tvst

DOI: 10.1167/tvst.5.3.3

Article

Development of Visual Field Screening Procedures: A Case Study of the Octopus Perimeter

Andrew Turpin[1], Jonathan S. Myers[2], and Allison M. McKendrick[3]

[1] Department of Computing and Information Systems, The University of Melbourne, Melbourne, Australia
[2] Wills Eye Hospital, Philadelphia, PA, USA
[3] Department of Optometry and Vision Science, The University of Melbourne, Melbourne, Australia

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

Received: 29 September 2015
Accepted: 22 January 2016
Published: 9 May 2016

Keywords: visual field; screening; glaucoma

Citation: Turpin A, Myers JS, McKendrick AM. Development of visual field screening procedures: a case study of the Octopus perimeter. Trans Vis Sci Tech. 2016;5(3):3, doi:10.1167/tvst.5.3.3

Purpose: We develop a methodology for designing perimetric screening procedures, using Octopus perimeters as a case study.

Methods: The process has three stages: analytically determining specificity and number of presentations required for different multisampling suprathreshold schemes at a single location of the visual field, ranking visual field locations by their positive predictive value (PPV) for glaucoma, and determining a pass/fail criteria for the test. For the case study the Octopus G-program visual field test pattern is used, and a dataset of 385 glaucoma and 86 normal patients.

Results: Using a 1-of-3 sampling strategy at a level equal to the 95 percentile of normal observers gave the most robust specificity under the influences of false-negative responses using an average of 1.5 presentations per location. The PPV analysis gave 19 locations that completely classified our glaucomatous data. A further 9 points were added to screen for nonglaucomatous loss. The final stage found that insisting that 3 locations are missed for the screening to fail gave a simulated specificity and sensitivity of approximately 95% for unreliable responders.

Conclusions: Our method gives a principled approach to choosing between the many parameters of a visual field screening procedure. We have developed a procedure for the Octopus that should terminate in less than 1 minute for normal observers with high specificity and sensitivity to glaucoma.

Translational Relevance: Visual field screening is used in community settings and eye care practice. This study provides a principled approach to the development of such screening procedures and details a new procedure.

Introduction

There is a long history of visual field screening in community settings and within clinical environments with the aim of rapidly measuring the visual field to detect visual disorder (for example, see prior reports[1–5]). Most studies that report on the effectiveness of visual field screening have used some form of computerized perimetry, although noncomputerized, manual methods also have been reported (for example, see report of Matsumoto et al.[6]). Screening procedures must try to achieve a difficult balance between test sensitivity (the proportion of truly abnormal visual fields that are detected by the procedure as being abnormal) and specificity (the proportion of truly normal fields that are classified by the procedure as normal), with test duration being minimized to enable rapid screening. We formalized a methodology for designing a computerized visual field screening procedure, and tested the predicted effectiveness of the screening strategy using computer simulation.

When faced with the task of developing a screening procedure, many choices may be made. Firstly, the hardware platform and test stimulus features (for example, white-on-white targets, size III) are chosen. Once these are selected, there are 3 main parameter sets to investigate: (1) the number and pattern of visual field locations to test, (2) the stimuli levels to show at each location, and (3) a rule to allow interpretation of the result as "pass" or "fail."

We present a methodology split into three experiments that set these parameters. We use the Octopus perimeters (Haag-Streit AG, Köniz, Switzerland) as the platform. The Octopus 600 and 900 use white-on-

white Goldmann Size III targets as in most current conventional perimeters. These three stages allow the methodical development of a computer simulation validated procedure that is ready for clinical trials. The methods described could be used more generally to develop visual field screening strategies where the design requirements may be different.

Methods

For a general screening procedure, we should make no assumptions about spatial relationships between locations that are tested in the patient's visual field. For example, we do not assume a glaucomatous or neurologic pattern of visual field damage: each location is tested independent of any other. We exploit this property by first experimenting with stimulus levels and sampling schemes at a single normal location to examine the trade-off between test time and specificity, particularly in the presence of false-negative responses. Once these parameters are set (Experiment 1), we can estimate the time required to test a single location, and then we can go about choosing the number and spatial distribution of test locations (Experiment 2). Finally, we can simulate the outcomes of the procedure on a database of fields and choose the number of locations that must fail the screening for the whole field to be labelled abnormal (Experiment 3). We describe each of these experiments in turn.

Experiment 1: Stimulus Level and Sampling Scheme

At each location, a small number of stimulus presentations must be made to allow the decision of "normal" or "damaged" to be derived. Further, the test must be suitable for participants who are new to perimetry, and who may only receive brief instructions (for example in rapid community screening programs). Suprathreshold targets are used for screening purposes, and their ease of identification in normal regions of visual field assists in patient instruction and compliance with the test. However, a key decision to be made is the suprathreshold level (in this case, the luminance of the Size III target).

The number of chances that an observer is given to respond to a stimulus affects the sensitivity–specificity trade-off of the test at that location. For example, consider a single location for a particular observer. If the stimuli level is set at a suprathreshold level that the observer can see 95% of the time and the stimulus was presented only once we would expect a specificity of 95% from that location (assuming no false-responses). If the stimulus was presented a second time, if not seen the first, then the chance of our observer not seeing both presentations is 0.25%, and specificity rises to 99.75%. If the observer has visual field damage at this location (that results in their true sensitivity being poorer than the 95% normal level), they will only respond "seen" as a false-positive. If there is a 15% chance of a false-positive response, then showing the stimuli once will give a sensitivity of 85%, and showing it twice gives the observer more chance of falsely pressing the button, and so sensitivity drops to 72.25%. Thus, presenting the stimuli more than once per location increases specificity, but decreases sensitivity (assuming false-positive responses). As a variant, we could present the stimulus 3 times, and require the observer to see it twice to be "normal." Again, this trades increased sensitivity against decreased specificity. Artes et al.[7] give a principled approach to examine these trade-offs, and for selecting a multisampling scheme. As we were aiming for high specificity in a screening procedure, we initially computed the expected specificity for the following four multisampling suprathreshold schemes from their selections at a single location: (1) 1-of-1, where 1 presentation must be seen for a "normal" classification; (2) 1-of-2, where 1 of 2 presentations must be seen for a "normal" classification; (3) 1-of-3, where 1 of 3 presentations must be seen for a "normal" classification; and (4) 2-of-3, where 2 of 3 presentations must be seen for a "normal" classification.

Note that some screening procedures have altered the stimulus level after the first presentation,[8] but we did not explore these variants. Rather, we kept the stimulus level constant at any given location set to either the population's 0.5%, 1%, 2%, or 5% sensitivity levels (brightest) of age-matched normal observers for that location. These values were taken from the normative database of the Octopus 600 as supplied by Haag-Streit AG. By keeping the stimuli level constant for all presentations at a location, we can compute the expected specificity and expected number of presentations at a single location, and be confident that it will apply to all locations in the visual field. If we used an adaptive level of stimuli, spatial logic between neighbors, or did not adjust stimuli levels for age and eccentricity, then we could not make such an extrapolation.

The combination of the four sampling methods and four stimulus intensities gives 16 different procedures to compare. For the moment, let us assume that we know the probability that a normal

Table 1. Expected Specificity and Number of Presentations at a Single Location Where the Probability of a Normal Observer Seeing Stimuli that is Presented According to the Sampling Scheme in the First Column is p

Sampling Scheme	Expected Specificity	Expected Number of Presentations
1-of-1	p	1
1-of-2	$p + (1 - p)p$	$p + 2(1 - p)$
1-of-3	$p + (1 - p)p + (1 - p)^2p$	$p + 2p(1 - p) + 3(1 - p)^2$
2-of-3	$p^2 + 2p^2(1 - p)$	$2p^2 + 3(1 - p)^2$

observer will say "yes" to a stimulus of s dB, say p. We can compute from this the expected specificity and number of presentations as in Table 1.

For perimetric stimuli, the probability of responding "seen" to stimulus at a level s, p, is usually modelled as a cumulative Gaussian frequency-of-seeing curve[9–12] as follows:

$$P(s,t,f_+,f_-) = f_+ + (1 - f_+ - f_-)\Big(1 - Gauss\big(s, \mu = t, \sigma = \exp(-0.078*t + 2.896)\big)\Big),$$

where f_+ and f_- are the probabilities of a false-positive and false-negative response respectively, *Gauss* is a cumulative Gaussian distribution with mean μ and standard deviation σ, t is the known threshold for the location, and the value of σ is taken from Table 1 of Henson et al.[9] altered for the Octopus dB scale. The Henson et al.[9] model was built assuming that 0 dB is 10,000 apostilbs, whereas the Octopus dB scale assumes a 4000 apostilb light for 0dB. In our simulations we do not allow σ to exceed 6 dB.

So now, to compute the expected specificity for a scheme over a population of normal observers we can take a normative database to give us $Q(t)$, the probability of a particular threshold occurring in the population at this location, and sum the product of that probability with the value from Table 1. For example, for the 1-of-2 scheme we would compute:

$$p(t) = P(s,t,f_+,f_-)$$
$$\text{Expected specificity} = \sum_{\forall t} \mathrm{Q(t)}\Big(p(t) + p(t)(1 - p(t))\Big), \quad (1)$$

where s, f_+, and f_- would be constants chosen for the simulation. In our case, s can be one of the 0.5%, 1%, 2%, or 5% levels taken from the normative database; f_+ will be zero, as any false-positives will only increase specificity calculations; and f_- will vary from 0% to 50% in steps of 5%. Using this approach we computed all of the expected specificities and number of presentations of the 16 schemes for a variety of false-negative conditions. This will allow us to choose the scheme with the best specificity-time trade-off for an individual location. The sensitivity of the entire test is explored in Experiment 3.

Experiment 2: Location Selection

The results of Experiment 1 will give us a stimuli level and sampling scheme that gives a good specificity-time trade-off for individual locations, particularly when false-negative lapses are high. In this experiment we decide on locations within the visual field at which the scheme will be applied.

Choosing locations to test for a visual field screening requires balancing test times with spatial coverage and likelihood to detect vision deficits. As visual field screening is often directed to glaucoma screening, a priority was choosing locations that are likely to be damaged by glaucoma. Wang and Henson[13] recently used a method to choose locations in the 24-2 pattern of the Humphrey Field Analyzer (HFA; Carl Zeiss Meditec, Inc., Dublin, CA) that held the highest diagnostic value in separating visual fields obtained from normal eyes from those with glaucoma. They did this by computing the positive predictive value (PPV) of each location in a database of 2344 normal visual fields and 2222 visual fields classified as having moderate glaucoma, taking out the highest scoring location, and repeating.

We adopted the same methodology for the G pattern, using a database of 385 glaucomatous and 86 eyes with normal visual fields collected on the Octopus 900 machine (Haag-Streit, Köniz, Switzerland) at Wills Eye Hospital. To be included in the databases, eyes were required to have visual acuity of better than 20/40, and for the visual fields to have reliability indices less than 30% (false-positives, false-negatives, fixation losses). Glaucoma eyes were required to have a diagnosis of primary open angle glaucoma. Both groups were required to have no other ocular disorder of significance. The mean age of the glaucoma eyes was 64.6 years (SD = 9.4 years), with an average visual field mean defect of −7.6 dB

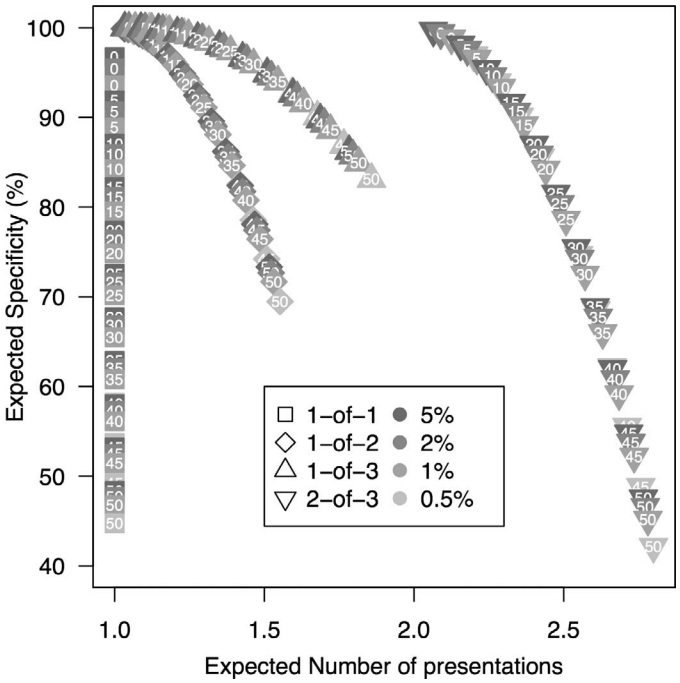

Figure 1. Expected specificity and number of presentations for the 4 different sampling schemes (shapes), 4 different stimuli levels *s* (colors), and 11 false-negative values f_- shown in *white text* inside each *symbol*.

(SD = 5.5 dB). The mean age of the normal eyes was 57.8 years (SD = 9.1 years) with an average visual field mean defect of 0.5 dB with a SD of 0.7 dB.

The locations chosen using this method are only relevant for detection of glaucoma, thus to ensure the screening procedure would detect vision deficits due to other causes, we added further locations manually, taking into account previous studies that have investigated the importance and information content of locations in the G and 24-2 patterns.[13–15]

Experiment 3: Sensitivity and Specificity

From Experiments 1 and 2 we have a test pattern and a scheme for deciding stimuli to present at the locations in the pattern. In this experiment we use standard perimetric computer simulation to examine the sensitivity and specificity of the test on two data sets. The final parameter to be decided in this experiment is the number of locations that must fail the screening for the whole test to be called "failed" for a given individual.

We used the *SimHenson* mode of the Open Perimetry Interface[16] to run our simulations. Briefly, each presentation of the test has a probability of responding as described by $P(s,t,f_+,f_-)$ used in Experiment One. For the initial simulations we used f_+ and f_- in the range 0% to 30%, a common cut-off value for false-positive and -negative responses as measured by catch-trials in the literature. To simulate a screening environment, where perimetrically naïve subjects are perhaps tested with minimal supervision, we increased these to 50% for some final tests.

The input thresholds, *t*, for the simulation were taken from two sources. We generated 500 normal fields directly from the Haag-Streit Octopus 600 normative data base, adding a random general height adjustment drawn from a Standard Normal distribution. We generated 163 glaucomatous fields were generated by converting Humphrey Field Analyzer (HFA; Carl Zeiss Meditec) 24-2 visual fields to the Octopus dB scale, and using natural neighbor interpolation to extract the G pattern. We have used this set of glaucomatous visual fields measured on the HFA previously in other simulation studies of visual field procedures (for example, see prior reports[17–19]). We also report the sensitivity and specificity on the Wills data set that was used in Experiment Two.

Results

Experiment 1: Stimuli Level and Sampling Scheme

Figure 1 shows the results of applying Equation 1 to compute the specificity-time trade off of the 16 choices of stimuli. It reveals several obvious trends. Firstly, increasing the false-negative rate (increasing white numbers within each symbol type) decreases specificity for all 4 sampling schemes, most dramatically for the 1-of-1 and 2-of-3 schemes (boxes and down-triangles). The 1-of-3 scheme (up-triangles) maintains the highest specificity as false-negative responses increase. A second observation is that the 2-of-3 scheme (down-triangles) requires approximately twice as many presentations as the other schemes, as expected. The third observation is that the choice of the level of stimuli (4 colors within each shape-number group) has little effect on the specificity and time for the test. Generally, the 5% level (darkest) has the highest specificity and shortest time in each group.

This leads us to conclude that we should use the 1-of-3 scheme at the 5% stimuli level.

Experiment 2: Test Pattern Selection

The left panel of Figure 2 shows the locations determined by the PPV analysis of the Wills data set (dark circles). The 19 locations that are shown enable

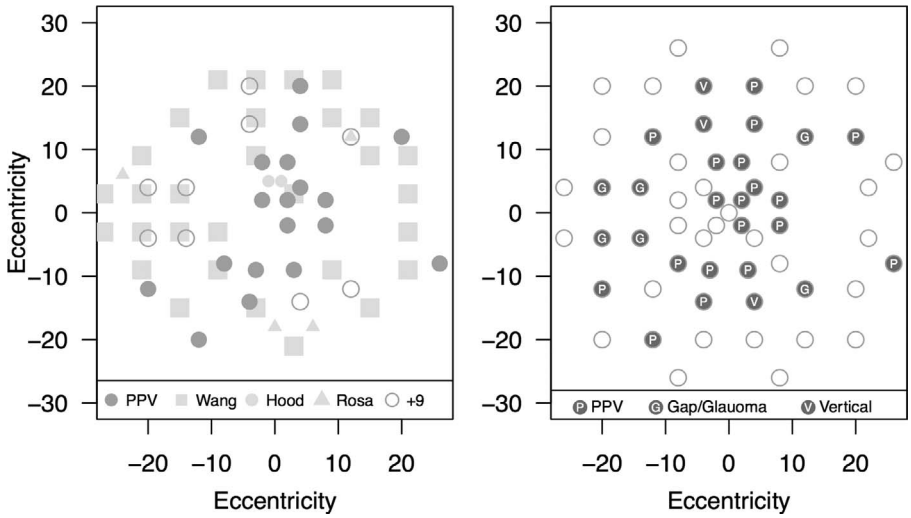

Figure 2. The *left* shows: the locations chosen using the PPV method on our data set (*dark circles*), the top 30 locations as published by Wang and Henson[13] using the same method on HFA 24-2 data (*squares*), the two locations recommended by Hood et al.[21] for detecting macular loss in glaucoma (*small circles*), the four locations advocated by Gonzalez de la Rosa et al.[14] as correlating with global measures of loss (*triangles*), and the 9 locations we added by hand to the PPV results (*open circles*). The *right* shows the final set of locations (*dark circles*) with *P* indicating the 19 PPV locations from the left panel. Locations tagged *G* and *V* are discussed in the text.

complete separation of the glaucomatous eyes from the normal eyes in this data set. Also shown in Figure 2 as squares are the top 30 locations found by Wang and Henson[13] using PPV on HFA 24-2 data. Interestingly, our locations are much more centrally located than the Wang locations. The Wills data used the G pattern, where 8 locations feature in the central 10°. For the Wang data only one location appears in the central 10°. This difference is most likely due to the test patterns used to collect the data, as the G pattern samples more densely in the macular region than the 24-2 pattern. Locations in the nasal-step region of the field (approximately −20, 0) did not feature in our PPV selections. A further observation is that Wang and Henson[13] and our selections included a location temporally close to the blind-spot. Temporal defects are associated with glaucoma but are much less common than nasal or arcuate scotomata.[20]

Also included in the left of Figure 2 are the 4 locations that Gonzalez de la Rosa et al.[14] found correlated with global defect measures (triangles). While these locations also appear near those suggested by the data of Wang and Henson,[13] they do not have a high PPV in our data set. A further two points are included in the panel as recommended by Hood et al.[21] and also confirmed by Chen et al.[22] for detecting macular changes in glaucoma (small circles). These are surrounded by locations with high PPV in our data set.

As can be seen, simply using these 19 locations as a test pattern for a general screening procedure would not be desirable. There are large areas of the nasal and inferotemporal field that are not tested at all. While it might be the case that the 19 locations are good for detecting glaucomatous field loss, other types of loss may be missed.

From Experiment 1 we know we should have less than 30 locations if we are to have a screening procedure that will take under 1 minute on a normal participant. Accordingly, we added 9 more locations manually to cover the missing regions. These locations are shown as the empty circles in the left of Figure 2, and the full test pattern is shown in the right. The right panel also labels the added locations as either "G," indicating that the location was chosen because it not only filled a gap in the field, but was likely to detect glaucomatous damage based on previous studies; or "V," indicating that the location was chosen to fill a gap and to detect defects that obey the vertical midline. As can be seen, the final pattern of 28 locations covers most areas within the central 20° of the G pattern.

Experiment 3: Sensitivity and Specificity

Using the 28 locations shown in Experiment 2, and the 1-of-3 sampling at the 5% limit of normal as the suprathreshold stimulus from Experiment 1, Figure 3 shows the sensitivity of the test on the interpolated HFA data and the Wills data sets, plotted against the specificity derived from the simulated normal dataset. Each series of colored dots depicts 961 different

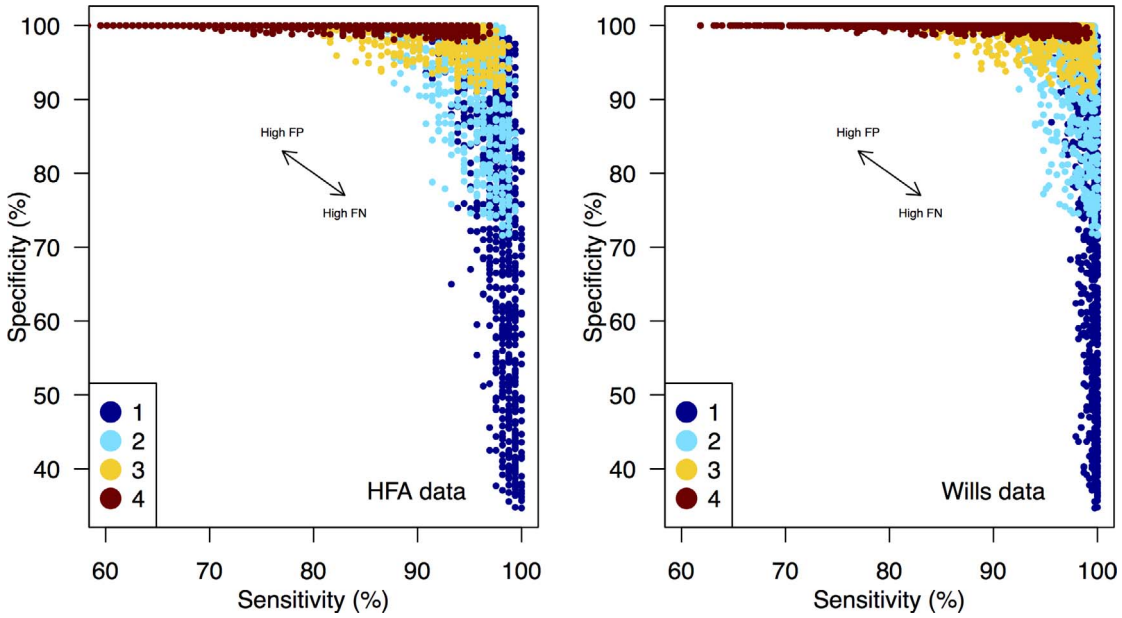

Figure 3. Sensitivity-specificity for the test when 1, 2, 3, or 4 locations are unseen (as indicated by the *colors* in the legend). Each *dot* is one of 961 possible combinations of false-positive and false-negative rates drawn from {0%, 1%, . . ., 30%}.

combinations of false-positive and -negative rates in the range 0% to 30%, and corresponds to a setting of the number of locations parameter. As can be seen, both data sets show the same pattern. The sensitivity for the Wills data is slightly higher than for the HFA data. The Wills dataset contains more locations with thresholds in the range 1 to 21 dB, thus the screening procedure is more likely to detect abnormal fields, because the datasets are not strictly matched for deficit depths. Note that this figure represents an average run of 100 simulations, hence due to stochastic variation any individual run may produce slightly different effects. Using only one location has high sensitivity (dark blue area) for all error conditions, but very low specificity when false-negative rates are high. Increasing to 2 locations (light blue area) increases the specificity for the high false-negative cases, but also gives decreased sensitivity for high false-positive rates. For a screening procedure, high specificity is desirable, and so increasing to requiring 3 locations (yellow area) gives even higher specificity in all cases, but now sensitivity can be as low as 70% for observers that make many false-positive responses. Increasing to requiring 4 presentations (brown area) guarantees even higher specificity, but allows sensitivity to drop to approximately 60%.

Table 2 shows the number of presentations required for the simulated normal, HFA, and Wills data sets for the 16 representative combinations of false-positive and negative rates. These numbers include an extra 2 presentations at the start of the test as practice trials, and a false-positive catch trial every 16 presentations. Also included are the sensitivity (Wills and HFA) and specificity values (Normals) for each data set.

Discussion

We present a three-step methodology for developing a perimetric screening procedure. While the specific experiments describe the development of a glaucoma screening strategy for the Octopus perimeter, the principles illustrated are more generally applicable.

While there are numerous reports in the literature of the clinical performance of visual field screening strategies (for example, see prior studies[1,3,5]), in-detail descriptions of the development decisions underpinning screening tests are fewer. Clearly, it is essential to test performance in a clinical setting; however, computer simulation studies allow the performance of a test procedure to be studied exhaustively, enabling robust selection of the test parameters and algorithmic choices beforre clinical evaluation. A clear advantage of computer simulation is that it allows the experimenter to know the "true" sensitivity of the simulated observer. This is never the case when testing humans as all measured thresholds are

Table 2. Number of Presentations Required, Specificity, and Sensitivity for the Final Procedure on the Three Data Sets

		Simulated Normals			HFA			Wills		
f_+	f_-	Mean	SD	Spec	Mean	SD	Sens	Mean	SD	Sens
0%	0%	29.1	1.2	100%	62.2	14.2	91%	56.2	14.2	88%
0%	3%	30.0	1.5	100%	62.3	14.1	93%	56.6	14.3	88%
0%	15%	34.0	2.6	100%	64.7	12.8	94%	59.2	13.2	92%
0%	30%	39.9	3.5	91%	67.4	11.7	97%	62.5	11.8	95%
3%	0%	29.1	1.2	100%	60.3	13.8	91%	55.1	13.7	86%
3%	3%	30.0	1.5	100%	61.0	13.9	92%	55.6	13.6	87%
3%	15%	34.1	2.5	100%	62.9	12.6	94%	57.9	12.6	90%
3%	30%	39.9	3.6	93%	65.8	10.4	97%	61.2	11.2	96%
15%	0%	29.0	1.0	100%	54.6	11.5	87%	50.3	11.6	76%
15%	3%	29.8	1.4	100%	55.5	11.0	84%	50.7	11.3	77%
15%	15%	33.8	2.5	100%	57.6	10.4	92%	53.0	10.4	84%
15%	30%	39.8	3.6	93%	60.0	8.8	94%	56.5	9.2	89%
30%	0%	28.8	0.9	100%	48.7	9.2	71%	45.0	9.1	59%
30%	3%	29.6	1.3	100%	49.0	9.0	71%	45.6	8.9	58%
30%	15%	33.6	2.4	100%	50.5	8.1	75%	47.6	7.9	63%
30%	30%	39.5	3.4	94%	53.2	6.9	86%	50.9	7.0	76%

estimates and, therefore, are susceptible to variability, response errors, and any biases in the algorithms used to estimate the thresholds. Of course, simulation studies have their own limitations, specifically the various assumptions regarding the performance of the simulated observers. In the case of perimetry using size III white-on-white targets, however, there is a wealth of empirical data documenting typical response errors, frequency of seeing curves, and typical thresholds in those with normal and abnormal vision that enable reasonable assumptions to be made for the purposes of the simulations.

Indeed, assessing the success of a screening procedure in a clinical or community setting is nontrivial and depends on the choice of "gold-standard" against which the screening test results are compared. Screening tests compare to some normative limits, hence the relevance of the normative population in the database to the population being screened is key to the outcomes. In this study, the normative data used were from the database of the O600 perimeter which included participants free from eye disease, with refractive error between +6 and −6 diopters (D) and with no systemic health conditions known to affect visual function or likely to impact on perimetric performance. We used a different database of normal and glaucomatous visual fields to assess and tailor the performance of various possible variants of the screening procedure. We considered these populations to be fairly representative of most people with normal vision and those with glaucoma; however, the performance of the screening procedure in any given population will depend on the validity of the norms. We also assumed significant numbers of response errors might occur in untrained populations; however, with careful instruction and the very short test procedure, we expect these should be reduced in practice

In this study, we chose to reduce the number of visual field locations from the more standard G (Octopus perimeter) or 24-2 (Humphrey Field Analyzer) test, to a subset of locations deemed most likely to carry unique information about the likelihood of visual field abnormality. The purpose of the screening procedure is to identify an abnormality, not to characterize the spatial nature of the abnormality, hence the decision to reduce test-time by testing fewer locations. Our experiments evaluating the performance of the test procedures were biased towards the detection of glaucoma; however, it should be noted that additional test locations were added to the locus of test points to ensure that visual field defects typical of cortical origin, and those due to common retinal abnormalities, such as macular degeneration, should also be identified. Nevertheless, it should be noted that all visual field test strategies only sample a tiny

minority of the total spatial area of the visual field so will always struggle to identify small areas of localized loss.

In summary, we described a framework for the development of a screening visual field test procedure and highlighted many of the factors that must be considered in such development. The test used a multisampling approach that is predicted to take less than 1 minute to complete in individuals with normal vision, and tests spatial locations likely to carry significant information content regarding the likelihood of disease. Further evaluation of performance in a variety of screening settings is now warranted.

Acknowledgments

The authors thank Chris Johnson for providing the 24-2 Humphrey Field Analyzer data.

Supported by Haag-Streit AG (AT, AMM, JSM). The authors alone are responsible for the content and writing of this paper.

Disclosure: **A. Turpin**, Haag-Striet AG (F); **J.S. Myers**, Haag-Striet AG (F); **A.M. McKendrick**, Haag-Striet AG (F)

References

1. Mansberger SL, Edmunds B, Johnson CA, Kent KJ, Cioffi GA. Community visual field screening: prevalence of follow-up and factors associated with follow-up of participants with abnormal frequency doubling perimetry technology results. *Ophthalmic Epidemiol*. 2007;14:134–140.
2. de Vries MM, Stoutenbeek R, Muskens RP, Jansonius NM. Glaucoma screening during regular optician visits: the feasibility and specificity of screening in real life. *Acta Ophthalmol*. 2012;90:115–121.
3. Yamada N, Chen PP, Mills RP, et al. Screening for glaucoma with frequency-doubling technology and Damato campimetry. *Arch Ophthalmol*. 1999;117:1479–1484.
4. Lee AJ, Wang JJ, Rochtchina E, Healey P, Chia EM, Mitchell P. Patterns of glaucomatous visual field defects in an older population: the Blue Mountains Eye Study. *Clin Exp Ophthalmol*. 2003;31:331–335.
5. Katz J, Tielsch JM, Quigley HA, Javitt J, Witt K, Sommer A. Automated suprathreshold screening for glaucoma: the Baltimore Eye Survey. *Invest Ophthalmol Vis Sci*. 1993;34:3271–3277.
6. Matsumoto C, Eura M, Okuyama S, et al. CLOCK CHART (r): a novel multi-stimulus self-check visual field screener. *Jpn J Ophthalmol*. 2015;59:187–193.
7. 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.
8. Spry PGD, Hussain HM, Sparrow JM. Performance of the 24-2-5 frequency doubling technology screening test: a prospective case study. *Br J Ophthalmol*. 2007;91:11345–11349.
9. 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.
10. Spry PGD, Johnson CA, McKendrick AM, Turpin A. Variability components of standard automated perimetry and frequency doubling technology perimetry. *Invest Ophthalmol Vis Sci*. 2001;42:1404–1410.
11. Wall M, Maw RJ, Stanek KE, Chauhan BC. The psychometric function and reaction times of automated perimetry in normal and abnormal areas of the visual field in patients with glaucoma. *Invest Ophthalmol Vis Sci*. 1996;37:878–885.
12. 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.
13. Wang Y, Henson DB. Diagnostic performance of visual field test using subsets of the 24-2 pattern for early glaucomatous field loss. *Invest Ophthalmol Vis Sci*. 2013;54:756–761.
14. Gonzalez de la Rosa M, Reyes JA, Gonzalez Sierra MA. Rapid assessment of the visual field in glaucoma using an analysis based on multiple correlations. *Graefes Arch Clin Exp Ophthalmol*. 1990;228:387–391.
15. Zeyen TG, Zulauf M, Caprioli J. Priority of test locations for automated perimetry in glaucoma. *Ophthalmology*. 1993;100:518–522.
16. Turpin A, Artes PH, McKendrick AM. The Open Perimetry Interface: an enabling tool for clinical visual psychophysics. *J Vis*. 2012;12:1–10.
17. Chong LX, McKendrick AM, Ganeshrao SB, Turpin A. Customised, automated stimulus loca-

tion choice for assessment of visual field defects. *Invest Ophthalmol Vis Sci*. 2014;55:3265–3274.
18. Turpin A, Jankovic D, McKendrick AM. Retesting visual fields: utilising prior information to decrease test-retest variability in glaucoma. *Invest Ophthalmol Vis Sci*. 2007;48:1627–1634.
19. McKendrick AM, Turpin A. Combining perimetric supra-threshold and threshold procedures to reduce measurement variability in areas of visual field loss. *Optom Vis Sci*. 2005;82:43–51.
20. Schiefer U, Papageorgiou E, Sample PA, et al. Spatial pattern of glaucomatous visual field loss obtained with regionally condensed stimulus arrangements. *Invest Ophthalmol Vis Sci*. 2010; 51:5685–5689.
21. Hood DC, Nguyen M, Ehrlich AC, et al. A test of a model of glaucomatous damage of the macula with high-density perimetry: Implications for the locations of visual field test points. *Transl Vis Sci Technol*. 2014;3:5, collection 2014.
22. Chen S, McKendrick AM, Turpin A. Choosing two points to add to the 24-2 pattern to better describe macular visual field damage due to glaucoma. *Brit J Ophthal*. 2015;99:1236–1239.