Article

Assessing the GOANNA Visual Field Algorithm Using Artificial Scotoma Generation on Human Observers

Luke X. Chong[1,2], Andrew Turpin[3], and Allison M. McKendrick[1]

[1] Department of Optometry & Vision Sciences, The University of Melbourne, Melbourne, Australia
[2] School of Optometry, University of California Berkeley, Berkeley, CA, USA
[3] Department of Computing & Information Systems, The University of Melbourne, Melbourne, Australia

Correspondence: Luke X. Chong, University of California Berkeley, School of Optometry, 591A Minor Hall, Berkeley, CA 94720, USA; email: lxchong@berkeley.edu

Purpose: To validate the performance of a new perimetric algorithm (Gradient-Oriented Automated Natural Neighbor Approach; GOANNA) in humans using a novel combination of computer simulation and human testing, which we call Artificial Scotoma Generation (ASG).

Methods: Fifteen healthy observers were recruited. Baseline conventional automated perimetry was performed on the Octopus 900. Visual field sensitivity was measured using two different procedures: GOANNA and Zippy Estimation by Sequential Testing (ZEST). Four different scotoma types were induced in each observer by implementing a novel technique that inserts a step between the algorithm and the perimeter, which in turn alters presentation levels to simulate scotomata in human observers. Accuracy, precision, and unique number of locations tested were measured, with the maximum difference between a location and its neighbors *(Max_d)* used to stratify results.

Results: GOANNA sampled significantly more locations than ZEST (paired *t*-test, $P < 0.001$), while maintaining comparable test times. Difference plots showed that GOANNA displayed greater accuracy than ZEST when *Max_d* was in the 10 to 30 dB range (with the exception of *Max_d* = 20 dB; Wilcoxon, $P < 0.001$). Similarly, GOANNA demonstrated greater precision than ZEST when *Max_d* was in the 20 to 30 dB range (Wilcoxon, $P < 0.001$).

Conclusions: We have introduced a novel method for assessing accuracy of perimetric algorithms in human observers. Results observed in the current study agreed with the results seen in earlier simulation studies, and thus provide support for performing larger scale clinical trials with GOANNA in the future.

Translational Relevance: The GOANNA perimetric testing algorithm offers a new paradigm for visual field testing where locations for testing are chosen that target scotoma borders. Further, the ASG methodology used in this paper to assess GOANNA shows promise as a hybrid between computer simulation and patient testing, which may allow more rapid development of new perimetric approaches.

Introduction

In conventional perimetry, the spatial sampling resolution of commonly used programs such as the Humphrey Field Analyser (HFA; Zeiss Humphrey Systems, Dublin, CA) is typically conducted on an equidistant 6° × 6° rectangular grid. However, it has been shown that a 6° grid may be too coarse to adequately characterize visual field loss due to spatial aliasing through undersampling of the field.[1] Furthermore, Wyatt and colleagues[2] not only demonstrated that visual field measurement variability was correlated with gradient, but also suggested that accurate scotomata characterization requires sampling at a higher spatial resolution than 6°.

Previously, we introduced a new perimetric algorithm that autonomously selects a subset of test locations from a larger pool of 150 possible locations spaced 3° apart, called Gradient-Oriented Automated Natural Neighbor Approach (GOANNA).[3,4] GOANNA allows for sampling off the conventional 6° grid by increasing the spatial resolution in regions surrounding scotoma edges. An important feature of GOANNA is that its test time is comparable to current commercially implemented procedures despite

sampling a greater number of spatial locations. Upon completion of the procedure, thresholds for locations that have not been measured are determined through natural neighbor interpolation.[5,6] As a result, the final output is a 3° grid with threshold values returned for all 150 locations. Previous studies indicate that we get reasonable estimates at all points.[3,4]

Through simulation, we have demonstrated that GOANNA can be more accurate and precise than a baseline Bayesian perimetric strategy.[3] We also showed that this improvement in accuracy and precision around scotoma borders can result in earlier detection of simulated glaucomatous progression.[4] Although computer simulation is useful in allowing investigators to assess performance of a procedure, it may not be completely indicative of the true responses seen in human observers. Fatigue and learning effects can manifest in human observers,[7–12] but were not modelled in our earlier simulation studies, and may differentially alter locations that are tested at the beginning relative to the end of the test procedure. Furthermore, testing a procedure on humans is essential to determine what they experience while undertaking the test. Therefore, both computer simulation and human testing is required in order to obtain a thorough understanding of procedure performance.

There have been many studies in the past that have compared the performance of perimetric procedures in human subjects,[13–19] but investigators have only been able to explore outcome measures, such as test–retest variability, threshold comparison among perimetric procedures, and test time. One notable absence in these studies is the measure of accuracy, as the true underlying threshold for any given observer can only be estimated through repeat testing.

Here, we introduce a novel technique of obtaining an estimate of algorithmic accuracy in humans in addition to testing other performance characteristics of the GOANNA procedure. We call this technique Artificial Scotoma Generation (ASG). This was achieved by inducing a number of different scotoma types in healthy observers (arcuate, multiple defects, nasal step, and quadrantanopia). The defects were induced by inserting a module between the algorithm and the perimeter that altered presentation levels to simulate scotomata in human observers, which allowed us to have control over the size and depth of defect induced. By taking carefully measured baseline fields as an estimate of true normal sensitivity, we were then able to estimate test accuracy in human observers. A similar methodology has been adopted for evaluation of a computer-based self-administered visual field screening test for glaucoma (Tsamis E, et al. *IOVS* 2016;57:ARVO E-Abstract 2279).

We tested a small cohort of healthy observers intensively over three separate visits to explore accuracy and precision of the procedures. Although defect depth and size were simulated, we did not simulate the associated increase in variability with visual loss. In addition, there was no monitoring of head or eye movement other than direct observation by the examiner.

Methods

Participants

One eye of 15 healthy observers (mean age = 28 years, range = 21–45 years) was tested. We recruited both experienced ($n = 9$) and inexperienced ($n = 6$) psychophysical human observers. To be included in the study, observers were required to have normal visual field sensitivities (mean deviation > -2 dB), false positive rate lower than 20% on the Octopus 900 (Haag Streit AG, Koeniz, Switzerland), no history of diabetes or other systemic diseases, no evidence of ocular disease, trauma or surgery, and could not be taking medications known to affect visual field sensitivity or contrast sensitivity. All participants were prepresbyopic and wore a refractive correction with which they had Snellen visual acuity of 6/6 or better in the study eye. Full details of participant age and refractive error are given in Table 1. Before testing, all subjects provided written informed consent after explanation of the nature and possible consequences of the study. Data collection followed the tenets of the Declaration of Helsinki. Human research ethics approval for this study was obtained from The University of Melbourne Human Research Ethics Committee.

Scotoma Types

Four different scotoma types were investigated in our study: arcuate, multiple defects, nasal step, and quadrantanopia (Fig. 1). The arcuate, multiple defect, and nasal step fields were taken from an empirical 3° degree field dataset of glaucomatous patients and an age-appropriate correction of −0.1 dB per year was performed. This dataset was the same as the dataset of 23 input visual fields that were used in our earlier simulation experiments.[3,4] The quadrantanopic field was adapted from published literature.[20] The imple-

Table 1. Details of Each Participant's Age and Refractive Error

Participant	Age, y	Refractive Error, Diopters
A	27	+0.50 / −0.50 × 75
B	24	Plano
C	21	Plano / −0.25 × 90
D	45	−0.50 DS
E	28	−4.75 / −0.25 × 170
F	29	−2.25 DS
G	25	Plano
H	29	−4.75 DS
I	25	Plano / −0.50 × 90
J	21	−3.50 DS
K	30	−1.75 DS
L	29	−1.00 DS
M	29	Plano
N	28	+0.50 DS
O	29	+0.50 / −1.00 × 90

Refractive error is given in the form sphere (DS) / cylinder (DC) × axis (degrees).

mentation of GOANNA used herein, and in our previous work, enables the sampling of visual field at a resolution of 3°, hence the necessity for scotomata to be defined at 3° spacing.

Artificially Inducing Scotomata in Healthy Observers

We induced the various aforementioned visual field defects in healthy observers using ASG (Fig. 2). To create sensitivity loss at a particular location, the algorithm presented a dimmer stimulus than the intensity calculated to present. This in turn increased the probability that the observer would respond "no" to that stimulus and therefore the resultant sensitivity was reduced. To elaborate, let's say we want the final sensitivity of a given location to be 9 dB (see Fig. 2). However, the true sensitivity of the observer is 31 dB at that location. Therefore, for every stimulus presentation during the test, the intensity needs to be 22 dB dimmer than what would normally be presented. For example, if a stimulus was calculated to be presented at 25 dB, a 47 dB stimulus would be presented instead, which more than likely would result in the observer responding "no." Nevertheless, the algorithm still thinks that the observer responded "no" to a true 25 dB stimulus. This is how scotomata were artificially induced in healthy observers using the ASG method.

Test Procedures

To provide an indication of how the induced scotomata were characterized by current conventional automated perimetry, a Zippy Estimation of Sequential Testing (ZEST) procedure with a bimodal probability mass function (PMF) was implemented.[21–23] GOANNA was also implemented with the same bimodal PMF. As ZEST is based on maximum likelihood principles, it possesses similarities with the family of Swedish Interactive Threshold Algorithms (SITA), but is not as computationally involved. ZEST has been shown to outperform a SITA-like strategy and full threshold in terms of accuracy and precision, especially when the initial threshold estimate is far from true threshold.[22] We did not include SITA as a comparator procedure as the mechanics of SITA are not available in the public domain.

The selection of optimal parameters for GOANNA and ZEST were based on results from previous simulations,[3,4] and are described in detail below. Both test procedures were performed on an Octopus 900 perimeter (Haag Streit AG). The Octopus decibel scale is defined as $10 \times \log(1273 / L)$ where L is the stimulus luminance (cd/m^2). The background intensity was fixed at 10 cd/m^2 (31.4 apostilbs). The maximum stimulus luminance corresponding to 0 dB was set at 1273 cd/m^2 (4000 apostilbs). A Goldmann Size III target was used. A participant's false positive rate was measured by presenting a catch trial every 15 presentations.

ZEST Algorithm

Like GOANNA, ZEST was implemented using the open perimetry interface (OPI).[24] The ZEST procedure was based on a maximum-likelihood determination described elsewhere in the literature.[21,25] Variants of the ZEST procedure used in this study have been implemented in several commercially available machines such as the Medmont perimeter (Medmont Australia, Pty. Ltd., Camberwell, Australia) and the Humphrey Matrix (Carl Zeiss Meditec, Dublin, CA), although the specific details of the implementations differ between instruments. For our version, at each tested stimulus location, a PMF over the domain −5…40 dB is chosen as a prior, and the expected mean of the prior PMF is presented. This PMF defines, for all possible sensitivities, the probability that a given observer will have that sensitivity. Depending upon the observer's response, either a 'yes' or 'no' likelihood function is then multiplied with the prior PMF to generate a new PMF. The likelihood function represents the likelihood of an
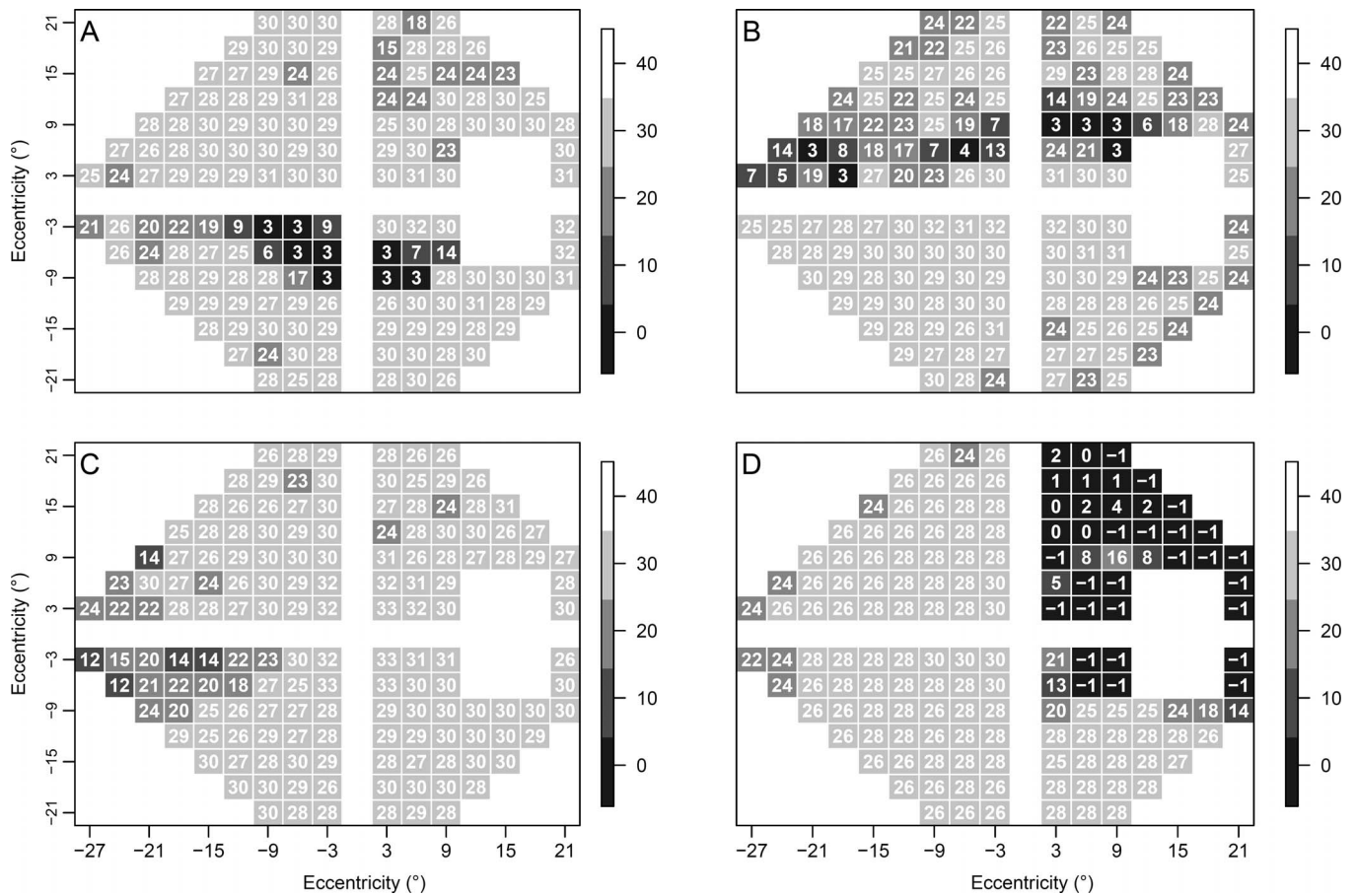
Figure 1. Threshold values of the four different visual field defects induced in healthy observers, shown both numerically and in a greyscale plot. Values describe the sensitivity in decibels at a given location. (A) Inferior arcuate; (B) multiple defects; (C) nasal step; (D) quadrantanopia. Sensitivities less than 0 dB indicate that the observer did not respond "yes" to a stimulus of 0 dB. Note: the spacing between points is 3°.

observer seeing a stimulus, and is the same function as used in previous studies.[3,4,22] The expected mean of the new PMF is then presented for the next response. This process continues until the termination criterion (standard deviation of the PMF is less than 1.5 dB) is met. The mean of final PMF upon termination is taken as the final output threshold.

A bimodal prior PMF is applied at each location, with one peak at 0 dB modelling thresholds of damaged locations, and a second peak at M modelling thresholds of healthy locations. Each location in the visual field requires an initial estimate of M. The approach of the Humphrey Field Analyzer 24-2 "growth pattern" for determining these initial estimates was adopted.[26] In our "growth pattern" implementation, four primary locations ($\pm 9°$, $\pm 9°$) have M set to 30 dB. Once the final thresholds were determined for these locations, their immediate neighbors set their M value to the value of the
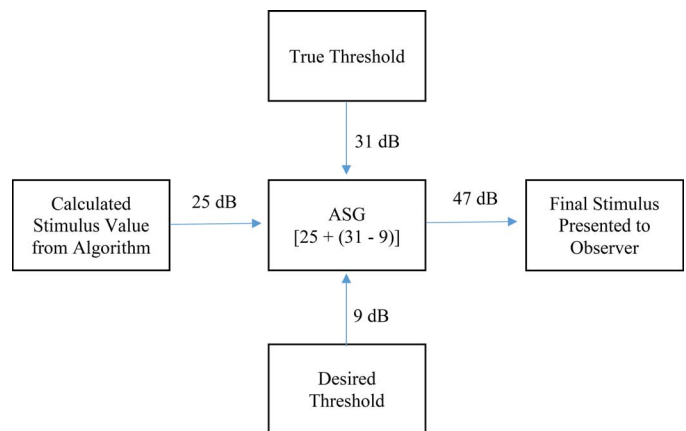

Figure 2. Flow diagram illustrating how scotomata are induced within the algorithm. In this case, the participant's true threshold is 31 dB, but we want the final threshold to be 9 dB. Therefore, a stimulus 22 dB dimmer than what would normally be presented is shown instead (47 dB in this example).
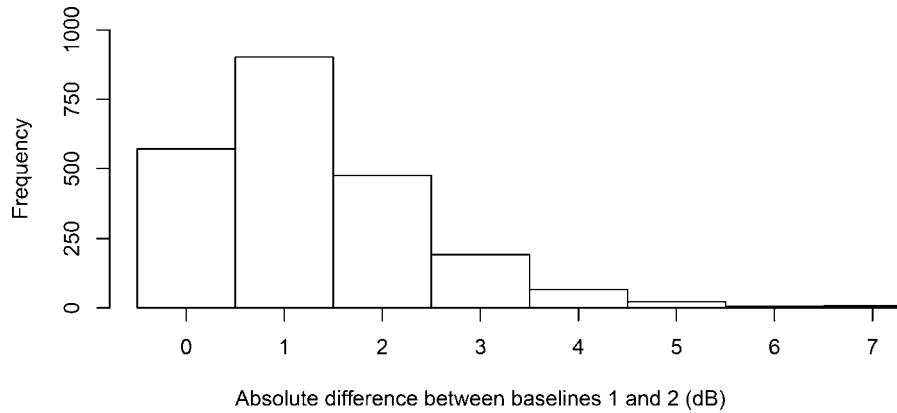
Figure 3. Histogram of the sensitivity differences between the two baseline fields for each location, with data pooled across all observers.

primary location. After these 24 locations terminated, their immediate neighbors take their M values as the mean of their neighbors, and so on. ZEST is run on 52 locations, with sensitivities of the remaining 98 locations obtained through natural neighbor interpolation.[6]

GOANNA Algorithm

The GOANNA procedure that was implemented in the current study has been described in detail previously.[3] In brief, there are 150 potential test locations, which GOANNA can select from. GOANNA interleaves both 24-2 and non 24-2 locations, but does not necessarily test all of these points. GOANNA first samples 36 predetermined seed locations until the standard deviation of their respective PMFs are less than 6 dB. From these seed locations, gradients (expected mean of probability mass function [dB] / distance between locations [degrees]) between all pairs of seed locations are calculated. The five location pairs, which correspond to the greatest gradients are then identified and a location pair out of these five is chosen at random. GOANNA then tests a location at the midpoint of this location pair. If the locations that form a chosen location pair are direct neighbors and are both unfinished, one of these locations is chosen at random to be tested. If one of the two locations in that pair is unfinished, the unterminated location is chosen. Until the termination criteria is met, the same process is used to select subsequent test locations, where gradients are calculated between locations that have already had at least one presentation. Upon termination of the procedure, threshold values for untested locations are obtained using natural neighbor interpolation.

Test Protocol

All participants underwent three testing sessions, each of approximately 1 hour in duration. The right eye of each patient was selected for testing. Two separate bimodal ZEST procedures on a 3° grid of 150 locations were undertaken in order to establish a baseline measure. The baseline field was taken as the mean of these two tests.

Once a baseline field was established, GOANNA and ZEST were performed three times each for each scotoma type, resulting in a total of 24 visual field test results per observer. Rest breaks were allowed during and between tests as required by the observers, and all observers were provided with the same instructions before initiation of testing. The order of testing was randomized and counterbalanced in order to minimize the possible confounding fatigue and learning effects. Any test that was deemed unreliable ($\geq$20% false positives) was discarded and repeated again.

Results

Baselines

A histogram illustrating the quality of the baselines can be seen in Figure 3. For each participant, the absolute threshold difference between the two baseline tests were calculated on a pointwise basis for each location in the visual field. A histogram was created from the pooled data.

Locations Tested

As the sampling grid for GOANNA varies from test to test depending on scotoma type and participant responses, we explored how many unique

Table 2. Description of the Unique Number of Locations Tested across Three Separate Visual Field Examinations for Each Defect Type, as Well as the Average Number of Locations That are Sampled at any Given Test for GOANNA

	Mean Number of Locations Tested at 1 of the 3 Visits	Mean Number of Locations Tested at 2 of the 3 Visits	Mean Number of Locations Tested at all 3 Visits	Mean Total Number of Locations Tested on a Given Visit
Arcuate	7	6	65	71
Multiple	12	12	49	61
Nasal step	12	8	55	64
Quadrantanopia	10	9	61	70

In comparison, ZEST tested the same 52 locations at all three visits. Results from all 15 observers were pooled.

locations GOANNA sampled from one visit to the next and investigated the number of locations that GOANNA samples for a given test (Table 2). Furthermore, we looked at whether the number of unique locations varied across scotoma type (Table 2). For each scotoma type, we pooled results across all 15 observers and calculated the mean number of locations tested.

For all defect types, GOANNA sampled significantly more locations than ZEST (paired t-test, $P <$ 0.001). On average, GOANNA sampled between 61 to 71 locations, depending on the scotoma type. Unlike ZEST, which samples the same locations each test, it can be seen that GOANNA did not sample the identical set of locations from one visit to the next.

Absolute Error and Presentation Number

Data was pooled across all observers and scotoma types to compare absolute error (| measured threshold – true threshold |) and total number of

presentations between GOANNA and ZEST (Fig. 4). There was no clinically significant difference in absolute error between GOANNA and ZEST. On average, GOANNA required more presentations to terminate compared with ZEST (283 vs. 258.5, Wilcoxon $P < 0.001$).

Absolute Error and Precision with Respect to Max_d

Because GOANNA exploits spatial information to improve characterization of scotomata, analysis that looks at spatial relationships is expected to reveal differences between GOANNA and ZEST. We examined absolute error and precision (interquartile range [IQR]) with respect to Max_d, which is defined as the largest difference in sensitivity (dB) between a location and any eight of its immediate neighbors. A location with a high Max_d corresponds to a scotoma edge. On the other hand, a location with a low Max_d corresponds to an area of uniform sensitivity. Three
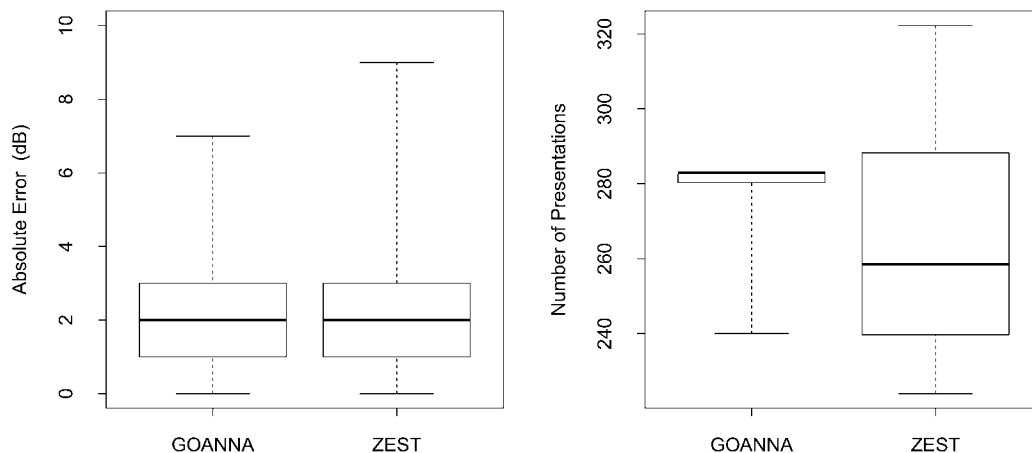

Figure 4. Boxplots of GOANNA and ZEST describing absolute error (*left plot*) and number of presentations (*right plot*). *Lower hinge*: 25th percentile; *upper hinge*: 75th percentile; *middle line in box*: median (50th percentile); *lower whisker*: 5th percentile; *upper whisker*: 95th percentile.
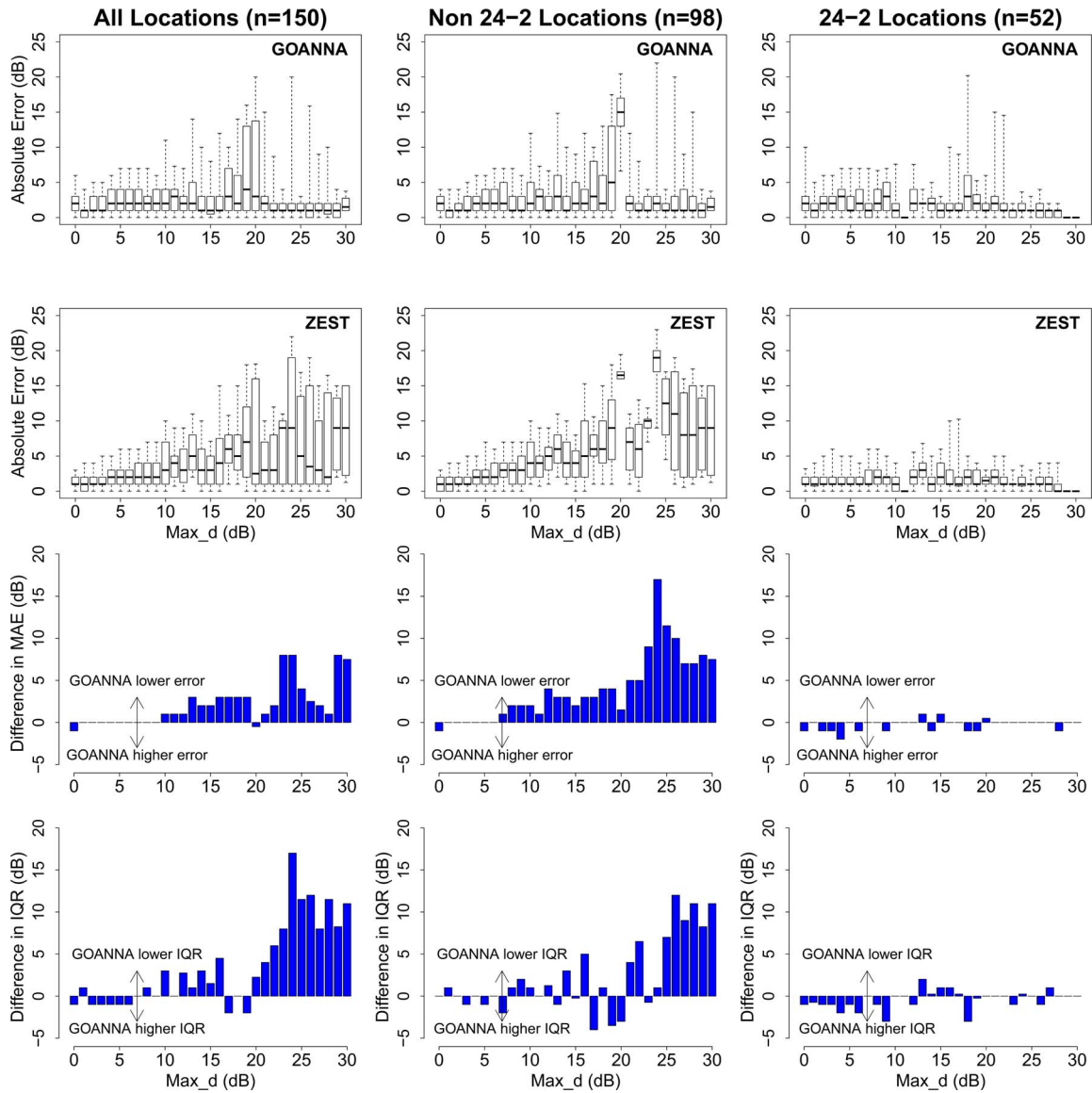
Figure 5. *Top*: boxplots of absolute error for GOANNA and ZEST with respect to *Max_d*. *Bottom*: difference plots of ZEST less GOANNA with respect to *Max_d*. Values in the positive range indicate where GOANNA performed better than ZEST. *Left column*: analysis performed on all 150 locations. *Middle column*: analysis performed on non 24-2 locations only, comparing GOANNA with interpolation from ZEST. *Right column*: analysis performed on 24-2 locations only.

location groups were studied: (1) all locations, (2) non 24-2 locations, and (3) 24-2 locations. We chose these groupings in order to analyze GOANNA's performance across locations that were measured versus interpolated in ZEST.

Boxplots in Figure 5 revealed that precision of ZEST decreased as *Max_d* increased (second row of Fig. 5). In contrast, GOANNA did not follow this trend (top row of Fig. 5). Difference plots showed that GOANNA displayed significantly greater accuracy than ZEST (Wilcoxon, $P < 0.001$) when *Max_d* was in the 10 to 30 dB range (with the exception of

Max_d = 20 dB; third row of Fig. 5). Outside this range, accuracy between GOANNA and ZEST was the same, with the exception at *Max_d* = 0 dB, where ZEST was more accurate by 1 dB. Similarly, GOANNA demonstrated greater precision than ZEST (Wilcoxon, $P < 0.001$) when *Max_d* was in the 20 to 30 dB range (bottom row of Fig. 5). Therefore, as we observed in the simulation experiments, GOANNA has a significant advantage over ZEST in regions that are near a scotoma edge.

From Figure 6, GOANNA demonstrated lower median absolute error than ZEST at scotoma edges
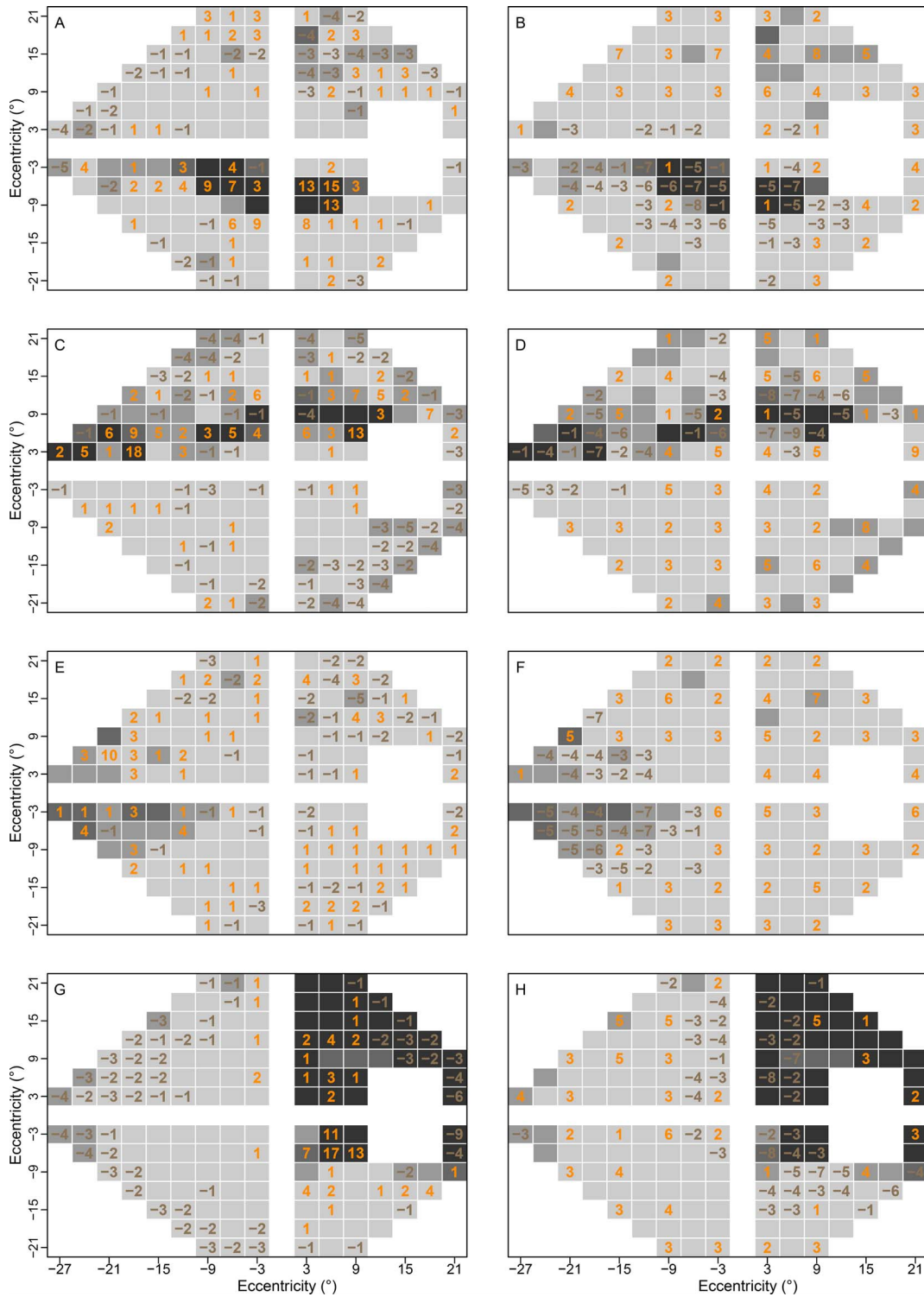
Figure 6. Difference plots outlining the difference in median absolute error (*left column*) and number of presentations (*right column*) between GOANNA and ZEST for the arcuate (A, B), multiple (C, D), nasal step (E, F), and quandrantanopic (G, H) defects. *Orange numbers* indicate values where GOANNA had lower error/presentations. Similarly, *brown numbers* indicate where ZEST had lower error/presentations. Locations without a number indicate that GOANNA and ZEST were equal at that location. *Grayscale plots* of true threshold are also plotted to provide spatial information about the underlying scotoma (see Fig. 1 for numerical values of true threshold).

and greater median absolute error in areas of uniform sensitivity. We also observed that GOANNA spent more presentations along scotoma edges and less presentations in parts of the hill of vision where sensitivity was uniform.

Discussion

Previously, we showed through computer simulation that exploiting spatial information to direct customized stimulus location choice leads to improved accuracy and precision around scotoma edges.[3] We also demonstrated that this improvement has the potential to translate to more accurate and earlier detection of glaucomatous progression.[4] Through this study of human observers, using ASG, we have now been able to replicate and validate the cross-sectional performance of GOANNA predicted from these previous simulation experiments.

An alternate approach that measures sensitivities at non 24-2 locations using gradient has been reported previously.[27] This method adds a fixed number of extra test locations to a completed 24-2 pattern by calculating the gradient between previously tested locations.[27] As a result, test duration is longer than testing the standard 24-2 grid,[27] whereas GOANNA does not increase test times.[3,4]

In Figure 4 it can be seen that the median number of presentations of GOANNA was greater than that of ZEST. Due to the differences in termination criteria between the two algorithms, GOANNA's median number of presentations was also equal to its maximum presentation number. On the other hand, ZEST's presentation number distribution was more evenly spread around its median. Therefore, it would be unfair to match the median number of presentations to that of ZEST as this would result in ZEST spending more presentations than GOANNA during half of all tests. Hence, setting the majority of GOANNA's presentations to equal less than the third quartile or ZEST's presentation distribution provided a fair compromise.

GOANNA devoted a large number of presentations to targeted areas of the visual field depending on where a given scotoma lies (right column of Fig. 6). These regions of localized loss are known to be highly variable.[2] ZEST on the other hand, has its presentations spread more evenly across the whole field.

Through this study, we were able to gain an understanding about the human factors and ergonomic aspects relating to GOANNA that we were unable to assess in previous computer simulation studies. We were also able to include any learning or fatigue effects experienced by human observers in the performance measures, which we previously did not simulate. Observers did not report any difficulty in undergoing testing with GOANNA and described it as comfortable to perform as standard visual field procedures used in clinic, but did report that they noticed a heavier concentration of presentations in particular areas of the visual field at times.

Several studies have suggested that it may not be necessary to measure sensitivities less than 15 dB, as responses become too unreliable beyond this lower limit.[28,29] The argument is that any apparent change in sensitivity within the range 0 to 15 dB may not be indicative of true progression of the disease state, but rather response noise.[29] If this is indeed true, there is potential to truncate the stimulus domain to 15...40 dB. This would allow GOANNA to either sample more locations per test or more thoroughly threshold the same number of locations, resulting in further improvements to performance.

One of the concerns with the GOANNA approach was that if GOANNA concentrated its presentations in one particular region of interest, it could potentially miss subtle scotomata if there were multiple defects present within a visual field. To test for this, the case of multiple scotomata was investigated. The results not only suggested that GOANNA was successful in identifying both scotomata, but it was able to better characterize these scotomata than ZEST (Fig. 6C, 6D).

As GOANNA only superficially samples a number of seed locations at the beginning of a test, false responses made early in an examination may misinform GOANNA, resulting in stimulus presentations being assigned to erroneous locations and potentially allowing subtle scotomata to go undetected. The relationship between location choice and false responses is an interesting factor to consider when developing next-generation customized procedures such as GOANNA, where locations are chosen automatically during the test. One possible strategy to mitigate the negative effect on performance due to false responses could be to assign more presentations to seed locations before proceeding to test nonseed points where gradients are largest. However, this would reduce the number of presentations that can be given to nonseed locations and therefore GOANNA would sample fewer locations overall. Finding a balance between thorough testing of a few locations versus superficial testing of many locations will be

important to explore during future development of GOANNA in the future.

Although the methodology of artificially inducing scotomata in healthy observers enabled estimation of accuracy, it was not able to completely mimic the behavior of people with disease due to the likely flattening of the frequency-of-seeing slope in those people.[30] As only healthy observers underwent testing, the frequency-of-seeing slope remained steep, even in areas of reduced sensitivity. One could argue that this could be solved by testing glaucomatous observers instead, but then a reliable measure of test accuracy would not have been able to be obtained. Another alternative is to add some random variation to the threshold chosen by ASG to simulate noise due to frequency-of-seeing slope (Henson DB, oral communication, 2015).

There were a number of assumptions that were made in this study. Firstly, in order to obtain an estimate of absolute error, we needed to assume that the baseline visual field was stable during the entire testing period. If there was a large degree of fluctuation between visits then the measurement of accuracy would in turn be affected. In order to minimize the effects of intertest fluctuation, observers were given ample breaks both between and during tests when needed. Additionally, the observers' head and eye position were closely monitored by the perimetrist using the in-built camera throughout testing.

In this study, we did not monitor eye movement with gaze or fundus tracking. Several studies have suggested that eye movements can influence test–retest variability and the slope of scotoma edges.[31–33] However, unlike these studies that involved participants with real visual field defects, the current study simulated defects in healthy individuals with smooth underlying visual fields. Therefore, in our study cohort, neighboring locations have similar thresholds to each other, and thus small eye movements would have minimal impact on the final outcome.

Furthermore, false positive catch trials were included in the procedures to obtain a measure of test reliability. The average false positive rate exhibited in this study was low (mean rate ± 95% confidence interval: 1.69% ± 0.41). We also assumed that the responses obtained from false positive catch trials was representative of response reliability during the entire course of a test. This assumption is one that is already widely accepted in current commercialized instruments.

The success of this experiment gives confidence that future experimentation involving a larger cohort of human observers over a larger number of visits would be worthwhile to pursue. A longitudinal study of glaucomatous patients to assess performance of GOANNA in detection of glaucomatous progression would be something to consider in future studies.

Acknowledgments

References

1. Maddess T. The influence of sampling errors on test-retest variability in perimetry. *Invest Ophthalmol Vis Sci*. 2011;52:1014–1022.
2. Wyatt HJ, Dul MW, Swanson WH. Variability of visual field measurements is correlated with the gradient of visual sensitivity. *Vision Res*. 2007;47: 925–936.
3. 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.
4. Chong LX, Turpin A, McKendrick AM. Targeted spatial sampling using GOANNA improves detection of visual field progression. *Ophthalmic Physiol Opt*. 2015;35:155–169.
5. Braun J, Sambridge M. A numerical method for solving partial differential equations on highly irregular evolving grids. *Nature*. 1995;376:655–660.
6. Sambridge M, Braun J, McQueen H. Geophysical parameterization and interpolation of irregular data using natural neighbours. *Geophys J Int*. 1995;122:837–857.
7. Fujimoto N, Adachi-Usami E. Fatigue effect within 10 degrees visual field in automated perimetry. *Ann Ophthalmol*. 1993;25:142–144.
8. Gonzalez de la Rosa M, Pareja A. Influence of the "fatigue effect" on the mean deviation

measurement in perimetry. *Eur J Ophthalmol.* 1997;7:29–34.

9. 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.

10. Lamparter J, Schulze A, Schuff AC, Berres M, Pfeiffer N, Hoffmann EM. Learning curve and fatigue effect of flicker defined form perimetry. *Am J Ophthalmol.* 2011;151:1057–1064, e1051.

11. Marra G, Flammer J. The learning and fatigue effect in automated perimetry. *Graefes Arch Clin Exp Ophthalmol.* 1991;229:501–504.

12. Wild JM, Searle AE, Dengler-Harles M, O'Neill EC. Long-term follow-up of baseline learning and fatigue effects in the automated perimetry of glaucoma and ocular hypertensive patients. *Acta Ophthalmol (Copenh).* 1991;69:210–216.

13. Morales J, Weitzman ML, Gonzalez de la Rosa M. Comparison between Tendency-Oriented Perimetry (TOP) and octopus threshold perimetry. *Ophthalmology.* 2000;107:134–142.

14. Springer C, Bultmann S, Volcker HE, Rohrschneider K. Fundus perimetry with the Micro Perimeter 1 in normal individuals: comparison with conventional threshold perimetry. *Ophthalmology.* 2005;112:848–854.

15. Spry PG, Johnson CA. Within-test variability of frequency-doubling perimetry using a 24-2 test pattern. *J Glaucoma.* 2002;11:315–320.

16. Su WW, Wu SC, Chang SH, Shen SC. Comparison of standard white-on-white automated perimetry and short-wavelength automated perimetry in early glaucoma patients. *Chang Gung Med J.* 2004;27:188–192.

17. 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.

18. Bengtsson B, Heijl A. Evaluation of a new perimetric threshold strategy, SITA, in patients with manifest and suspect glaucoma. *Acta Ophthalmol Scand.* 1998;76:268–272.

19. Bengtsson B, Heijl A, Olsson J. Evaluation of a new threshold visual field strategy, SITA, in normal subjects. Swedish Interactive Thresholding Algorithm. *Acta Ophthalmol Scand.* 1998;76:165–169.

20. Rowe F. *Visual Fields via the Visual Pathway.* 1 ed. Oxford, UK: Wiley-Blackwell; 2006.

21. Vingrys AJ, Pianta MJ. A new look at threshold estimation algorithms for automated static perimetry. *Optom Vis Sci.* 1999;76:588–595.

22. 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.

23. 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.

24. Turpin A, Artes PH, McKendrick AM. The Open Perimetry Interface: an enabling tool for clinical visual psychophysics. *J Vis.* 2012;12(11):pii 22.

25. 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.

26. Anderson DR, Patella VM. *Automated Static Perimetry. 2nd ed.* Philadelphia: Mosby; 1999.

27. Aoyama Y, Murata H, Tahara M, et al. A method to measure visual field sensitivity at the edges of glaucomatous scotomata. *Invest Ophthalmol Vis Sci.* 2014;55:2584–2591.

28. Gardiner SK, Demirel S, Goren D, Mansberger SL, Swanson WH. The effect of stimulus size on the reliable stimulus range of perimetry. *Transl Vis Sci Technol.* 2015;4:10.

29. Gardiner SK, Swanson WH, Goren D, Mansberger SL, Demirel S. Assessment of the reliability of standard automated perimetry in regions of glaucomatous damage. *Ophthalmology.* 2014;12:1359–1369.

30. 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.

31. Henson DB, Evans J, Chauhan BC, Lane C. Influence of fixation accuracy on threshold variability in patients with open angle glaucoma. *Invest Ophthalmol Vis Sci.* 1996;37:444–450.

32. Fendrich R, Wessinger CM, Marshall D, Johnson CA. Stabilized image perimetry: Evaluating the influence of eye movements on perimetry data. Vision science and its applications. *J Opt Soc Am Technical Digest Series.* 1995:254–257.

33. Demirel S, Johnson CA, Fendrich R, Vingrys AJ. The slope of frequency-of-seeing curves in normal amblyopic and pathologic vision. Vision science and its applications. *J Opt Soc Am Technical Digest Series.* 1997:244–247.

translational vision science & technology