1901–1911.
2. Weinreb RN, Khaw PT. Primary open-angle glaucoma. *Lancet*. 2004;363:1711–1720.
3. Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. *Br J Ophthalmol*. 2006;90:262–267.
4. Kingman S. Glaucoma is second leading cause of blindness globally. *Bull World Health Organ*. 2004;82:887–888.
5. Lau LI, Liu CJ, Chou JC, Hsu WM, Liu JH. Patterns of visual field defects in chronic angle-closure glaucoma with different disease severity. *Ophthalmology*. 2003;110:1890–1894.
6. Bengtsson B, Heijl A. A visual field index for calculation of glaucoma rate of progression. *AM J Ophthalmol*. 2008;145:343–353.
7. Bengtsson B, Bizios D, Heijl A. Effects of input data on the performance of a neural network in distinguishing normal and glaucomatous visual fields. *Invest Ophthalmol Vis Sci*. 2005;46:3730–3736.
8. Bizios D, Heijl A, Bengtsson B. Trained artificial neural network for glaucoma diagnosis using visual field data: a comparison with conventional algorithms. *J Glaucoma*. 2007;16:20–28.
9. Bowd C, Chan K, Zangwill LM, et al. Comparing neural networks and linear discriminant functions for glaucoma detection using confocal scanning laser ophthalmoscopy of the optic disc. *Invest Ophthalmol Vis Sci*. 2002;43:3444–3454.
10. Bowd C, Lee I, Goldbaum MH, et al. Predicting glaucomatous progression in glaucoma suspect eyes using relevance vector machine classifiers for combined structural and functional measurements. *Invest Ophthalmol Vis Sci*. 2012;53:2382–2389.
11. Bowd C, Medeiros FA, Zhang Z, et al. Relevance vector machine and support vector machine classifier analysis of scanning laser polarimetry retinal nerve fiber layer measurements. *Invest Ophthalmol Vis Sci*. 2005;46:1322–1329.
12. Bowd C, Zangwill LM, Medeiros FA, et al. Confocal scanning laser ophthalmoscopy classifiers and stereophotograph evaluation for prediction of visual field abnormalities in glaucoma-suspect eyes. *Invest Ophthalmol Vis Sci*. 2004;45:2255–2262.
13. Burgansky-Eliash Z, Wollstein G, Chu T, et al. Optical coherence tomography machine learning classifiers for glaucoma detection: a preliminary study. *Invest Ophthalmol Vis Sci*. 2005;46:4147–4152.
14. Chan K, Lee TW, Sample PA, Goldbaum MH, Weinreb RN, Sejnowski TJ. Comparison of machine learning and traditional classifiers in glaucoma diagnosis. *IEEE Trans Biomed Eng*. 2002;49:963–974.
15. Demirel S, Fortune B, Fan J, et al. Predicting progressive glaucomatous optic neuropathy using baseline standard automated perimetry data. *Invest Ophthalmol Vis Sci*. 2009;50:674–680.
16. Goldbaum MH, Sample PA, Chan K, et al. Comparing machine learning classifiers for diag-

nosing glaucoma from standard automated perimetry. *Invest Ophthalmol Vis Sci*. 2002;43:162–169.
17. Goldbaum MH, Sample PA, White H, et al. Interpretation of automated perimetry for glaucoma by neural network. *Invest Ophthalmol Vis Sci*. 1994;35:3362–3373.
18. Kelman SE, Perell HF, D'Autrechy L, Scott RJ. A neural network can differentiate glaucoma and optic neuropathy visual fields through pattern recognition. In: Mills RP, Heijl A, eds. *Perimetry Update 1990/1991*. New York: Kugler & Ghedini Publications; 1991:287–290.
19. Lietman T, Eng J, Katz J, Quigley HA. Neural networks for visual field analysis: how do they compare with other algorithms? *J Glaucoma*. 1999;8:77–80.
20. Madsen EM, Yolton RL. Demonstration of a neural network expert system for recognition of glaucomatous visual field changes. *Mil Med*. 1994;159:553–557.
21. Nagata S, Kani K, Sugiyama A. A computer assisted visual field diagnosis system using neural netowrks. In: Mills RP, Heijl A, eds. *Perimetry Update 1990/1991*. New York: Kugler & Ghedini Publications; 1991:291–295.
22. Spenceley SE, Henson DB, Bull DR. Visual field analysis using artificial neural networks. *Ophthalmic Physiol Opt*. 1994;14:239–248.
23. Wroblewski D, Francis B, Chopra V, et al. Glaucoma detection and evaluation through pattern recognition in standard automated perimetry data. *Graefe's Arch Clin Exp Ophthalmol*. 2009;247:1517–1530.
24. Bowd C, Weinreb RN, Balasubramanian M, et al. Glaucomatous patterns in Frequency Doubling Technology (FDT) perimetry data identified by unsupervised machine learning classifiers. *PLoS One*. 2014;9:e85941.
25. Goldbaum MH. Unsupervised learning with independent component analysis can identify patterns of glaucomatous visual field defects. *Trans Am Ophthalmol Soc*. 2005;103:270–280.
26. Goldbaum MH, Jang G-J, Bowd C, et al. Patterns of glaucomatous visual field loss in sita fields automatically identified using independent component analysis. *Trans Am Ophthalmol Soc*. 2009;107:136–144.
27. Sample PA, Chan K, Boden C, et al. Using unsupervised learning with variational bayesian mixture of factor analysis to identify patterns of glaucomatous visual field defects. *Invest Ophthalmol Vis Sci*. 2004;45:2596–2605.
28. Yousefi S, Goldbaum MH, Zangwill LM, Medeiros FA, Bowd C. Recognizing patterns of visual field loss using unsupervised machine learning. *Proc SPIE Int Soc Opt Eng*. 2014;2104:pii:90342M.
29. Sample PA, Boden C, Zhang Z, et al. Unsupervised machine learning with independent component analysis to identify areas of progression in glaucomatous visual fields. *Invest Ophthalmol Vis Sci*. 2005;46:3684–3692.
30. Goldbaum MH, Lee I, Jang G, et al. Progression of patterns (POP): a machine classifier algorithm to identify glaucoma progression in visual fields. *Invest Ophthalmol Vis Sci*. 2012;53:6557–6567.
31. Yousefi S, Goldbaum MH, Balasubramanian M, et al. Learning from data: recognizing glaucomatous defect patterns and detecting progression from visual field measurements. *IEEE Trans Biomed Engin*. 2014;61:2112–2124.
32. Goldbaum MH, Jang GJ, Bowd C, et al. Patterns of glaucomatous visual field loss in sita fields automatically identified using independent component analysis. *Trans Am Ophthalmol Soc*. 2009;107:136–144.
33. Fitzke FW, Hitchings RA, Poinoosawmy D, McNaught AI, Crabb DP. Analysis of visual field progression in glaucoma. *Br J Ophthalmol*. 1996;80:40–48.
34. 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.
35. Karakawa A, Murata H, Hirasawa H, Mayama C, Asaoka R. Detection of progression of glaucomatous visual field damage using the point-wise method with the binomial test. *PLoS One*. 2013;8:e78630.
36. Murata H, Araie M, Asaoka R. A new approach to measure visual field progression in glaucoma patients using variational Bayes linear regression. *Invest Ophthalmol Vis Sci*. 2014;55:8386–8392.
37. Sample PA, Girkin CA, Zangwill LM, et al. The African Descent and Glaucoma Evaluation Study (ADAGES): design and baseline data. *Arch Ophthalmol*. 2009;127:1136–1145.
38. Åsman P, Heijl A. Glaucoma hemifield test: automated visual field evaluation. *Arch Ophthalmol*. 1992;110:812–819.
39. Lee T-W, Lewicki MS, Sejnowski TJICA. Mixture models for unsupervised classification of non-Gaussian classes and automatic context

switching in blind signal separation. *IEEE Trans Pattern Anal Mach Intell*. 2000;22:1078–1089.
40. Balasubramanian M, Kriegman DJ, Bowd C, et al. Localized glaucomatous change detection within the proper orthogonal decomposition framework. *Invest Ophthalmol Vis Sci*. 2012;53: 3615–3628.
41. Zhu H, Russell RA, Saunders LJ, Ceccon S, Garway-Heath DF, Crabb DP. Detecting changes in retinal function: Analysis with non-stationary Weibull error regression and spatial enhancement (ANSWERS). *PLoS One*. 2014;9.
42. Yousefi S, Goldbaum MH, Zangwil LM, et al. Glaucomatous retinal nerve fiber layer patterns of loss identified by unsupervised Gaussian model with expectation maximization (GEM) analysis. *21st International Visual Field and Imaging Symposium*. New York, NY; 2014.
43. Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. New York: Chapman & Hall; 1993.