University Library

MINERVA ACCESS

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Rubinstein, NJ;McKendrick, AM;Turpin, A

Title:
Incorporating spatial models in visual field test procedures

Date:
2016-03-11

Citation:
Rubinstein, N. J., McKendrick, A. M. & Turpin, A. (2016). Incorporating spatial models in visual field test procedures. Translational Vision Science and Technology, 5 (2), https://doi.org/10.1167/tvst.5.2.7.

Persistent Link:
https://hdl.handle.net/11343/102419

License:
CC BY-NC-ND

tvst

DOI: 10.1167/tvst.5.2.7

Article

# Incorporating Spatial Models in Visual Field Test Procedures

Nikki J. Rubinstein[1,2], Allison M. McKendrick[1], and Andrew Turpin[2] ✉

[1] Department of Optometry and Vision Sciences, The University of Melbourne, Melbourne, Australia
[2] Computing and Information Systems, The University of Melbourne, Melbourne, Australia

**Correspondence:** Andrew Turpin, Computing and Information Systems, The University of Melbourne, Parkville, 3010, Victoria, Australia. e-mail: aturpin@unimelb.edu.au

**Received:** 26 June 2015
**Accepted:** 30 January 2016
**Published:** 11 March 2016

**Keywords:** perimetry; visual field; glaucoma

**Citation:** Rubinstein NJ, McKendrick AM, Turpin A. Incorporating spatial models in visual field test procedures. Trans Vis Sci Tech. 2016;5(2):7, doi:10.1167/tvst.5.2.7

**Purpose:** To introduce a perimetric algorithm (Spatially Weighted Likelihoods in Zippy Estimation by Sequential Testing [ZEST] [SWeLZ]) that uses spatial information on every presentation to alter visual field (VF) estimates, to reduce test times without affecting output precision and accuracy.

**Methods:** SWeLZ is a maximum likelihood Bayesian procedure, which updates probability mass functions at VF locations using a spatial model. Spatial models were created from empirical data, computational models, nearest neighbor, random relationships, and interconnecting all locations. SWeLZ was compared to an implementation of the ZEST algorithm for perimetry using computer simulations on 163 glaucomatous and 233 normal VFs (Humphrey Field Analyzer 24-2). Output measures included number of presentations and visual sensitivity estimates.

**Results:** There was no significant difference in accuracy or precision of SWeLZ for the different spatial models relative to ZEST, either when collated across whole fields or when split by input sensitivity. Inspection of VF maps showed that SWeLZ was able to detect localized VF loss. SWeLZ was faster than ZEST for normal VFs: median number of presentations reduced by 20% to 38%. The number of presentations was equivalent for SWeLZ and ZEST when simulated on glaucomatous VFs.

**Conclusions:** SWeLZ has the potential to reduce VF test times in people with normal VFs, without detriment to output precision and accuracy in glaucomatous VFs.

**Translational Relevance:** SWeLZ is a novel perimetric algorithm. Simulations show that SWeLZ can reduce the number of test presentations for people with normal VFs. Since many patients have normal fields, this has the potential for significant time savings in clinical settings.

## Introduction

Glaucoma is a disease of the optic nerve, which results in the progressive loss of peripheral vision.[1] Visual function is typically quantified clinically in glaucoma using automated perimetry, which enables measurement of contrast detection thresholds at various locations across the visual field (most typically the central 24–30 degrees). Static white-on-white automated perimetry (SAP) is the most common form of clinical perimetry and has been used as a measure of visual function in a number of large clinical trials.[2–5]

Since SAP's conception, there has been a demand for tests to be faster. SAP requires observers to maintain fixation whilst attending to small dim lights for the duration of the test. As the length of the test increases, observers start to fatigue and have attentional lapses.[6–9] The Full Threshold algorithm was the first SAP procedure and had test times of approximately 9 minutes per eye in normally sighted observers.[10]

Swedish Interactive Threshold Algorithm (SITA) was developed in the 1990s and superseded Full Threshold in the Humphrey Field Analyzer, reducing test times to 4 to 5 minutes for normally sighted observers.[11,12] Most of this time saving was not algorithmic; SITA matched the response window to patient response times and used post hoc analysis to determine false response rates, instead of using catch trials. Using the Zippy Estimation by Sequential Testing (ZEST) algorithm engineered for multiple locations has a similar test time to SITA, with improved accuracy and precision of visual sensitivity estimates.[13,14] Note that these test times increase for observers with visual field loss. While the ZEST





This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

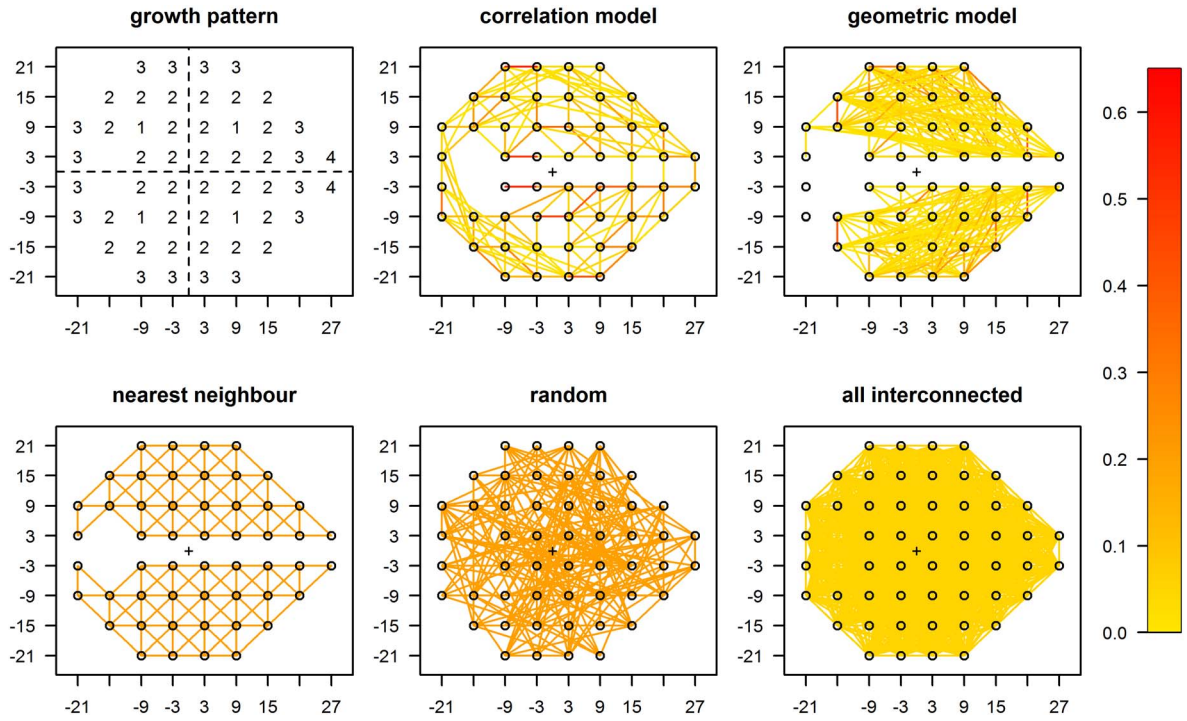



**Figure 1.** Graphical models for SWeLZ using the 24-2 test pattern. The *growth pattern* is the same as that used in SITA and was used for the ZEST procedure in the simulations described here. The four seed locations labelled '1′ (±9, ±9) are tested first. After these four locations have reached the termination criteria, the PMFs of the '2′ locations are seeded with initial values based on the visual sensitivity estimates of the neighboring '1′ locations. After all of the '2′ locations have terminated, the PMFs of the '3′ locations are seeded based on their neighboring '2′ locations and then finally the PMFs of the '4′ locations are seeded based on their neighboring '3′ locations. The *correlation model*, *geometric model*, *nearest neighbor, random*, and *all-interconnected* panels show the interlocation relationships used for the respective models. The edge weights can be interpreted using the colored bar on the right.

algorithm was originally described as an algorithm to threshold a single location in the visual field, we will use "ZEST" in this article to refer to the procedure where ZEST is used at multiple locations.

We hypothesized that incorporating spatial information into the test algorithm would further reduce test times. Currently, spatial information is only loosely incorporated into clinical perimetric algorithms, such as via a growth pattern in SITA[15] (Fig. 1, top left panel). The growth pattern relates neighboring locations by seeding a location's initial stimulus test value from its neighbors. The sequence of test presentations is then performed independently for each location in the visual field. Spatial information is also used to calculate the final visual sensitivity estimates in the SITA postprocessing phase.[16] However, the details of this postprocessing phase are not available in the public domain.

Incorporating spatial information into a perimetric algorithm requires knowledge of the spatial relationships present in the visual field. One such source of spatial information is the arcuate trajectories of the retinal nerve fibres. In glaucoma, when nerve fibres are damaged, defects project into the visual space along the nerve fibre trajectories.[17–19] Spatial information can therefore be inferred at the retinal level from models of nerve fibre trajectories,[20–24] as confirmed in visual space from data sets of glaucomatous visual field defects.[25–30]

In this article, we develop a new algorithm, Spatially Weighted Likelihoods in Zest (SWeLZ), which uses spatial information to help determine the order and luminance of perimetric stimulus presentations. SWeLZ extends the ZEST procedure to update visual sensitivity estimates across multiple locations after each test presentation. The online updating of visual sensitivity estimates across the visual field should allow locations to reach termination in fewer test presentations, which should result in an overall reduction in test time.

The aim of SWeLZ is to reduce test times relative to ZEST, without compromising on the accuracy and precision of visual sensitivity estimates. ZEST was chosen as a comparison clinical algorithm because it

performs with improved accuracy and precision relative to SITA. SITA is commonly used in the clinic but cannot be simulated, since the full details of the procedure are not available in the public domain. The procedures were tested through computer simulation, which allow thousands of simulated observers with known underlying true sensitivities to be tested within a short period of time. Testing new algorithms via computer simulations prior to commencing clinical trials is a well-established methodology in the literature, which has been used to test SITA,[31] as well as other perimetric algorithms.[14,32–36] Computer simulations are also widely used in the development of techniques for both visual field analysis and the detection of visual field progression.[37–39] Comparisons were made on test time and sensitivity estimates.

# Methods

## Software

Custom software was written in C and C++ using the MinGW environment (Minimalist GNU for Windows, version 4.8.1, VA Software). Simulations were run using the Victorian Life Sciences Computation Initiative (VLSCI) supercomputing facility. Analysis was performed in the open-source environment, R (http://www.r-project.org/, in the public domain), using RStudio (version 0.98.501, RStudio, Inc.).

## Test Procedures

### Zippy Estimation by Sequential Testing

ZEST employs a maximum likelihood Bayesian approach to determine the order and luminance of test presentations.[13,14,34,40] The ZEST implementation used in this article was parameterized based on previous literature that has explored how to efficiently run ZEST for perimetry.[34,41–43] The ZEST procedure employed a growth pattern with four waves (Fig. 1, top left panel).

Before testing begins, each location in the visual field is assigned a probability mass function (PMF). The PMF defines the probability that each possible visual sensitivity value (−5 to 40 dB) is the true visual sensitivity of the observer. Note that, although negative intensity values cannot be physically tested, they are included in the domain to prevent a floor effect on visual sensitivity estimation. Negative dB values cannot be tested because luminance values increase with decreasing dB values, and 0 dB represents the maximum intensity that the visual field machine can display.

A bimodal PMF was created using a weighted combination of normal and abnormal visual sensitivity population data (4:1, respectively).[14,43] The resultant curve has two peaks: one peak representing healthy visual sensitivities and another smaller peak representing dysfunctional sensitivities, centered at 0 dB (Fig. 2, top panels). The position of the PMF peak representing healthy visual sensitivities varied according to the wave number of the growth pattern: the four locations in the first wave (±9°, ±9°) had their peaks centered on 30 dB,[43] while locations in waves 2 to 4 had the position of their peaks calculated from the mean of the eccentricity-corrected visual sensitivity estimates of their already terminated neighbors.[33] Eccentricity corrections were performed according to the formula by Hermann, et al.[44] describing the normal hill of vision.

During the test, stimuli are presented with intensity equal to the mean of a location's PMF. After each presentation, a new PMF is generated for the tested location by multiplying the current PMF by a likelihood function. The likelihood function represents the probability that the observer will see the stimulus and is defined by a cumulative Gaussian centered on the test presentation intensity with a standard deviation of 1 dB (see Fig. 2, left-most panels). The test terminates when the standard deviation of the PMF at each location is $< 1.5$ dB.[41,42] The final estimate of visual sensitivity for each location is the mean of the final PMF for that location.

### Spatially Weighted Likelihoods in ZEST

SWeLZ is a newly developed algorithm that extends the ZEST procedure to propagate information across the visual field after each test presentation. SWeLZ uses a spatial graph to define relationships between visual field locations. A spatial graph is defined by edges and edge weights: each edge relates two locations together and the edge weight specifies the strength of this relationship. Larger edge weights indicate stronger relationships. The spatial graphs used in this article are detailed later (see Fig. 1).

Prior to test commencement, each location is assigned a bimodal PMF with peaks at 0 dB and 30 dB (as per wave 1 locations in ZEST). The test then identifies the location with the PMF with the largest standard deviation (i.e., the most uncertainty regarding the sensitivity estimate). If many locations have

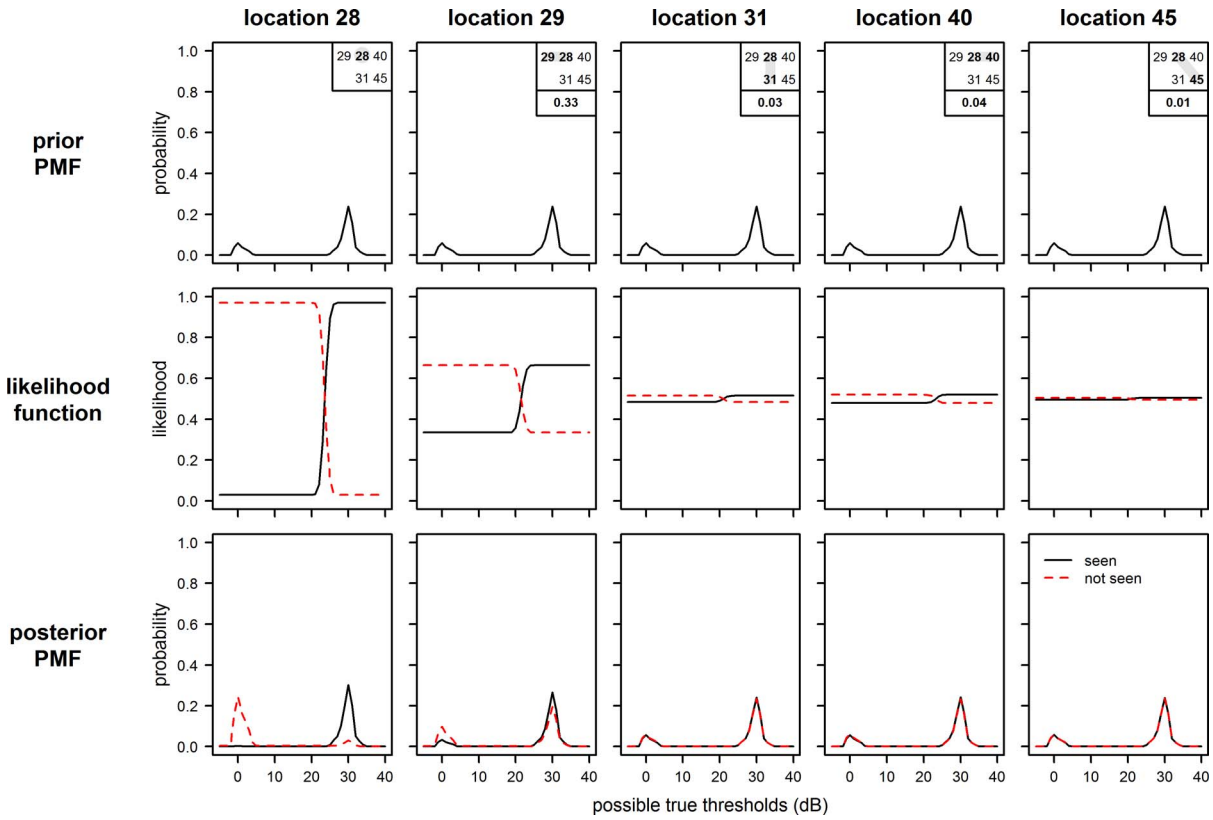



**Figure 2.** SWeLZ example. In this example, the first location tested was location 28. Location 28 is related to locations 29 (weight 0.33), 31 (weight 0.03), 40 (weight 0.04), and 45 (weight 0.01), according to the geometric model. Along the *top row*, the spatial configurations of locations 29, 31, 40, and 45 with respect to location 28 are given in the insets. The value at the bottom of the inset is the edge weight between the tested location 28 and the related locations from the geometric model. The *top row* shows the bimodal PMF assigned to each location at the beginning of testing. A test stimulus is presented at the mean of the PMF of location 28 (24 dB). The PMF is then multiplied by a likelihood function (shown in the *middle row*), whose form is based on whether the observer saw the stimulus (*black*) or not (*red dashed*). The resultant PMF is shown in the *bottom row*. The PMF at each location that shares an edge with location 28 is updated by a modified likelihood function: the height (the vertical distance between upper and lower asymptotes) is equivalent to the edge weight; and the function is centered on the intensity value that was presented at location 28 corrected by an eccentricity factor.[44]

PMFs with the same standard deviation—as occurs at the test beginning—the location is chosen randomly from this pool. The presented stimulus intensity is equal to the mean of the location's PMF.

After each presentation, a new PMF is generated for the tested location, as well as any location that shares an edge with the tested location. The likelihood function used to update the PMFs is eccentricity-corrected and has a height equivalent to the relevant edge weight. Thus, updating strongly related locations has more effect on the shape of the PMF than updating weakly related locations (for an example, see Fig. 2). The test terminates when the PMF at every location is $< 1.5$ dB. The final estimate of visual sensitivity is the mean of the final PMF at each location.

## Defining Spatial Relationships

SWeLZ requires the spatial relationships between locations in the visual field to be defined. Because high edge weights result in extensive spatial smoothing, preliminary experiments optimized the edge weights for the trade-off in speed and error (data not shown). Optimization was achieved using a fabricated data set of 52 visual fields, each with a single location defect of −10 dB (one for each location in the visual field). Edge weights were scaled down until the defect was no longer smoothed out (the median error at the defective location from 100 simulations was <5 dB).

We limited spatial relationships to the 24-2 test pattern, to test SWeLZ on previously collected clinical data. However, it should be noted that

SWeLZ can be used with any test pattern. In order to investigate the effect of using different spatial relationships on the performance of SWeLZ, five different spatial graphs were derived from either published data or heuristic models (Fig. 1), and are explained in turn below.

### Correlation Model

The spatial graph for the correlation model was derived from a previously published spatial filter.[27] The filter was derived from a study data set of 98,821 visual fields from 14,675 individuals, collected using the 30-2 and 24-2 test patterns on the Humphrey Field Analyzer (HFA; Zeiss Humphrey Systems, Dublin, CA).[27] Only the 24-2 locations were used for the derivation of the filter.

For the purpose of creating a spatial graph, the average of two filter values was used to determine the edge weight of the edge shared between each pair of locations. Edge weights were rescaled linearly to have a maximum weight of 0.55 and a minimum weight of 0 (Fig. 1, top middle panel).

### Geometric Model

The spatial graph for the geometric model was derived from a computational model relating retinal ganglion cells to the angle of their insertion at the optic disc.[23,24] The model requires input of axial length and optic nerve head position.

For the purpose of deriving a spatial graph, population average biometric data were used: axial length 25 mm; optic nerve head position (−15°, −2°). Locations that were separated by less than 30° at the disc were considered to share an edge.[45] The edge weight was defined by an exponential decay function on the degree of angular separation at the optic nerve ($\alpha$), and a linear decay function on the Euclidean distance between locations ($\delta$), which was optimized to have a maximum edge weight of 0.47 and a minimum edge weight of 0: $(0.4747 \times 0.0099^{\alpha/30} - 0.0047) \times (6 / \delta)$ (Fig. 1, top right panel).

### Nearest Neighbor

The growth pattern used in SITA is based on nearest neighbor relationships (see Fig. 1, top left panel).[31] In order to have a comparison model for SWeLZ, we created a spatial graph based on nearest neighbors. Locations were connected to their nearest spatial neighbors, without crossing the horizontal midline to be consistent with the SITA growth pattern. All edges were given a weight of 0.2 (Fig. 1, bottom left panel).

### Random

In order to determine whether the specific configuration of edges affects error rates, a spatial graph with randomly assigned edges was generated. Each location in the visual field was randomly assigned edges to four other locations. A location could not be related to itself, and pairings were not allowed to be duplicated. All edges were given a weight of 0.2 (Fig. 1, bottom middle panel).

### All Interconnected

To further explore the effects of edge configurations on error rates, we created a spatial graph in which every location shared an edge with every other location in the visual field. All edges were given a weight of 0.05 (Fig. 1, bottom right panel).

## Computer Simulation

Computer simulations were used to test the newly developed algorithm. The simulation begins by reading in visual sensitivities, which are assumed to represent the true underlying visual sensitivities of the simulated patient. For each simulation, the test procedure is run until all locations have reached the termination criteria. Visual sensitivity estimates are returned rounded to the nearest integer. Patient responses are simulated based on a frequency of seeing curve calculated using Abbott's formula, which has been shown to effectively model human response characteristics.[46,47]

$$\Psi(x,t,s) = 1 - \mathrm{FN} - (1 - \mathrm{FN} - \mathrm{FP}) \cdot G(x,t,s) \quad (1)$$

where FN is the false-negative response rate; FP is the false-positive response rate; $t$ is the true sensitivity; $s$ is the standard deviation of a cumulative Gaussian; $G(x,t,s)$ is the value of a cumulative Gaussian with mean $t$ and standard deviation $s$ at visual sensitivity $x$. The standard deviation of the frequency of seeing curve is known to increase with decreasing visual field sensitivity.[48,49] Thus, the standard deviation of the cumulative Gaussian was varied according to a clinically derived formula, capped at a maximum of 6 dB.[49]

$$s = \min\Big(\exp(-0.081 \times t + 3.27), 6\Big) \quad (2)$$

Both reliable (FP = 3%, FN = 3%) and unreliable (FP = 15%, FN = 3%) observers were simulated. Simulations were run 1000 times for each visual field using each test procedure. Test procedures were assessed by comparing the error—the true sensitivity minus the returned sensitivity estimates—and the number of presentations required for the test to reach completion.

Table. Summary Statistics for ZEST and SWeLZ with Five Different Spatial Models

| | Glaucoma | | Normal | |
|---|---|---|---|---|
| Perimetric Algorithm | Error (dB) | Number of Presentations per Test | Error (dB) | Number of Presentations per Test |
| FP 3% | | | | |
| ZEST | 0 (−1, 1) | 257 (219, 322) | 0.5 (0, 1) | 245 (213, 289) |
| Correlation | 0 (−1, 1) | 252 (162, 371) | 0 (−1, 1) | 191 (155, 281) |
| Geometric | 0 (−1, 1) | 263 (167, 384) | 0 (−1, 1) | 196 (160, 292) |
| Nearest neighbor | 0 (−1, 0.5) | 253 (160, 372) | 0 (−1, 1) | 188 (153, 282) |
| Random | 0 (−1, 1) | 244 (146, 389) | 0 (−1, 1) | 173 (140, 270) |
| All interconnected | 0 (−1, 0.5) | 217 (119, 364) | 0 (−1, 1) | 151 (116, 239) |
| FP 15% | | | | |
| ZEST | 0 (−1.5, 1) | 257 (219, 319) | 0 (−1, 1) | 244 (212, 288) |
| Correlation | −1 (−2, 0) | 238 (160, 365) | 0 (−2, 1) | 186 (154, 256) |
| All interconnected | −1 (−2, 0) | 199 (119, 371) | 0 (−2, 1) | 148 (117, 224) |

Data are presented as median (fifth percentile, 95$^{th}$ percentile). The top six rows indicate statistics for a reliable observer (FP = 3%, FN = 3%), and the bottom three rows indicate statistics for an unreliable observer (FP = 15%, FN = 3%).

## Input Visual Fields

For the computer simulations, we used visual field data that has been used previously for published simulations.[14,42,50–52] Note that these are different visual fields to both the fabricated visual fields used to train the spatial models and the visual fields used to determine the bimodal pattern of the initial PMF. The data set comprises 233 normal and 163 glaucomatous visual fields, obtained using the 24-2 Full Threshold procedure on the HFA (Zeiss Humphrey Systems). These visual fields were collected for a previous study, at which time participants provided written informed consent to have their perimetric data kept in a deidentified database for future research purposes. Normal participants were aged 47 ± 16 years, and glaucomatous participants were aged 61 ± 13 years, with visual field damage ranging from early to severe (median MD = −1.81 dB, fifth percentile = +2.14 dB, 95th percentile = −22.55 dB). Visual fields were altered by 1 dB/decade in order to age-correct them to 45 years and were converted to left eye format. The locations directly adjacent to the blind spot were excluded from analysis (−15° ±3°), resulting in a total of 52 locations for each visual field (see Fig. 1).

## Analysis Measures

The performance of SWeLZ using each of five different spatial graphs (see Fig. 1) was compared with ZEST. Since the data did not follow the form of a normal distribution, the median was used as a measure of central tendency and the 5th to 95th percentile range was used as a measure of spread or variability. Errors were calculated by subtracting the returned estimate of visual sensitivity from the true sensitivity value, resulting in signed errors: negative indicates overestimation and positive indicates underestimation. Thus, the distribution of errors is equivalent to a vertical translation of the distribution of output values. Global errors were calculated for each visual field simulation by taking the median error across all locations. Global errors and numbers of presentations for each visual field simulation were pooled for analysis. Group comparisons were performed using a Wilcoxon rank sum test. Boxplots and difference plots of error by input sensitivity were plotted, to give an indication of how the algorithm performs across different severities of visual field loss.

Very few locations had input sensitivities above 36 dB. Results from these sensitivities are not represented in the boxplots and difference plots, as there is not enough data to make a useful inference regarding the outcome measures. Also note, no locations in the normal data set had input sensitivities below 16 dB, which is reflected in the lower limits of the x-axes in Figures 5 and 7.

# Results

All comparisons of error and number of presentations between ZEST and SWeLZ were statistically

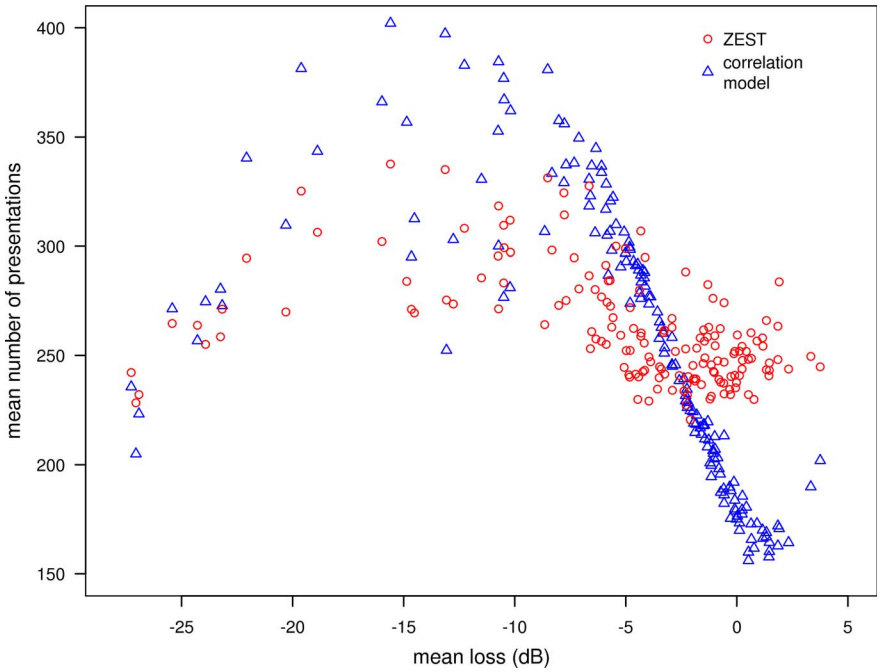



Figure 3. Mean number of presentations against mean loss for the glaucoma data set (ZEST and the correlation model). Mean loss is calculated by subtracting the median of the normal visual fields from each of the glaucomatous input visual field and averaging across the 52 locations.

significant (Wilcoxon $P < 0.001$). However, in large data sets such as those used in this study, small magnitude differences that are statistically significant may not be clinically meaningful. The summary statistics of global error and number of presentations are listed in the Table. The number of presentations plotted against the level of visual field impairment is shown in Figure 3. Figures 4, 5, and 7 show boxplots and difference plots of error across input sensitivity. Figure 6 shows the simulated performance of ZEST and SWeLZ on visual fields with typical glaucomatous visual field defects.

## Reliable Observer (FP = 3%, FN = 3%)

For simulations using a reliable observer, global error performance was similar for all implementations of SWeLZ and ZEST (Table, first and third columns). Median global error was 0 dB for both SWeLZ and ZEST simulated on the glaucoma data set; and 0.5 dB (underestimation) for ZEST and 0 dB for all implementations of SWeLZ simulated on the normal data set. The spread of the global error, as indicated by the 5th to 95th percentile range, was similar for all models simulated on the glaucoma data set (3–4 dB), but was 1 dB greater for all implementations of SWeLZ (2 dB) relative to ZEST (1 dB) simulated on the normal data set.

All implementations of SWeLZ showed a reduction in the number of presentations per visual field relative to ZEST for simulations on the normal data set, while maintaining similar numbers of presentations to ZEST for simulations on the glaucoma data set (Table, second and fourth columns). The median number of presentations was reduced by 20% to 38% for simulations run on the normal data set for SWeLZ relative to ZEST, with the smallest gain seen for the geometric model and the greatest gain seen for the all interconnected model.

Plotting the mean number of presentations against the degree of visual field loss (Fig. 3) reveals that the number of presentations follows a parabola-like shape for both ZEST and the correlation model. The number of presentations is lowest for both mild (mean loss: −10 to 4 dB) and severe (mean loss: −28 to −20 dB) glaucoma and is highest for moderate (mean loss: −10 to −20 dB) glaucoma. However, the range of the number of presentations for the correlation model is approximately twice that of ZEST: the correlation model terminates faster than ZEST in very mild glaucoma (160 vs. 220 presentations) but shows more variability in the number of presentations in moderate glaucoma than ZEST (ZEST range: 270–340; correlation model range: 250–400), on average terminating slightly slower than ZEST.

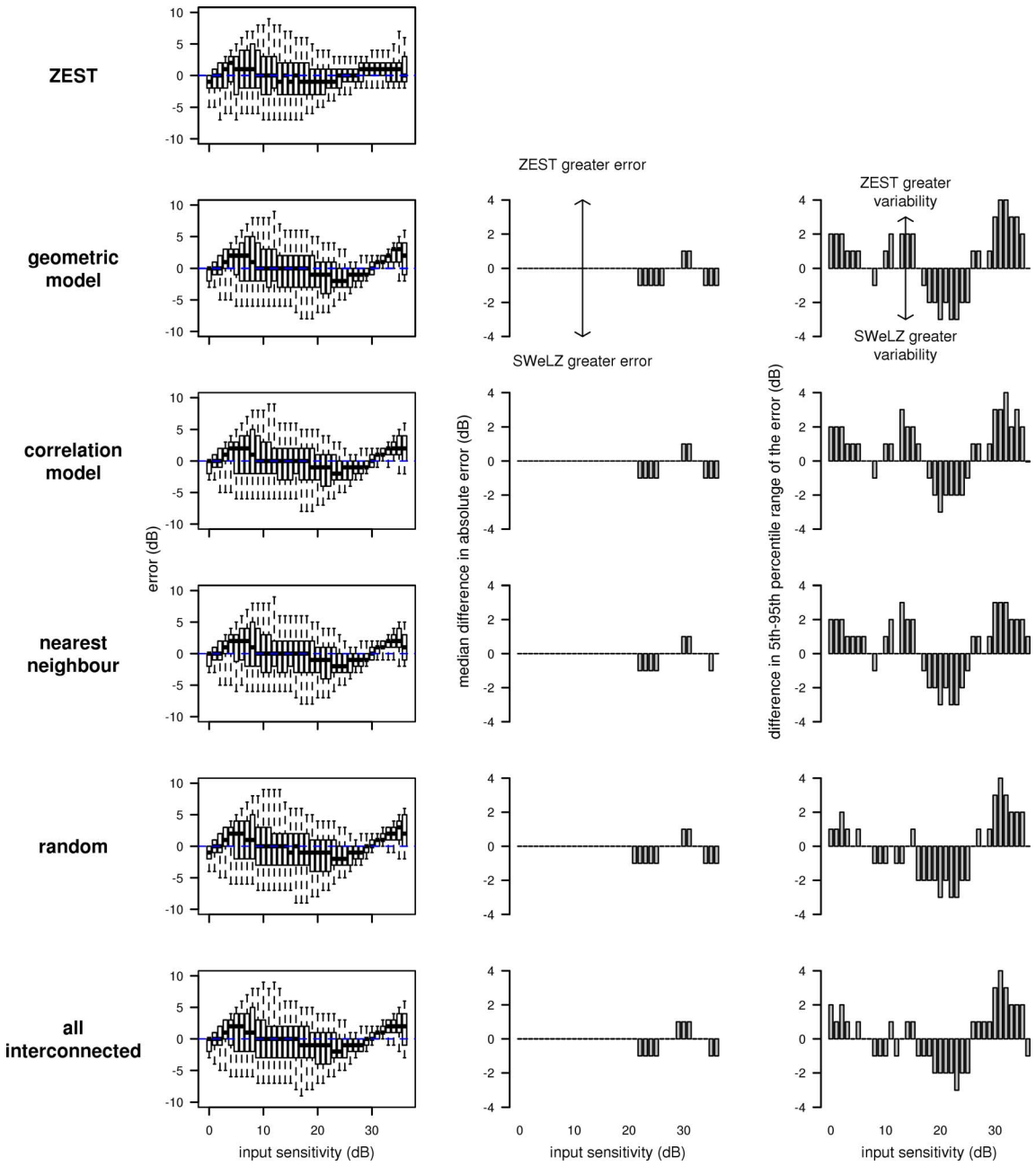



Figure 4. Results from glaucoma data set (FP = 3%, FN = 3%), split by input sensitivity. *Left column*: boxplots of error. *Middle column*: difference plots of median absolute error (ZEST less SWeLZ). *Right column*: difference plots of 5th to 95th percentile range of the error (ZEST less SWeLZ). Positive values in the middle and left columns indicate where SWeLZ performs better than ZEST. Boxplots are drawn such that the bottom and top edges of the box indicate the first and third quartiles (interquartile range [IQR]), respectively; the band inside the box indicates the second quartile (median); and the whiskers represent the 5th and 95th percentiles.

All implementations of SWeLZ had a greater spread of the number of presentations, as shown by the 5th to 95th percentile range. This increased spread manifested as skewed distributions with a tail in the direction of increasing numbers of presentations for simulations on the normal data set, and bimodal distributions for simulations performed on the glaucoma data set.

To compare algorithm performance across varying degrees of visual field loss, we looked at error across input sensitivity. Figures 4 and 5 show the results from simulations performed on the glaucoma and normal data sets, respectively. All imple-

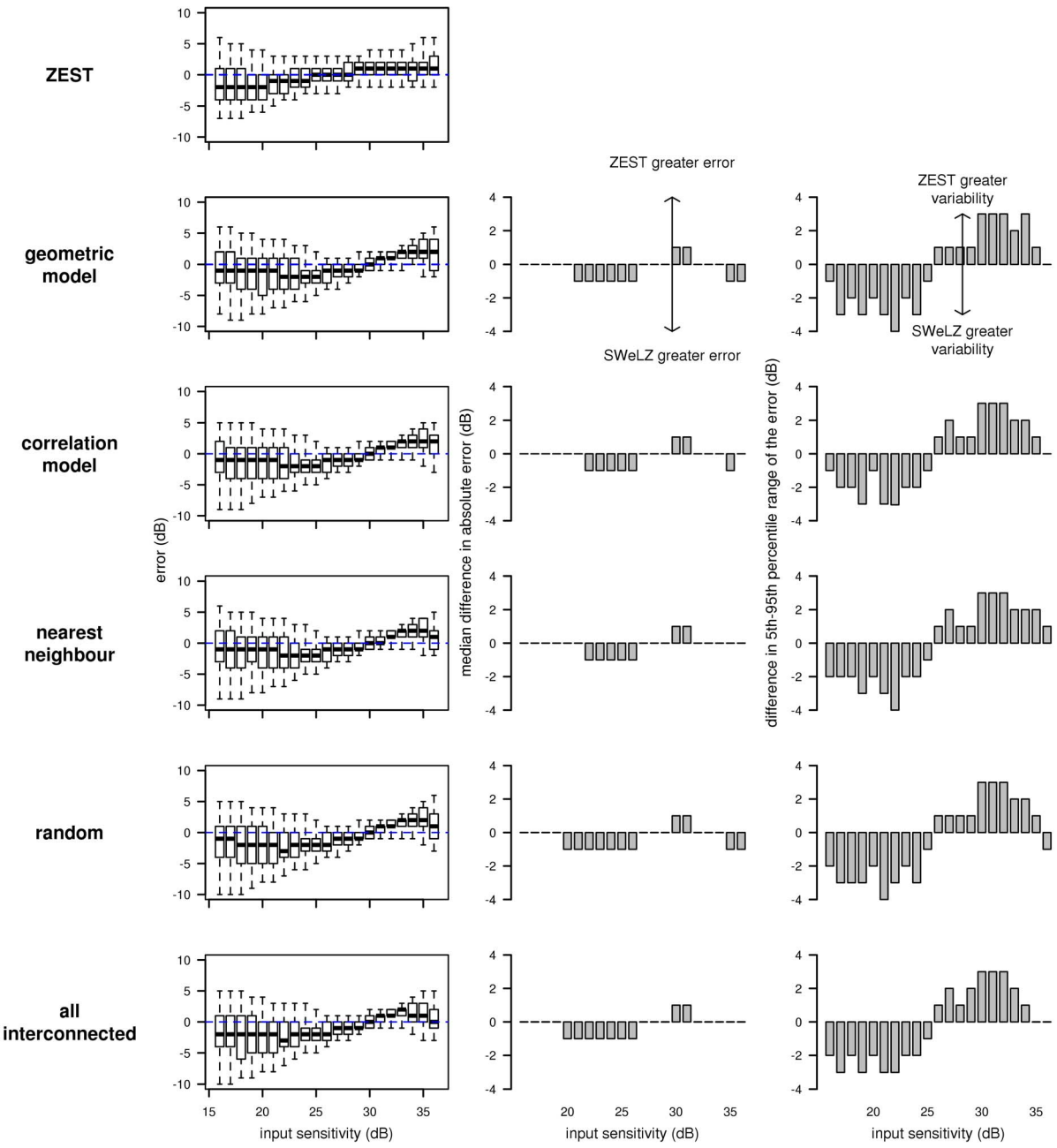



Figure 5. Results from normal data set (FP = 3%, FN = 3%), split by input sensitivity. Formatting is the same as for Figure 4.

mentations of SWeLZ performed the same across input sensitivity for both data sets. Error profiles of ZEST and SWeLZ were also similar across all conditions.

For the glaucoma data set, comparisons of SWeLZ with ZEST revealed no median difference in absolute errors for intensity values < 22 dB and differences in the order of 1 dB for sensitivity values > 21 dB (SWeLZ error greater than ZEST: 22–26 dB and 34–36 dB; ZEST error greater than SWeLZ: 29–31 dB). SWeLZ had reduced spread of error over the lower (0–16 dB) and higher ends (26–36 dB) of the dynamic range of intensities, while ZEST had reduced spread centrally (17–25 dB). Differences in spread were of the magnitudes: 1–4 dB.

For the normal data set, comparisons of SWeLZ with ZEST revealed median differences in absolute error in the order of 1 dB (SWeLZ error greater than ZEST: 20–26 dB and 35–36 dB; ZEST error greater than SWeLZ: 30–31 dB). SWeLZ has reduced spread

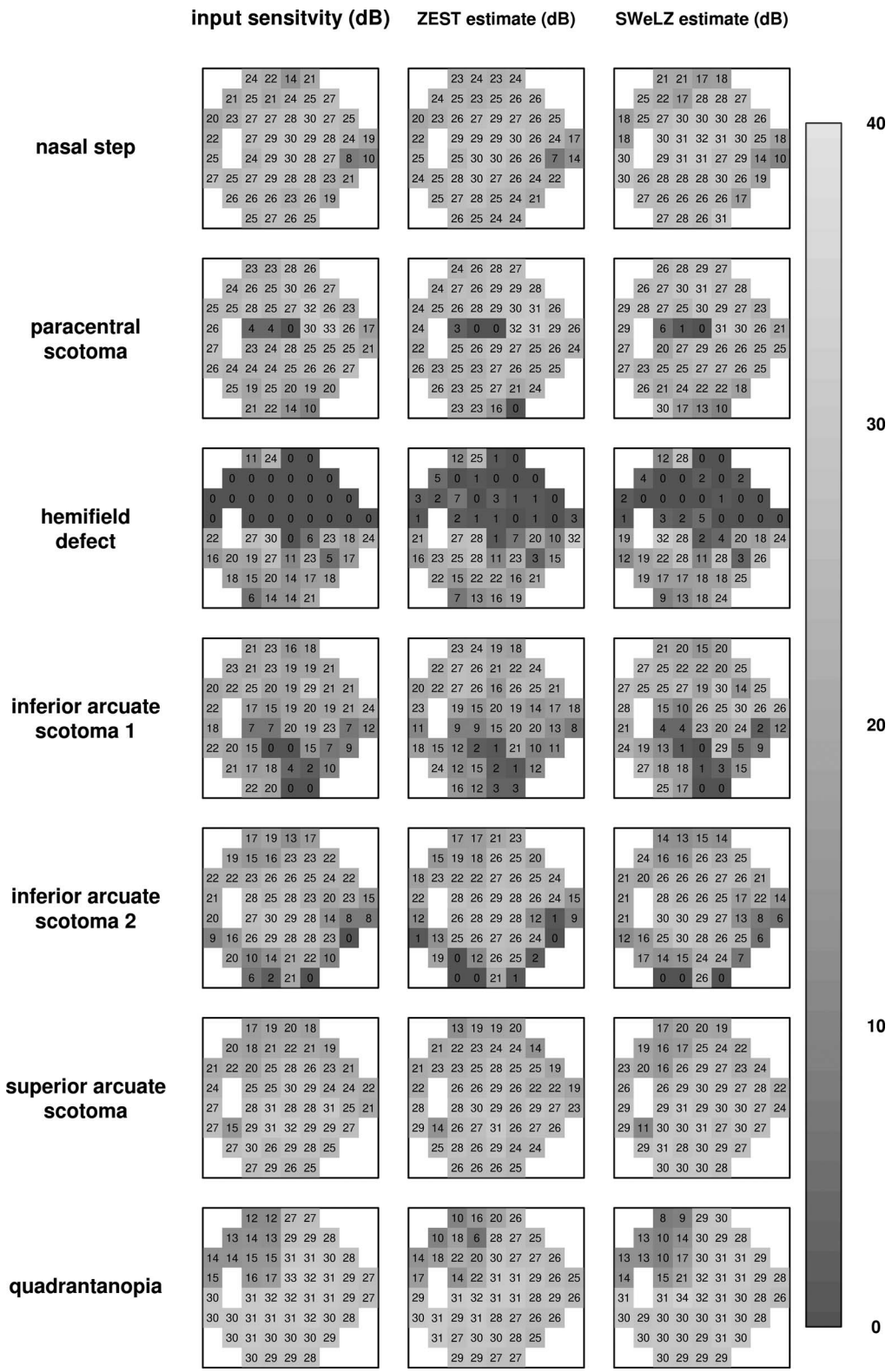



Figure 6. Estimates of ZEST and SWeLZ (correlation model) for visual fields with typical glaucomatous defects and a quadrantanopia for a randomly chosen simulation (FP = 3%, FN = 3%). *Left column*: input sensitivity. *Middle column*: estimate for ZEST procedure. *Right column*: estimate of SWeLZ (correlation model) procedure.

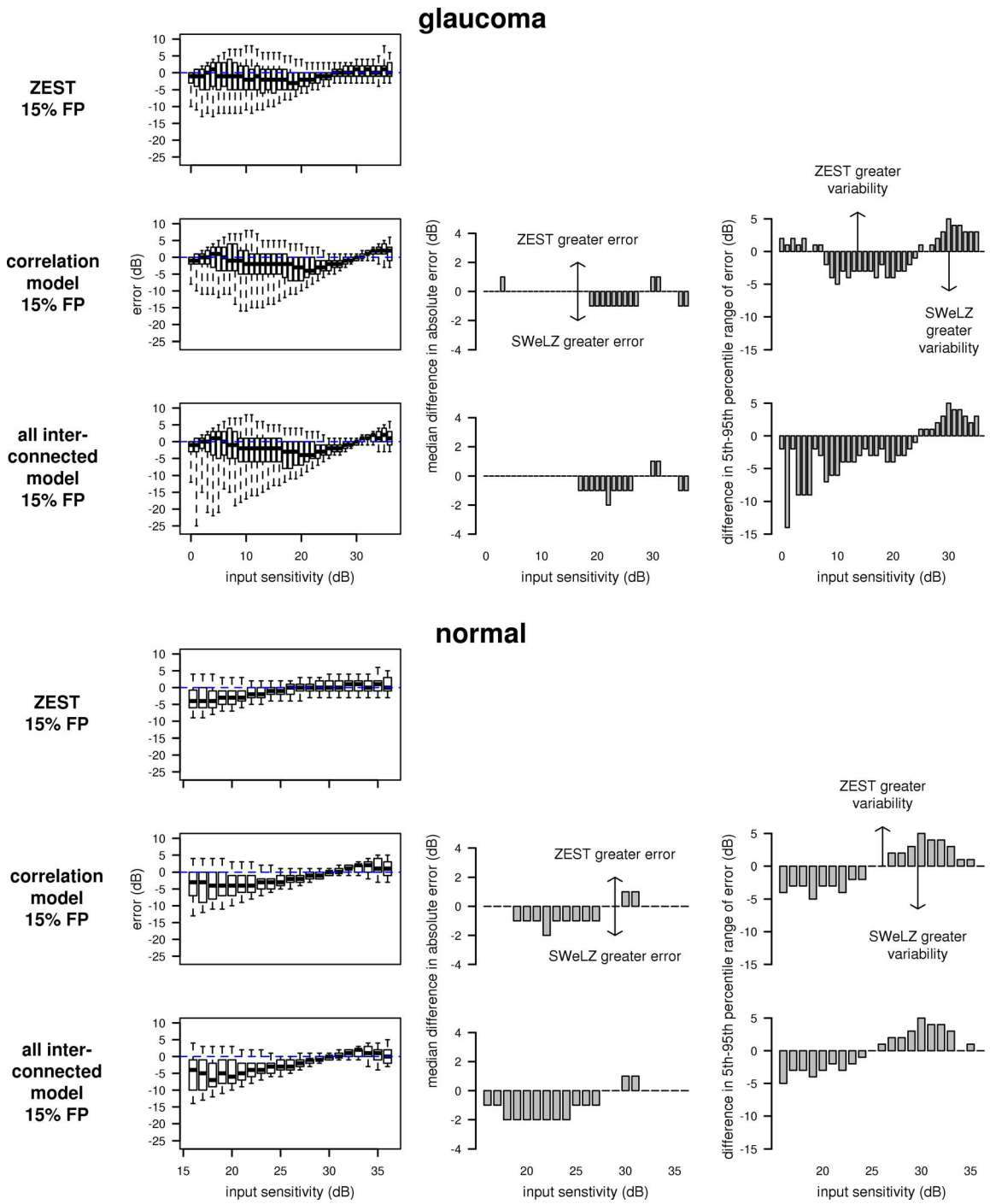



Figure 7. Results for an unreliable observer (FP = 15%, FN = 3%), split by input sensitivity. *Top three rows*: results from glaucoma data set. *Bottom three rows:* results from normal data set. Column formatting is the same as Figure 4.

at the higher end of the range of intensities (26–36 dB), while ZEST had reduced variability relative to ZEST at lower dB values (16–25 dB). The magnitude of the difference in spread spanned: 1–4 dB.

In order to verify that localized defects were not being smoothed out, we looked at visual fields with typical patterns of glaucomatous loss. The results of randomly chosen simulations of ZEST and the correlation model are shown in Figure 6. Both ZEST and SWeLZ were capable of detecting a paracentral scotoma, nasal step, hemifield loss, and arcuate defects.

### Unreliable Observer (FP = 15%, FN = 3%)

A typical false-positive responder was simulated using both ZEST and SWeLZ (correlation and all-interconnected models). The correlation model was chosen for SWeLZ, because it showed the greater reduction in test times of the two structurally and functionally derived models. The all-interconnected model was chosen as a nonspecific spatial model to contrast with the literature-derived correlation model. The relative performance of the correlation model with ZEST was similar to the simulations using the reliable observer. However, the tests simulated using the all-interconnected model tended to smooth out localized defects.

Median global error was 0 dB for ZEST and SWeLZ (both geometric and all-interconnected models) for simulations on the normal data set, with SWeLZ having a larger spread of global error (3 dB) relative to ZEST (2 dB) (Table, bottom three rows). Simulations using the glaucoma data set, revealed a median error of 0 dB for ZEST and −1 dB (overestimation) for SWeLZ, with a similar spread of global error for the two procedures (2–2.5 dB).

The reduction in the number of presentations for SWeLZ relative to ZEST for the normal data set was of a similar magnitude to simulations using the reliable observer (correlation model: 24% reduction; all interconnected model: 39% reduction). ZEST terminated in a similar number of presentations (257) for the unreliable as for the reliable observer using the glaucoma data set. However, results from the glaucoma data set showed that SWeLZ required slightly fewer presentations for the unreliable than the reliable observer (correlation model: 238; all interconnected model: 199). This is likely because false-positive responses give the semblance of a smoother field.

Splitting the error by input sensitivity revealed greater spread for the unreliable relative to the reliable observer. However, the relationship between the median error and spread of the error for the correlation model and ZEST was the same as for the reliable observer (Fig. 4 versus Fig. 7 top rows; Fig. 5 versus Fig. 7 bottom rows). While the median error was similar for the reliable and unreliable all-interconnected model results, localized defects tended to be smoothed out for the unreliable observer. This is shown by the large increase in the 5th to 95th percentile range of locations with input sensitivities < 15 dB (Fig. 4 bottom row versus Fig. 7 third row).

## Discussion

We developed a novel algorithm that incorporates spatial information into the choice of where and at what intensity to present perimetric stimuli. The key difference between SWeLZ and previous uses of ZEST for perimetry is that PMFs are updated at multiple locations after each response, rather than just at the location where the stimulus was presented.

Computer simulations were used to compare the performance of SWeLZ with ZEST. Computer simulations provide a powerful method for evaluating an algorithm's performance prior to the initiation of clinical trials. Thousands of tests can be run in a relatively short space of time, with known underlying visual sensitivities and controlled patient response variability. These benefits are not possible with real human observers.

There is a constant demand for perimetric tests to be faster. SWeLZ was designed to meet this demand. Relative to ZEST, simulated test times were reduced by 20% to 38% for normal visual fields. These test times are shorter than those expected for SITA by approximately 32% to 47%, which takes an average of 287 presentations on normal visual fields.[12] Many people who regularly undergo visual field assessment in primary care clinical settings have normal visual fields. SWeLZ provides a reduction in test time for these patients, without increasing error rates. A further advantage of SWeLZ relative to the growth pattern used in SITA is that peripheral locations do not have to be tested last. When the peripheral locations are always tested last, these same locations will always be tested when the patient is most fatigued by the test.

The reduction in test times is present only for normal visual fields because the difference between neighboring thresholds in normal visual fields is small: the hill of vision is smooth. Connecting locations across a relatively smooth surface allows locations to reach the termination criteria faster, without detriment to the final visual sensitivity estimate.

It is likely that the correlation and geometric models, while based on population data and representing the population average, are not true representations for a particular individual. These models do not take into account an individual's biometric data, which determines the distribution of nerve fibres and thus the spatial relationships within glaucomatous visual fields.[22–24,53] For this article, we used

population average biometric data to create models of spatial relationships, but tested these models with data that likely comprised a variety of different spatial relationships. Individuals within the data set may not be well described by the population average in some cases.

While the spatial graphs may be able to be refined, our data demonstrate that the exact configuration of the spatial graph is not the limiting factor on SWeLZ's performance. Although the spatial graphs used in this study had spatial configurations that differed from each other, similar error profiles were produced for each of them (Figs. 4, 5). The greatest improvement in test times for both data sets was achieved with a graph that had no spatial information; all locations were interconnected with equal edge weights (see Table). In effect, SWeLZ with this graph produces a raising or lowering of the height of the hill of vision with each stimulus presentation. SWeLZ is thus not glaucoma-specific, as can be seen by the ability to detect the quadrantanopia shown in Figure 6. However, sensible relationships should be chosen so as not to smooth out localized defects when the observer is an unreliable responder (Fig. 7). Thus, it may be prudent to incorporate a real-time decision-making process at the front-end of SWeLZ that allows for changes in spatial models depending on the suspected cause of visual field loss, as determined by a clinician.

Additional simulations were run to see whether increasing the number of test presentations beyond the test termination criteria (SD $<$ 1.5 dB) would improve the error profile (not shown) to create a procedure of the same duration as current procedures but more accurate. Running SWeLZ for longer did not improve the error profile. As SWeLZ is based on ZEST, it may be that this reflects the best performance achievable with ZEST and that either a different approach is required to improve the accuracy and precision of sensitivity estimates, or that performance simply cannot be improved beyond these levels due to other variability factors as described by the frequency of seeing curve of the responder. Alternately, using a more principled approach to incorporate spatial information, such as conditional random fields—a form of statistical modeling—may improve error rates, by incorporating combined probability functions across multiple visual field locations.[16,54]

SWeLZ is not a screening approach but provides genuine threshold estimates. There are several other perimetric approaches that are rapid for normal visual fields (e.g., Tendency Oriented Perimetry [TOP] and SITA-Fast). Key differences are that while SWeLZ dynamically chooses the number of presentations at each location, TOP presents only one stimulus at each location, commonly underestimating defect depth.[55,56] SITA-Fast terminates earlier than SITA accepting a lower accuracy of test results.[57,58] SWeLZ provides the benefits of a fast test procedure for smooth, normal visual fields, while expending a comparable or slightly greater number of test presentations to existing procedures for glaucomatous visual fields, without compromising on the precision and accuracy of test results in either case.

The next step for SWeLZ will be to test the algorithm on real observers. Computer simulations make assumptions about patient responses, such as the patient's frequency of seeing curve, attentional lapses and effects of fatigue, which do not necessarily reflect the response characteristics of a real observer. Thus, clinical testing is required to validate the findings described in this article.

In summary, a novel algorithm was developed to incorporate spatial information into the update instructions of a Bayesian maximum likelihood procedure. SWeLZ required 20% to 38% less presentations than ZEST to complete a visual field test for a normal visual field, without detriment to the error profile and while still maintaining accuracy on glaucomatous visual fields. Additionally, since the order in which locations are tested is not determined by a growth pattern, SWeLZ does not require peripheral locations to be the last locations tested. SWeLZ has the potential to reduce test times for patients requiring regular visual field tests, who do not have manifest visual field loss.

## Acknowledgments

The authors thank Chris Johnson for supplying the visual field data sets. The authors thank Stuart Gardiner for supplying details of the filter used to derive the correlation model spatial graph.

This research was supported by grant ARC LP130100055 (AT and AMM) and a Victorian Life Sciences Computation Initiative (VLSCI) grant [VR0280] on its Peak Computing Facility at the University of Melbourne, an initiative of the Victorian Government, Australia.

Disclosure: **N. J. Rubinstein**, None; **A. M. McKendrick**, Heidelberg Engineering GmbH, Haag-Streit

AG, CentreVue SpA, Carl-Zeiss Meditec; **A. Turpin**, Heidelberg Engineering GmBH, Haag-Streit AG, CentreVue SpA

## References


1. Wensor MD, McCarty CA, Stanislavsky YL, Livingston PM, Taylor HR. The prevalence of glaucoma in the Melbourne Visual Impairment Project. *Ophthalmology*. 1998;105:733–739.
2. Leske MC, Heijl A, Hyman L, Bengtsson B. Early Manifest Glaucoma Trial: design and baseline data. *Ophthalmology*. 1999;106:2144–2153.
3. Leske MC, Heijl A, Hyman L, et al. Predictors of long-term progression in the early manifest glaucoma trial. *Ophthalmology*. 2007;114:1965–1972.
4. Musch DC, Lichter PR, Guire KE, Standardi CL. The Collaborative Initial Glaucoma Treatment Study: study design, methods, and baseline characteristics of enrolled patients. *Ophthalmology*. 1999;106:653–662.
5. Ederer F, Gaasterland DE, Sullivan EK, AGIS Investigators. The Advanced Glaucoma Intervention Study (AGIS): 1. Study design and methods and baseline characteristics of study patients. *Control Clin Trials*. 1994;15:299–325.
6. Henson DB, Emuh T. Monitoring vigilance during perimetry by using pupillography. *Invest Ophthalmol Vis Sci*. 2010;51:3540–3543.
7. Heijl A, Drance SM. Changes in differential threshold in patients with glaucoma during prolonged perimetry. *Br J Ophthalmol*. 1983;67:512–516.
8. Hudson C, Wild JM, O'Neill EC. Fatigue effects during a single session of automated static threshold perimetry. *Invest Ophthalmol Vis Sci*. 1994;35:268–280.
9. Johnson CA, Adams CW, Lewis RA. Fatigue effects in automated perimetry. *Appl Opt*. 1988;27:1030–1037.
10. Wabbels BK, Diehm S, Kolling G. Continuous light increment perimetry compared to full threshold strategy in glaucoma. *Eur J Ophthalmol*. 2005;15:722–729.
11. Wall M, Punke SG, Stickney TL, Brito CF, Withrow KR, Kardon RH. SITA standard in optic neuropathies and hemianopias: a comparison with full threshold testing. *Invest Ophthalmol Vis Sci*. 2001;42:528–537.
12. Bengtsson B, Heijl A, Olsson J. Evaluation of a new threshold visual field strategy, SITA, in normal subjects. *Acta Ophthalmol Scand*. 1998;76:165–169.
13. King-Smith PE, Grigsby SS, Vingrys AJ, Benes SC, Supowit A. Efficient and unbiased modifications of the QUEST threshold method: theory, simulations, experimental evaluation and practical implementation. *Vision Res*. 1994;34:885–912.
14. Turpin A, McKendrick AM, Johnson CA, Vingrys AJ. Properties of perimetric threshold estimates from full threshold, ZEST, and SITA-like strategies, as determined by computer simulation. *Invest Ophthalmol Vis Sci*. 2003;44:4787–4795.
15. Anderson D, Patella V. *Automated Static Perimetry*. St Louis, MO: Mosby-Year Book; 1999.
16. Olsson J, Rootzén H. An image model for quantal response analysis in perimetry. *Scand J Stat*. 1994;375–387.
17. Westcott M, Garway-Heath D, Fitzke F, Kamal D, Hitchings R. Use of high spatial resolution perimetry to identify scotomata not apparent with conventional perimetry in the nasal field of glaucomatous subjects. *Br J Ophthalmol*. 2002;86:761–766.
18. Westcott MC, McNaught AI, Crabb DP, Fitzke FW, Hitchings RA. High spatial resolution automated perimetry in glaucoma. *Br J Ophthalmol*. 1997;81:452–459.
19. Nevalainen J, Paetzold J, Papageorgiou E, et al. Specification of progression in glaucomatous visual field loss, applying locally condensed stimulus arrangements. *Graefes Arch Clin Exp Ophthalmol*. 2009;247:1659–1669.
20. Garway-Heath DF, Poinoosawmy D, Fitzke FW, Hitchings RA. Mapping the visual field to the optic disc in normal tension glaucoma eyes. *Ophthalmology*. 2000;107:1809–1815.
21. Jansonius NM, Schiefer J, Nevalainen J, Paetzold J, Schiefer U. A mathematical model for describing the retinal nerve fiber bundle trajectories in the human eye: average course, variability, and influence of refraction, optic disc size and optic disc position. *Exp Eye Res*. 2012;105:70–78.
22. Lamparter J, Russell RA, Zhu H, et al. The influence of intersubject variability in ocular anatomical variables on the mapping of retinal locations to the retinal nerve fiber layer and optic nerve head. *Invest Ophthalmol Vis Sci*. 2013;54:6074–6082.
23. Denniss J, McKendrick AM, Turpin A. An anatomically customizable computational model relating the visual field to the optic nerve head in

individual eyes. *Invest Ophthalmol Vis Sci*. 2012; 53:6981–6990.
24. Turpin A, Sampson GP, McKendrick AM. Combining ganglion cell topology and data of patients with glaucoma to determine a structure-function map. *Invest Ophthalmol Vis Sci*. 2009;50: 3249–3256.
25. Crabb D, Fitzke F, McNaught A, Hitchings R. A profile of the spatial dependence of pointwise sensitivity across the glaucomatous visual field. In: Wall M, Heijl A, eds. *Perimetry Update 1996/ 1997, Proceedings of the XIIth International Perimetric Society Meeting*. Würzburg/Amsterdam: Kugler Publications; 1997:301–311.
26. Erler NS, Bryan SR, Eilers PH, Lesaffre EM, Lemij HG, Vermeer KA. Optimizing structure-function relationship by maximizing correspondence between glaucomatous visual fields and mathematical retinal nerve fiber models. *Invest Ophthalmol Vis Sci*. 2014;55:2350–2357.
27. Gardiner SK, Crabb DP, Fitzke FW, Hitchings RA. Reducing noise in suspected glaucomatous visual fields by using a new spatial filter. *Vision Res*. 2004;44:839–848.
28. González de la Rosa M, González-Hernández M, Abraldes M, Azuara-Blanco A. Quantification of interpoint topographic correlations of threshold values in glaucomatous visual fields. *J Glaucoma*. 2002;11:30–34.
29. Güerri N, Polo V, Larrosa J, Egea C, Ferreras A, Pablo L. Functional relationship between retinal sensitivity threshold values assessed by standard automated perimetry in glaucoma. *Arch Soc Esp Oftalmol*. 2013;88:223–230.
30. Heijl A, Lindgren A, Lindgren G. Inter-point correlations of deviations of threshold values in normal and glaucomatous visual fields. In: Heijl A ed. *Perimetry Update 1988/1989, Proceedings of the VIIIth International Perimetric Society Meeting*. Vancouver/Amsterdam: Kugler Publications; 1988:177–183.
31. Bengtsson B, Olsson J, Heijl A, Rootzén H. A new generation of algorithms for computerized threshold perimetry, SITA. *Acta Ophthalmol Scand*. 1997;75:368–375.
32. Artes PH, Henson DB, Harper R, McLeod D. Multisampling suprathreshold perimetry: a comparison with conventional suprathreshold and full-threshold strategies by computer simulation. *Invest Ophthalmol Vis Sci*. 2003;44:2582–2587.
33. Chong LX, McKendrick AM, Ganeshrao SB, Turpin A. Customized, automated stimulus location choice for assessment of visual field defects. *Invest Ophthalmol Vis Sci*. 2014;55:3265–3274.
34. Turpin A, McKendrick AM, Johnson CA, Vingrys AJ. Development of efficient threshold strategies for frequency doubling technology perimetry using computer simulation. *Invest Ophthalmol Vis Sci*. 2002;43:322–331.
35. Shapiro L, Johnson C, Kennedy R. Kraken: a computer simulation procedure for static, kinetic, supra-threshold and heuristic perimetry. In: Heijl A, ed. *Perimetry Update 1988/1989, Proceedings of the VIIIth International Perimetric Society Meeting*. Vancouver/Amsterdam; Kugler Publications; 1988:431–438.
36. Johnson C, Shapiro L. A comparison of MOBS (modified binary search) and staircase test procedures in automated perimetry. *Noninvasive Assessment of the Visual System, 1989 Tech Digest Ser 7*. 1989:84–87.
37. Russell RA, Garway-Heath DF, Crabb DP. New insights into measurement variability in glaucomatous visual fields from computer modelling. *PLoS One*. 2013;8:e83595.
38. Jansonius NM. Progression detection in glaucoma can be made more efficient by using a variable interval between successive visual field tests. *Graefes Arch Clin Exp Ophthalmol*. 2007;245: 1647–1651.
39. Crabb DP, Garway-Heath DF. Intervals between visual field tests when monitoring the glaucomatous patient: wait-and-see approach. *Invest Ophthalmol Vis Sci*. 2012;53:2770–2776.
40. Turpin A, McKendrick AM, Johnson CA, Vingrys AJ. Performance of efficient test procedures for frequency-doubling technology perimetry in normal and glaucomatous eyes. *Invest Ophthalmol Vis Sci*. 2002;43:709–715.
41. McKendrick AM, Denniss J, Turpin A. Response times across the visual field: Empirical observations and application to threshold determination. *Vision Res*. 2014;101:1–10.
42. McKendrick AM, Turpin A. Advantages of terminating Zippy Estimation by Sequential Testing (ZEST) with dynamic criteria for white-on-white perimetry. *Optom Vis Sci*. 2005;82:981–987.
43. Vingrys AJ, Pianta MJ. A new look at threshold estimation algorithms for automated static perimetry. *Optom Vis Sci*. 1999;76:588–595.
44. Hermann A, Paetzold J, Vonthein R, Krapp E, Rauscher S, Schiefer U. Age-dependent normative values for differential luminance sensitivity in automated static perimetry using the Octopus 101. *Acta Ophthalmol*. 2008;86:446–455.

45. Denniss J, Turpin A, McKendrick AM. Individualised structure-function mapping for glaucoma: Practical constraints on map resolution for clinical and research applications. *Invest Ophthalmol Vis Sci*. 2014;55:1985–1993.
46. Treutwein B, Strasburger H. Fitting the psychometric function. *Percept Psychophys*. 1999;61:87–106.
47. Wichmann FA, Hill NJ. The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept Psychophys*. 2001;63:1293–1313.
48. Wall M, Kutzko KE, Chauhan BC. Variability in patients with glaucomatous visual field damage is reduced using size V stimuli. *Invest Ophthalmol Vis Sci*. 1997;38:426–435.
49. Henson DB, Chaudry S, Artes PH, Faragher EB, Ansons A. Response variability in the visual field: Comparison of optic neuritis, glaucoma, ocular hypertension, and normal eyes. *Invest Ophthalmol Vis Sci*. 2000;41:417–421.
50. McKendrick AM, Turpin A. Combining perimetric suprathreshold and threshold procedures to reduce measurement variability in areas of visual field loss. *Optom Vis Sci*. 2005;82:43–51.
51. Turpin A, Jankovic D, McKendrick AM. Retesting visual fields: utilizing prior information to decrease test-retest variability in glaucoma. *Invest Ophthalmol Vis Sci*. 2007;48:1627–1634.
52. Denniss J, McKendrick AM, Turpin A. Towards patient-tailored perimetry: automated perimetry can be improved by seeding procedures with patient-specific structural information. *Transl Vis Sci Technol*. 2013;2.
53. Chauhan BC, Burgoyne CF. From clinical examination of the optic disc to clinical assessment of the optic nerve head: A paradigm change. *Am J Ophthalmol*. 2013;156:218–227.e212.
54. Geman S, Geman D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE T Pattern Anal*. 1984;721–741.
55. Lachkar Y, Barrault O, Lefrancois A, Demailly P. Rapid Tendency Oriented Perimeter (TOP) with the Octopus visual field analyzer. *J Fr Ophtalmol*. 1998;21:180–184.
56. Morales J, Weitzman ML. González de la Rosa M. Comparison between tendency-oriented perimetry (TOP) and Octopus threshold perimetry. *Ophthalmology*. 2000;107:134–142.
57. Bengtsson B, Heijl A. SITA Fast, a new rapid perimetric threshold test. Description of methods and evaluation in patients with manifest and suspect glaucoma. *Acta Ophthalmol Scand*. 1998;76:431–437.
58. King AJ, Taguri A, Wadood AC, Azuara-Blanco A. Comparison of two fast strategies, SITA Fast and TOP, for the assessment of visual fields in glaucoma patients. *Graefes Arch Clin Exp Ophthalmol*. 2002;240:481–487.