OPO OPHTHALMIC & PHYSIOLOGICAL OPTICS
THE JOURNAL OF THE COLLEGE OF OPTOMETRISTS

Ophthalmic & Physiological Optics ISSN 0275-5408

Comparison of defect depths for sinusoidal and circular perimetric stimuli in patients with glaucoma

William H Swanson and Brett J King

Indiana University School of Optometry, Bloomington, USA

Citation information: Swanson WH & King BJ. Comparison of defect depths for sinusoidal and circular perimetric stimuli in patients with glaucoma. *Ophthalmic Physiol Opt* 2019; 39: 26–36. https://doi.org/10.1111/opo.12598

Keywords: perimetry, glaucoma, retinal nerve fibre layer, sinusoids

Correspondence: William H Swanson
E-mail address: wilswans@indiana.edu

Received: 13 June 2018; Accepted: 19 November 2018

Abstract

Purpose: Clinical use of perimetric testing in patients with glaucoma typically assumes that perimetric defects will be less deep for larger than smaller stimuli. However, studies have shown that very large sinusoidal stimuli can yield similar defects as small circular stimuli. In order to provide guidelines for new perimetric stimuli, we tested patients with glaucoma using five different stimuli and compared defects to their patterns of retinal nerve fibre layer (RNFL) damage.

Methods: Twenty subjects with glaucoma were imaged with optical coherence tomography (OCT) volume scans to allow for en face RNFL images and were also tested on a custom perimetry station with five stimuli: Goldmann sizes III and V, a two-dimensional Gaussian blob (standard deviation 0.5°) and a 0.5 cycle degree^{-1} sinusoidal grating presented two ways: flickered at 5 Hz, and pulsed for 200 ms instead of flickered. En face RNFL images were reviewed with the visual field locations overlaid, and each location was labelled for a patient as either no visible RNFL defect or as wedge, slit, edge, or diffuse defect. Nineteen age-similar controls were tested with the same stimuli to define depth of defect as difference from mean normal. Bland-Altman analysis was used to test three predictions of neural modelling by making five comparisons.

Results: Bland-Altman analysis confirmed the three predictions. The flickered sinusoid gave deeper defects in damaged areas than the pulsed sinusoid ($r = 0.25$, $p < 0.0001$). When comparing data for sizes III and V there was increased spread of the data in deeper defects in the direction of size III having deeper defect ($r = 0.35$, $p < 0.0001$). The size V stimulus yielded shallower defects than a stimulus of similar size but with blurred edges ($r = 0.20$, $p = 0.0004$).

Conclusions: On average, all stimuli produced similar results comparing across type of RNFL damage. However, there were systematic patterns consistent with predictions of neural modelling: in damaged areas, depth of defect tended to be greater for the flickered sinusoid than the pulsed sinusoid, greater for the size III stimulus than the size V stimulus, and greater for the Gaussian blob than for the size V stimulus.

Introduction

Diagnosis of glaucoma and assessment of progression involve analysis of both perimetric and imaging measures, but these measures can have substantial discordance.[1,2] One potential factor in this discordance is that conventional automated static perimetry for patients with glaucoma tests each retinal area with a stimulus that covers < 0.5% of the area. Furthermore, for such a small stimulus, fixational instability can cause it to fall on non-overlapping locations on different trials.[3,4]

The common 24–2 test pattern samples a grid of stimulus locations with spacing of 6° horizontal and vertical, using the Goldmann[5] size III stimulus, a circular luminance

increment with area of 0.15 deg^2. If ganglion cell damage is inhomogeneous in the retinal region serving this 36 deg^2 region of the visual field, then the small retinal region that the stimulus falls on may have greater or lesser damage than the average damage for the retinal region sampled.[3,6] Use of additional locations can improve detection of early perimetric defects,[7–9] but this can greatly increase testing time.[10] An alternate approach for testing more locations is to perform suprathreshold testing,[11] such as Rarebit perimetry[12] which is designed for detection of visual field damage but not for assessing progression.

One could seek to test a larger retinal area by using a larger circular stimulus, but a range of studies has found that smaller circular stimuli can give deeper defects in early glaucomatous damage.[13–17] Clinical thought has been that glaucomatous perimetric defects will be less deep for larger stimuli[18,19] and that the largest Goldmann stimulus (size V, area of 2.3 deg^2), can be useful for detecting residual function in visual field locations where the patient cannot see the size III stimulus.[6,20] However, test-retest variability is greater for smaller stimuli, so larger stimuli may have greater sensitivity to early defect despite giving milder defect depths.[14] We demonstrated that these findings are consistent with effects of ganglion cell loss on cortical pooling of retino-thalamic spike trains,[3,21–23] and have proposed that appropriate sinusoidal stimuli can allow cortical pooling across many more retinal ganglion cells than circular stimuli.[22,24] For instance, Matrix 24–2 perimetry uses a very large flickered sinusoidal stimulus (25 deg^2), and on average has yielded similar defect depths to size III in mild to moderate defects.[25,26]

A collaboration with Hao Sun[27] compared macular contrast thresholds of patients with glaucoma for the 4 deg^2 macular Matrix stimulus and the much smaller size III stimulus, and she concluded that the ganglion cells mediating detection showed response saturation for the size III stimulus at contrasts above 100% (equivalent to 25 dB on the Humphrey Field Analyser). She confirmed this prediction by recording spike trains of primate ganglion cells responding to the size III stimulus at different contrasts.[28] We confirmed her prediction that test-retest variability in glaucomatous defects would be lower for appropriate sinusoidal stimuli than for the size III stimulus.[29] Collaborations with Stuart Gardiner confirmed the predicted effects of response saturation on the frequency-of-seeing curves from patients with glaucoma, for both the size III stimulus and the 16-fold larger Goldmann size V stimulus.[30,31] In summary, our analysis concludes that cortical pooling is limited to a small retinal area for circular stimuli, but not for appropriate sinusoids such as the Matrix stimuli. The clinically-relevant inference is that, on average, defect depths will not be as great for the size V stimulus as for the size III stimulus, but defect depths for sinusoidal stimuli will on average be similar to defect depths for the size III stimulus (until response saturation comes into play).

The 5° × 5° Matrix 24–2 stimulus covers 69% of the retinal area sampled, so has the potential to provide a more accurate assessment of the amount of ganglion cell damage in that region than the size III stimulus. However, the Matrix uses rapid flicker (12–25 Hz), which means that contrast threshold will be elevated when retinal illuminance is reduced (such as from small pupils).[32–34] We have found that use of 5 Hz flicker overcomes this problem,[35] so the current study used 5 Hz flicker with the Matrix sinusoid. To assess the proposal that 5 Hz flicker yields a greater cortical pooling area than the 200 millisecond temporal pulses used in conventional automated perimetry, we compared depth of defect when the sinusoid was flickered and pulsed. We also compared depth of defect for three circular stimuli, to assess predictions of which stimulus properties are important for the benefits of cortical pooling across large retinal areas.

Methods

Participants

The research for this study adhered to the tenets of the Declaration of Helsinki and was approved by the institutional review board at Indiana University. Informed consent was obtained from each participant after explanation of the procedures and goals of the study, before testing began.

Twenty patients with glaucoma and 20 age-similar controls were enrolled from our clinic population; all had prior experience with perimetry. Details of inclusion and exclusion criteria have been published elsewhere,[26] with the additional requirement for patients in the current study that there be a repeatable nasal visual field defect and corresponding thin sector of the retinal nerve fibre layer (RNFL). Briefly, all subjects were required to have clear ocular media, corrected monocular distance visual acuity of at least 20/20 (20/40 over age 70), refractive corrections between +2 and −6 dioptre spherical equivalent and cylindrical correction within ±3.0 dioptres. Control subjects were required to have had a comprehensive eye exam (not more than 2 years before the study) finding normal retinal characteristics and visual acuity of at least 20/20 (6/6, 0.0 logMAR) on each study visit. After recruitment and testing, one of the controls was diagnosed as a glaucoma suspect and was removed from the study, leaving 19 controls. Patients with glaucoma were required to be under the care of an eye care practitioner and to have had a recent clinical examination finding normal retinal characteristics except for retinal disc changes and perimetric changes associated with glaucoma. For the patients with glaucoma, mean deviation (MD) ranged from −16 dB to +1 dB (median

−4 dB, interquartile range −5 dB to −1 dB), PSD ranged from 1.7 dB to 11.9 dB (median 5.8 dB, interquartile range 2.5 dB to 9.2 dB). Subjects with ocular or systemic disease (other than glaucoma) currently affecting visual function were excluded from this study. Eyes with epiretinal membranes were also excluded from this study because the membranes can make retinal measurements unreliable.

The age range was 56–84 years for the 20 patients, and 49–89 years for the 19 controls. The mean (S.D.) age was 70 (7) years for the patients and 67 (10) for the controls, so we considered these two groups to be age-similar. Because we compared mean defect across stimuli, and the groups were age-similar, we did not use age corrections in computing depth of defect.

Equipment

A custom perimetric testing station used a cathode-ray tube display driven by a ViSaGe visual stimulus generator (www.crsltd.com/tools-for-vision-science/visual-stimulation/visage) that provided a resolution of 800 × 600 pixels with 14-bit control of luminance for each pixel; details are given elsewhere.[29] Four different spatial patterns, shown in *Figure 1*, were used as stimuli: the size III stimulus as a circular luminance increment with diameter 0.4°; the size V stimulus as a circular luminance increment with diameter 1.7°; a Gaussian blob[36] with standard deviation 0.5°; and a 5° × 5° vertical 0.5 cycle degree^{-1} sinusoid.[25] Each of these four spatial patterns was presented with a 200 ms rectangular temporal pulse. The sinusoid was also presented with counterphase temporal flicker at 5 Hz for 600 ms, making five stimulus conditions in all.

A Spectralis (www.heidelbergengineering.com) was used to image the RNFL. Each B-scan was averaged over nine frames. For each eye, dense (30 μm spacing) vertical scans for 4–6 fixation locations were gathered to cover a retinal region approximating most of that tested with 24–2 perimetry, then SLO images for these scans were montaged using a custom MATLAB (www.mathworks.com/) program that operated i2K Retina montaging software (www.dualalign.com/). Vertical scans were used because this improves imaging in the temporal raphe where fibres transition from horizontal to arcuate. Details have been published elsewhere.[11,37,38]

Protocol

Each subject made three one-hour visits, at ~1-month intervals, and on each visit five perimetric tests were performed on one eye per subject, with the order of the five tests counter-balanced across visits and subjects. Each test measured contrast thresholds for a single stimulus condition at 18 locations in the nasal visual field, as shown in *Figure 2*. We tested locations in the nasal visual field that correspond to the temporal raphe because this is the region in the retina where it can be determined where individual nerve fibre bundles begin.[39] Details of the ZEST algorithm that was used are given elsewhere.[29]

The five tests were conducted after checking acuity and Pelli-Robson contrast sensitivity, and were followed by 4–6

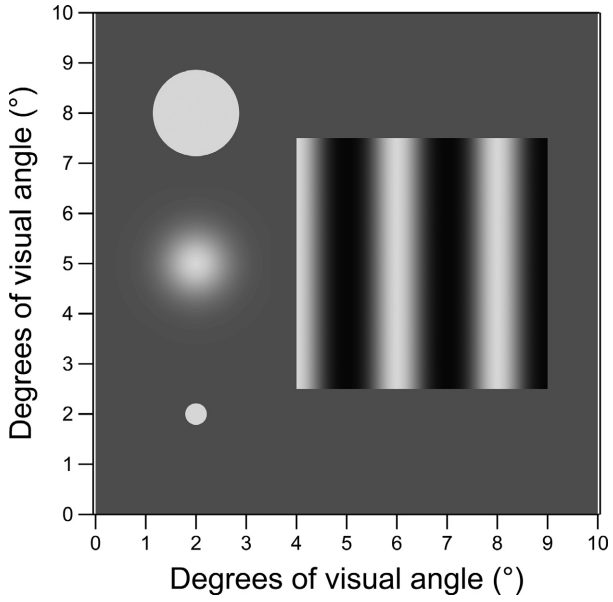

Figure 1. Spatial patterns used for perimetric stimuli: the size V stimulus (upper left), the Gaussian blob (middle left), the size III stimulus (lower left) and the 0.5 cycle degree^{-1} sinusoid (right).

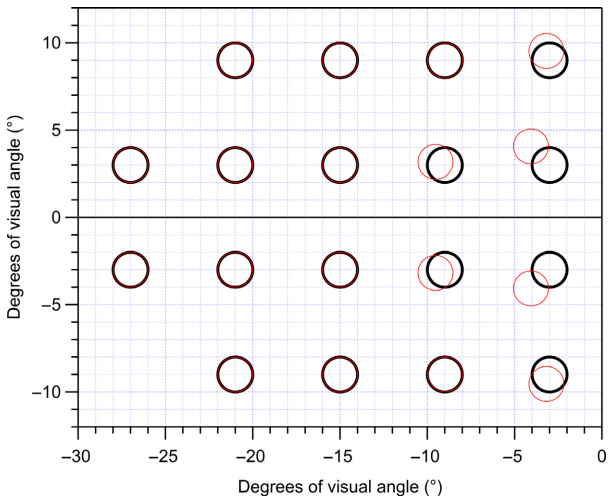

Figure 2. Perimetric locations tested, in right eye format. Thick black circles show visual field coordinates and thin red circles show effects of ganglion cell displacement. The circles indicate visual field location, not stimulus size.

volume scans for RNFL montages. If 1 h was not enough to complete the imaging, it was completed on the second visit. Repeat volume scans were gathered on the second and third visits when possible, and montages were made from the best sets of volume scans.

Data analysis

Mean log contrast threshold for each stimulus condition and location was computed from the data from the 19 controls, and patient data were converted to 'depth of defect' by subtracting their log contrast threshold from this mean value. Following Hao Sun,[27] we defined contrast for all stimuli as (peak luminance-background luminance)/(background luminance). This is equal to Weber contrast for a luminance increment and is equal to Michelson contrast for a sinusoidal grating. For each subject, contrast thresholds for each stimulus condition and location were averaged across the three sessions. To assess potential learning effects, for each stimulus we compared contrast thresholds on the first visit with the average from visits 2 and 3. The mean learning effects were small, ranging from 0.01 to 0.06 log unit as compared to test-retest standard deviations of 0.20 to 0.22 log unit, so we decided that learning effects were negligible.

En face RNFL montages were reviewed with the 18 visual field locations overlaid allowing for Henle fibre layer displacement[40] (red circles in *Figure 2*). Each location was independently labelled for a patient by a clinician (BJK) as either no visible RNFL defect ('normal') or defects listed as slit, edge, wedge, or diffuse. Examples are show in *Figure 3*. A 'slit' location falls within a very thin arcuate defect, just a thin arcuate reflectance defect slightly larger than the size III stimulus. A 'wedge' location is in a more severe form of damage that falls within a wide arcuate defect, which has sharp edges. A 'diffuse' location is in an area of widespread RNFL damage. An 'edge' pattern is normal-appearing RNFL at the edge of an RNFL reflectance defect, near the edge of a slit or a wedge. The clinician was masked to the perimetric data, and developed these classifications from reviewing the RNFL images. To assess internal consistency, a few weeks after the clinician initially scored the locations, he and his co-author sat down together and re-scored them without referring to the original scoring. Only 6% of locations had a different score, most changed from 'normal' to 'diffuse.' The second scores were used for the analysis.

Bland-Altman analysis[41,42] was used to compare defect depths across tests. Some tests had larger ranges between mean normal and maximum contrast, so when comparing depth of defect across tests a floor was used to equate possible ranges for defect depths.[29] For Bland-Altman analysis comparing depth of defect across two tests, the test with the lowest mean normal value was used to set a floor so that both tests had the same possible range of defect depths at all 18 locations.[26,29]

For the Bland-Altman analysis, the focus was on the predictions of our modelling for cortical pooling of ganglion cell responses.[3,21–23,27,28] This modelling predicts that cortical pooling will cover a larger retinal region for the flickered stimulus than for the pulsed stimulus, so when there are

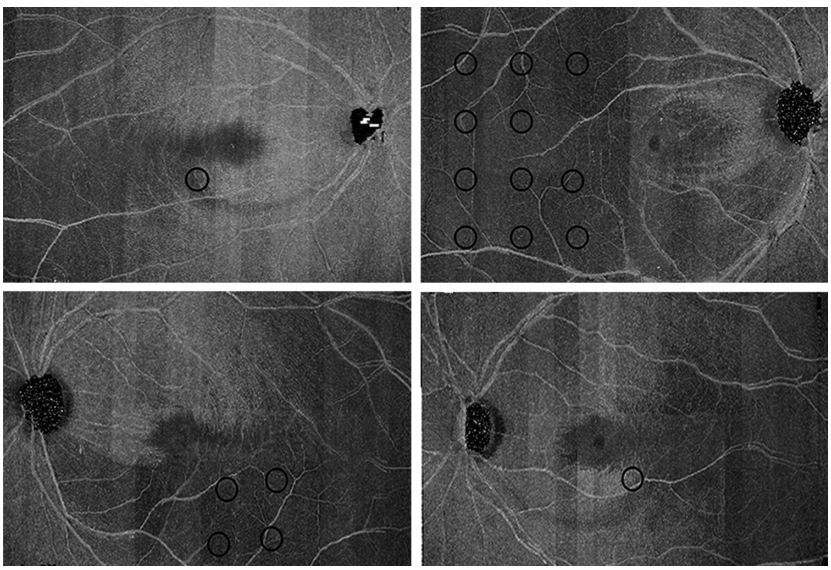

Figure 3. RNFL reflectance maps at 24 μm below the inner limiting membrane, illustrating perimetric locations corresponding to the four classes of retinal nerve fibre layer (RNFL) defects. Black circles show perimetric locations in these classes: slit (upper left), wedge (lower left), diffuse (upper right), edge (lower right). The circles indicate visual field location, not stimulus size.

residual patches of ganglion cells the pulsed stimulus will more often have shallower defects; this was assessed by analysis of agreement on depth of defect for the flickered sinusoid versus the pulsed sinusoid. This modelling also predicts that when one stimulus is much smaller than the other then there will be increased spread of the data in deeper defects, in the direction of the smaller stimulus having deeper defect. This was assessed two ways. First, by analysing agreement on defect depths for the flickered sinusoid versus the size III stimulus, for the flickered sinusoid versus the size V stimulus, and for the size III stimulus versus the size V stimulus. Second, we assessed spread of the data by linear regression of the absolute value of the difference in depth of defect as a function of the mean value for the defect depth. Finally, the modelling predicts that the circular size V stimulus with sharp edges will stimulate cortical mechanisms that pool ganglion cell responses over a region smaller than the stimulus (because the size V stimulus is much larger than Ricco's area in the central visual field, meaning that cortical processes mediating detection have a cortical pooling area much smaller than the stimulus[19,21,43]), and that in damaged areas it will yield shallower defects than stimuli of similar size but with blurred edges. This prediction was assessed by comparing defect depths for the size V stimulus and the Gaussian blob. This yielded 10 tests of statistical significance, for which we used a Bonferroni correction and required $p < 0.005$ for statistical significance.

Once these primary results were obtained, a secondary analysis explored whether asymmetry analysis could improve agreement. For each of the nine stimulus locations in superior visual field, asymmetry analysis computed difference in log contrast threshold for the mirror location across the horizontal midline, excluding all cases where the stimulus was not seen in one or both locations. Bland-Altman analysis was performed to determine whether asymmetry analysis reduced the width of the limits of agreement.

Another secondary analysis explored the extent to which test-retest variability contributed to the limits of agreement. The widths of the limits of agreement for test-retest variability were compared to those for between-stimulus agreement when just the third visit was used.

Sex as a biological variable

In compliance with National Institutes of Health (NIH) guidelines for reporting sex as a biological variable, we assessed sex differences, even though the study was not designed to look for them. For each control we computed mean threshold across all tests and locations, then compared these means for women and men. For patients with glaucoma, we compared frequency of the five different types of RNFL defect in women and men, counting the number of locations with each type.

Data sharing

In compliance with NIH and Indiana University policies and to protect the confidentiality of our human subject data and protected health information (PHI), Indiana University School of Optometry shares research data in the form of a limited data set pursuant to an approved data use agreement. Data and computer code used in this project will be shared with any research team whose institution executes an approved data use agreement with Indiana University.

Results

As shown in *Figure 4* and *Table 1*, across all subjects and locations the depth of defect was usually similar for all five stimulus conditions. However, for every patient at least one comparison had a difference by at least 0.28 log unit for at

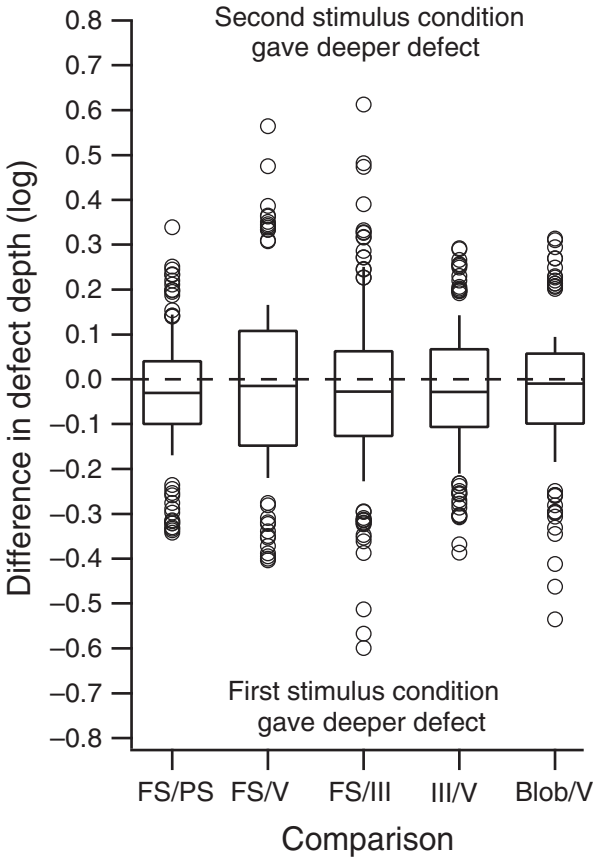

Figure 4. Box and whisker plots for the differences in depth of defect for pairs of the five stimuli: flickered sinusoid (FS), pulsed sinusoid (PS), sizes III and V, and the Gaussian blob.

Table 1. Summary of five comparisons of depth of defect for flickered sinusoid (FS), pulsed sinusoid (PS), the size III stimulus, the size V stimulus, and the Gaussian blob

Comparisons	FS/PS	FS/V	FS/III	III/V	Blob/V
Locations where at least one stimulus yielded defect depth above the floor, as percent of all locations where at least one stimulus yielded defect depth above the floor					
Y deeper	26%	30%	33%	27%	25%
X deeper	17%	19%	30%	16%	9%
Within ±0.1 log unit	57%	51%	37%	56%	65%
Locations where at least one stimulus yielded defect depth below the floor, as percent of all 360 locations					
Both not seen	7%	4%	16%	18%	11%
Y only not seen	2%	0%	3%	4%	2%
X only not seen	0%	3%	4%	1%	1%

X and Y axes are as in *Figure 6.*

least one location. *Table* 2 shows results of Bland-Altman analysis, which found weak but statistically significant correlations between difference and mean for three comparisons (flickered sinusoid versus pulsed sinusoid, size III versus size V, Gaussian blob versus size V) and correlations that did not reach our criterion for statistical significance for two comparisons (flickered sinusoid versus size V and versus size III). The limits of agreement ranged from ±0.22 to ±0.36 log unit. The analysis of spread of the data found that for all comparisons there was increased spread of the data in deeper defects, reaching statistical significance except for the flickered sinusoid versus the size III stimulus.

For the RNFL reflectance map labelling of the 360 retinal locations, 198 were scored as 'normal'; 102 as 'diffuse; 28 as 'wedge'; 18 as 'edge'; and 14 as 'slit'. *Figure 5* shows mean defect depth for locations assigned to the different types of RNFL defects, and *Figure 6* shows scatterplots for the five comparisons with a colour code to indicate type of RNFL defect. For clinical devices, a defect depth of −0.1 log unit corresponds to −1 dB for the size III stimulus and −2 dB for the Matrix stimulus. For both types of clinical devices, a defect depth of −0.5 log unit occurs in individuals free of disease less than 1% of the time.

The exploratory study on asymmetry analysis found that asymmetry analysis did not reduce the limits of agreement for any comparison. The exploratory study on test-retest variability found that the limits of agreement were ±0.40 to ±0.44 log unit across the five stimulus conditions, similar to the limits of agreement of ±0.33 to ±0.48 log unit for the comparisons of depth of defect for the third visit.

The assessment of sex as a biological variable found that the 14 female controls had 0.16 log unit lower mean log contrast threshold than the five male controls, and that the female patients had more locations with normal RNFL appearance and fewer with diffuse RNFL damage (*Table* 3).

Discussion

This study was conducted in the context of renewed research interest in revising perimetric stimuli. For circular stimuli as used in conventional perimetry, defects tend to be deeper with smaller stimuli.[13–16,18,19] However, two studies (one from our lab) found that large flickered sinusoids can yield similar defect depths as the much smaller size III stimulus, with increased spread of the data in deeper defects.[25,26] Our cortical pooling analysis[3,21–23,27,28] explains this by proposing that for appropriately chosen sinusoids there will be extensive cortical pooling of ganglion cell responses, so the defect depth for the 25 deg^2 sinusoid reflects the average amount of damage to much of the 36 deg^2 retinal region, while the much smaller size III stimulus will reflect damage to less than 1% of this region. We tested three predictions of the modelling, and performed a qualitative evaluation of effects of extent of ganglion cell damage based on en face RNFL images.

The first prediction was that when one stimulus was much smaller than the other then there would be increased spread of the data in deeper defects, and that the spread of the data would be in the direction of the smaller stimulus having deeper defect. This prediction was based on two

Table 2. Results of analysis of agreement on depth of defect and spread of the data for the five comparisons

	FS vs PS	FS vs V	FS vs III	III vs V	Blob vs V
r	0.25	−0.11	−0.08	0.35	0.20
p	<0.0001*	0.04	0.15	<0.0001*	0.0004*
Intercept	0.01	−0.04	−0.02	0.02	−0.01
Slope	0.12	−0.06	−0.07	0.23	0.09
LoA ±	0.25	0.35	0.36	0.26	0.22
Spread of the data	−0.18	−0.34	−0.15	−0.27	−0.26
p	0.0008*	<0.0001*	0.0096	<0.0001*	<0.0001*

Shown are: Pearson's *r* for difference against mean, the associated probability value (*p*), intercept and slope for the regression line, 1.96 times the standard deviation of the residuals (LoA±), Pearson's *r* for the absolute value of the difference versus the mean (spread of the data), and the corresponding *p* value. An asterisk (*) indicates a *p*-value that met our requirement for statistical significance.

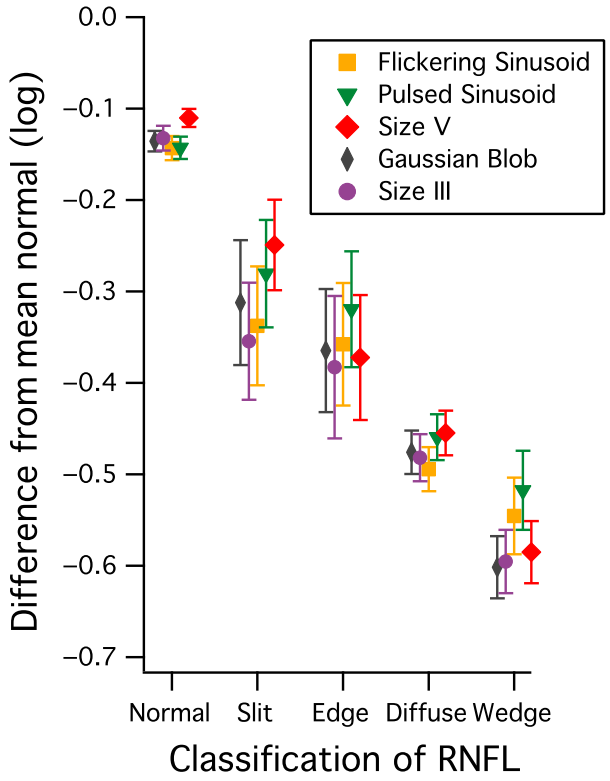

Figure 5. Mean defect depths for the five classes of retinal nerve fibre layer (RNFL) defect and the five stimuli, for a floor of −0.7 log unit. Error bars show standard error of the mean.

factors: heterogeneous damage and response saturation. Heterogeneous damage is expected when many but not all ganglion cells in a region have died, so that a small stimulus may not stimulate any remaining ganglion cells because it does not fall on their receptive field, or it may stimulate the normal number of ganglion cells if it happens to land on the receptive fields of a patch of remaining cells.[3,22] The spread of the data represents this variability. For large flickered sinusoids, we proposed that extensive cortical pooling averages responses of a large number of ganglion cells, and confirmed a reduction in test-retest variability.[29] The second factor, response saturation, comes into play because in healthy eyes it seems that at threshold for the size III stimulus a small number of cells each contributes ~1 stimulus-related spike and all spikes are in close enough temporal proximity that they summate at the cortex.[23,28] When half of these ganglion cells have died, the other ganglion cells need to contribute ~2 stimulus-related spikes. Response saturation for the size III stimulus means that doubling the stimulus contrast will not double the number of stimulus-related spikes, so the increase in contrast threshold would be more than two-fold. For large sinusoidal stimuli, at threshold each cell on average contributes much less than one stimulus-related spike and a doubling of contrast is expected to cause a doubling of stimulus-related spikes. Therefore, in more damaged areas we expect that the smaller stimuli will tend to give deeper defects due to effects of response saturation being greater for smaller stimuli.

The prediction of increased spread of the data was confirmed, in that the test for spread of the data (bottom of *Table* 2) found that the magnitude of difference in defect depth was greater in deeper defects for the size V stimulus versus the size III stimulus, and for the flickered sinusoid versus the size V stimulus and versus the size III stimulus; for the flickered sinusoid versus the size III stimulus this did not reach our criterion for statistical significance. The prediction that the spread of the data would be in the direction of the smaller stimulus was confirmed in that Bland-Altman analysis found slopes in the direction of the smaller stimulus having a deeper defect, although this only reached statistical significance for the size III stimulus versus the size V stimulus. This is consistent with the finding of Artes *et al.*[25] that in more damaged regions the size III stimulus tended to give deeper defects than the flickered sinusoid but some patients repeatedly showed deeper defects with the flickered sinusoid. Our analysis[21] concluded that there is insufficient cortical pooling for the size V stimulus to have this advantage. This conclusion is consistent with our finding that the strongest correlation was for the size III stimulus and the size V stimulus.

The increased spread of the data in comparing stimulus conditions at locations with perimetric defects confirms the finding reported by Artes *et al.*[25] that depth of defect can be similar across stimulus conditions in areas of mild to moderate loss with spread of the data in areas with deeper defect. This appears not to be limited to only the comparison of defects with the size III stimulus and Matrix 24–2 stimulus, as we found it with all four large stimulus conditions.

The second prediction was that, in damaged areas, the flickered sinusoid would give deeper defects than the pulsed sinusoid. This prediction is based on the well-known effect that appropriate use of flicker can increase contrast threshold for higher spatial frequencies,[44,45] which implies increased cortical pooling by psychophysical mechanisms mediating detection of low spatial frequencies.[21] The Bland-Altman analysis comparing the flickered sinusoid and the pulsed sinusoid supported this, because the slope of the difference in depth of defect versus the mean was positive, which is reflected in *Figure 6* as an excess of points falling below the diagonal in the left half of the graph.

The third prediction was that the circular size V stimulus with sharp edges stimulates cortical mechanisms that pool ganglion cell responses over a region smaller than the stimulus, so in damaged areas the size V stimulus will yield shallower defects than a stimulus of similar size, but blurred edges. This was supported by a positive slope for the Bland-Altman analysis comparing defect depths for the Gaussian blob and the size V stimulus, and is reflected in

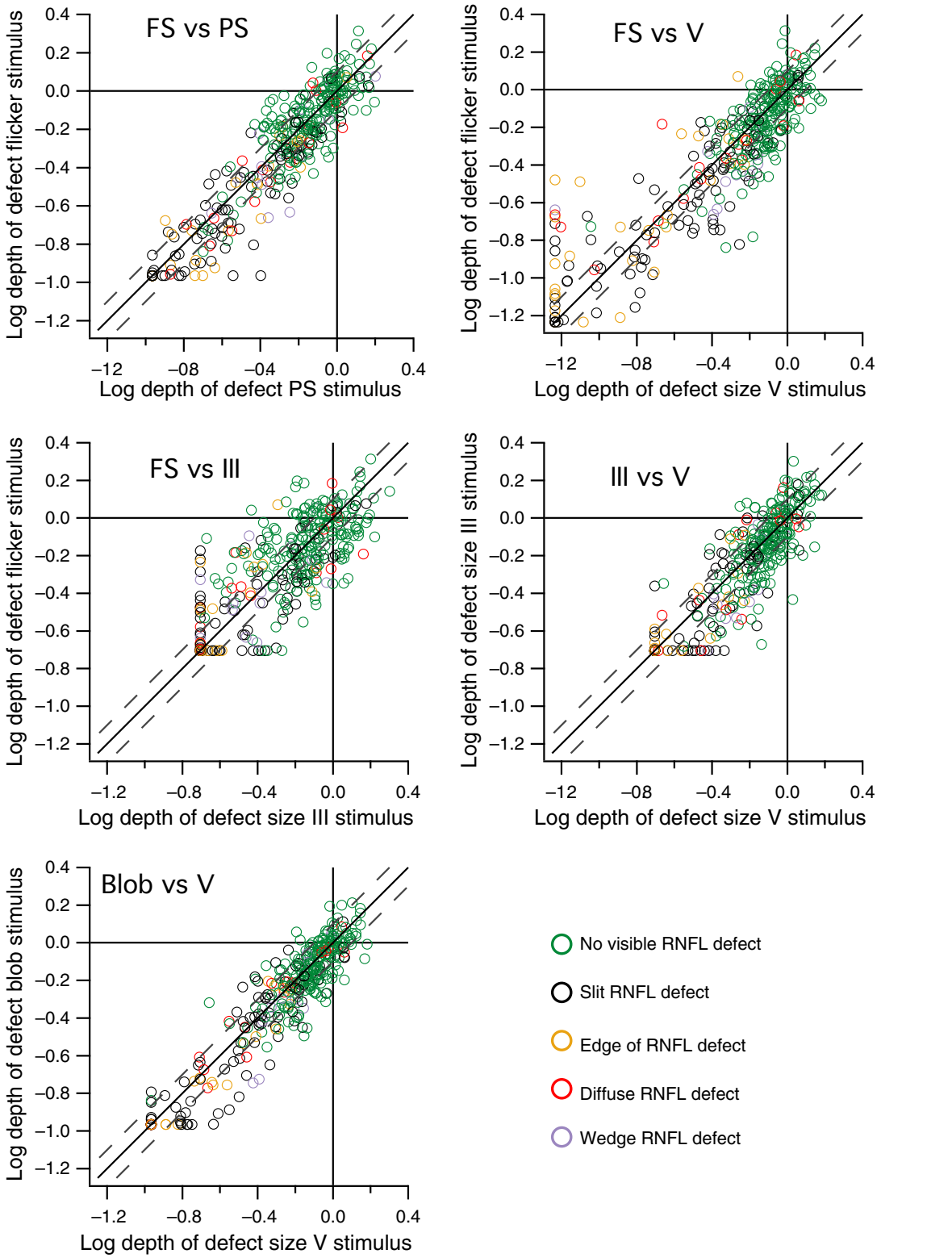

Figure 6. Scatterplots comparing depth of defect for different stimuli, averaged across three separate visits. Solid diagonal lines show equal defect depth, dashed lines show ± 0.1 log unit. The floor for depth of defect varied across comparisons, giving an appearance of truncation especially for the size III stimulus.

Figure 6 as an excess of points falling below the diagonal in the left half of the graph.

On average, all stimulus conditions produced similar results comparing across type of RNFL damage (*Figure 5*). For the locations labelled 'normal,' on average a mild perimetric abnormality was noted for all five stimulus conditions. A 'slit' defect was small compared to the size of the sinusoid, but the average defects were very similar for the flickered sinusoid and the size III stimulus. An 'edge' location appears normal, but because the average defect was similar as for slit defects, we infer that there is often ganglion cell damage. The variability across locations was greater for these two patterns than the other three, which we interpret as variability in the extent of damage to normal-appearing tissue. On average, the 'diffuse' and 'wedge' patterns tended to yield the deepest perimetric defects.

Our neural modelling predicts that, on average, all five stimulus conditions will give similar depths of defect when ganglion cell damage is relatively homogeneous in a retinal region, and this is seen in *Figure 5* for the 'normal' and 'diffuse' classes where all five mean defects are within 0.05 log unit of each other. Indeed, for the majority of the comparisons depth of defect was within 0.1 log unit for any two stimulus conditions, yet in some instances the differences were substantial (*Figure 4*). In particular, one-third of locations had deeper defect for the flickered sinusoid than for the size III stimulus. Given that test-retest variability was also substantial, we sought independent evidence from other studies that these patients had participated in. Twelve patients had at least 10 Matrix 24–2 fields, as well as at least 20 Humphrey Field Analyser (HFA) 24–2 fields with the size III stimulus, so we computed difference in depth of defect using the total deviation (TD) value provided by the devices, converting these to log units in order to have a common scale,[27] and used the same floor of −0.7 log unit. For each of the 18 visual field locations in each patient we performed a T-test between the defect depths for the repeated Matrix fields and the repeated HFA fields. Out of 216 locations, 26 had T-values lower than −3 (indicating that Matrix 24–2 defects were deeper), so we compared these to our results. We found that for 19 of these locations we also found deeper defects with the flickered sinusoid than with the size III stimulus. In the remaining five locations, we found deeper defects with the size III stimulus than with the flickered sinusoid; these locations were from two people with slit defects.

Table 3. Distribution of types of RNFL damage at the 18 locations in the patients with glaucoma, divided by sex

	Diffuse	Edge	Normal	Slit	Wedge
Female	15%	1%	69%	3%	11%
Male	37%	7%	46%	4%	6%

It is common to make age-corrections when computing perimetric defect depth, with each location in the visual field having a different age slope.[46] With only 19 controls, we did not have enough data to make reliable estimates of 90 age slopes. We have reported that sinusoids can have shallower age slopes than for the size III stimulus,[47] so it is possible that differences in age slopes contributed to spread of the data. Analysing age effects for published data[26] for the size III stimulus and a 5 Hz flickering sinusoid, we found that the median age slope was −0.10 log unit per decade for the size III stimulus and −0.03 log unit per decade for 5 Hz flickering sinusoid. This would produce deeper defects for size III in the older patients if the age-corrected defects were identical; for the oldest patient in the current study, this would produce a difference in depth of defect by 0.13 log unit. To assess the possibility that lack of age-correction had a substantial contribution to the spread of the data, we looked at the ages of the patients who yielded the 51 data points with at least an 0.3 log unit difference in depth of defect for the size III stimulus and the flickering sinusoid. Eight patients yielded 35 locations where size III yielded deeper defects, and six patients yielded 16 locations where the flickering stimulus yielded deeper defects, with two patients being in both groups. The mean age was 66 years (range 56–78 years) for the group with deeper defects for the size III stimulus, and was 74 years (range 63–84 years) for the group with deeper defects for the flickering stimulus. This is the opposite of what would be expected if lack of age-correction had a substantial contribution to the spread of the data.

We performed a comparison of sex as a biological variable to comply with NIH reporting standards, not to test any predictions. The purpose of these reporting standards is to accumulate findings in the literature that, if consistent across studies, can be used in a premise for a hypothesis-based grant proposal. For the controls, mean contrast threshold across all locations and subjects was 0.16 log unit lower for the women than the men. To assess whether this finding was repeatable, we compared mean contrast thresholds of 37 women (ages 47–77, mean 61 years) and 25 men (ages 50–85, mean 65 years) from a published[26] study, and found that the mean was only 0.02 log unit lower for the women than the men, so the finding was not confirmed and should not be considered support for a premise.

In summary, we used perimetric stimuli ranging in size from 0.15 deg^2 to 25 deg^2 and found that on average they all gave similar depth of defect, yet in a number of patients there were large and repeatable differences in depth of defect. We confirmed predictions of our neural modelling concerning when such differences would occur, and found

instances where the type of en face RNFL defect could account for these differences.

Acknowledgements

This work was supported by National Institutes of Health (NIH) grants R01EY024542 and R01EY028135.

Disclosure

Dr Swanson is an unpaid consultant for Heidelberg Engineering, which had no input into the design of this study. Dr King reports no conflicts of interest. Both authors have no proprietary interest in any of the materials mentioned in this article.

References

1. Hood DC & Kardon RH. A framework for comparing structural and functional measures of glaucomatous damage. *Prog Retin Eye Res* 2007; 26: 688–710.
2. Ashimatey BS & Swanson WH. Between-subject variability in healthy eyes as a primary source of structural-functional discordance in patients with glaucoma. *Invest Ophthalmol Vis Sci* 2016; 57: 502–507.
3. Swanson WH, Felius J & Pan F. Perimetric defects and ganglion cell damage: interpreting linear relations using a two-stage neural model. *Invest Ophthalmol Vis Sci* 2004; 45: 466–472.
4. 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.
5. Goldmann H. Fundamentals of exact perimetry. 1945. *Optom Vis Sci* 1999; 76: 599–604.
6. Fellman RL, Lynn JR, Starita RJ & Swanson WH. Clinical importance of spatial summation in glaucoma. In: *Perimetry Update 1988/1989*, Heijl A (ed.) Kugler & Ghedini: Berkeley, 1989; pp. 313–324.
7. Westcott MC, McNaught AI, Crabb DP, Fitzke FW & Hitchings RA. High spatial resolution automated perimetry in glaucoma. *Br J Ophthalmol* 1997; 81: 452–459.
8. Westcott MC, Garway-Heath DF, Fitzke FW, Kamal D & Hitchings RA. 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.
9. Schiefer U, Flad M, Stumpp F *et al.* Increased detection rate of glaucomatous visual field damage with locally condensed grids: a comparison between fundus-oriented perimetry and conventional visual field examination. *Arch Ophthalmol* 2003; 121: 458–465.
10. Schiefer U, Malsam A, Flad M *et al.* Evaluation of glaucomatous visual field loss with locally condensed grids using fundus-oriented perimetry (FOP). *Eur J Ophthalmol* 2001; 11(Suppl 2): S57–S62.
11. Alluwimi MS, Swanson WH, Malinovsky VE & King BJ. Customizing perimetric locations based on en face images of retinal nerve fiber bundles with glaucomatous damage. *Transl Vis Sci Technol* 2018; 7: 5.
12. Frisen L. New, sensitive window on abnormal spatial vision: rarebit probing. *Vision Res* 2002; 42: 1931–1939.
13. Zalta AH & Birchfield JC. Detecting early glaucomatous field defects with the size I stimulus and STATPAC. *Br J Ophthalmol* 1990; 74: 289–293.
14. Wall M, Doyle CK, Eden T, Zamba KD & Johnson CA. Size threshold perimetry performs as well as conventional automated perimetry with stimulus sizes III, V, and VI for glaucomatous loss. *Invest Ophthalmol Vis Sci* 2013; 54: 3975–3983.
15. Mulholland PJ, Redmond T, Garway-Heath DF, Zlatkova MB & Anderson RS. Spatiotemporal summation of perimetric stimuli in early glaucoma. *Invest Ophthalmol Vis Sci* 2015; 56: 6473–6482.
16. Phu J, Khuu SK, Bui BV & Kalloniatis M. A method using Goldmann stimulus sizes I to V - measured sensitivities to predict lead time gained to visual field defect detection in early glaucoma. *Transl Vis Sci Technol* 2018; 7: 17.
17. Rountree L, Mulholland PJ, Anderson RS, Garway-Heath DF, Morgan JE & Redmond T. Optimising the glaucoma signal/noise ratio by mapping changes in spatial summation with area-modulated perimetric stimuli. *Sci Rep* 2018; 8: 2172.
18. Bek T & Lund-Andersen H. The influence of stimulus size on perimetric detection of small scotomata. *Graefes Arch Clin Exp Ophthalmol* 1989; 227: 531–534.
19. Phu J, Khuu SK, Zangerl B & Kalloniatis M. A comparison of Goldmann III, V and spatially equated test stimuli in visual field testing: the importance of complete and partial spatial summation. *Ophthalmic Physiol Opt* 2017; 37: 160–176.
20. Zalta AH. Use of a central 10 degrees field and size V stimulus to evaluate and monitor small central islands of vision in end stage glaucoma. *Br J Ophthalmol* 1991; 75: 151–154.
21. Pan F & Swanson WH. A cortical pooling model of spatial summation for perimetric stimuli. *J Vis* 2006; 6: 1159–1171.
22. Pan F, Swanson WH & Dul MW. Evaluation of a two-stage neural model of glaucomatous defect: an approach to reduce test-retest variability. *Optom Vis Sci* 2006; 83: 499–511.
23. Swanson WH, Pan F & Lee BB. Chromatic temporal integration and retinal eccentricity: psychophysics, neurometric analysis and cortical pooling. *Vision Res* 2008; 48: 2657–2662.
24. Hot A, Dul MW & Swanson WH. Development and evaluation of a contrast sensitivity perimetry test for patients with glaucoma. *Invest Ophthalmol Vis Sci* 2008; 49: 3049–3057.
25. Artes PH, Hutchison DM, Nicolela MT, LeBlanc RP & Chauhan BC. Threshold and variability properties of matrix frequency-doubling technology and standard automated perimetry in glaucoma. *Invest Ophthalmol Vis Sci* 2005; 46: 2451–2457.

26. Swanson WH, Malinovsky VE, Dul MW *et al.* Contrast sensitivity perimetry and clinical measures of glaucomatous damage. *Optom Vis Sci* 2014; 91: 1302–1311.
27. Sun H, Dul MW & Swanson WH. Linearity can account for the similarity among conventional, frequency-doubling, and gabor-based perimetric tests in the glaucomatous macula. *Optom Vis Sci* 2006; 83: 455–465.
28. Swanson WH, Sun H, Lee BB & Cao D. Responses of primate retinal ganglion cells to perimetric stimuli. *Invest Ophthalmol Vis Sci* 2011; 52: 764–771.
29. Swanson WH, Horner DG, Dul MW & Malinovsky VE. Choice of stimulus range and size can reduce test-retest variability in glaucomatous visual field defects. *Transl Vis Sci Technol* 2014; 3: 6.
30. 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; 121: 1359–1369.
31. 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.
32. De Lange Dzn H. Research into the dynamic nature of the human fovea-cortex systems with intermittent and modulated light. I. Attenuation characteristics with white and colored light. *J Opt Soc Am* 1958; 48: 777–784.
33. Kelly DH. Spatiotemporal variation of chromatic and achromatic contrast thresholds. *J Opt Soc Am* 1983; 73: 742–750.
34. Swanson WH, Ueno T, Smith VC & Pokorny J. Temporal modulation sensitivity and pulse-detection thresholds for chromatic and luminance perturbations. *J Opt Soc Am A* 1987; 4: 1992–2005.
35. Swanson WH, Dul MW, Horner DG, Liu T & Tran I. Assessing spatial and temporal properties of perimetric stimuli for resistance to clinical variations in retinal illumination. *Invest Ophthalmol Vis Sci* 2014; 55: 353–359.
36. Horner DG, Dul MW, Swanson WH, Liu T & Tran I. Blur-resistant perimetric stimuli. *Optom Vis Sci* 2013; 90: 466–474.
37. Ashimatey BS, King BJ, Burns SA & Swanson WH. Evaluating glaucomatous abnormality in peripapillary OCT enface visualisation of the retinal nerve fibre layer reflectance. *Ophthalmic Physiol Opt* 2018; 38: 376–388.
38. Alluwimi MS, Swanson WH, Malinovsky VE & King BJ. A basis for customising perimetric locations within the macula in glaucoma. *Ophthalmic Physiol Opt* 2018; 38: 164–173.
39. Vrabec F. The temporal raphe of the human retina. *Am J Ophthalmol* 1966; 62: 926–938.
40. Drasdo N, Millican CL, Katholi CR & Curcio CA. The length of Henle fibers in the human retina and a model of ganglion receptive field density in the visual field. *Vision Res* 2007; 47: 2901–2911.
41. Bland JM & Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1: 307–310.
42. Bland JM & Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999; 8: 135–160.
43. Anderson RS. The psychophysics of glaucoma: improving the structure/function relationship. *Prog Retin Eye Res* 2006; 25: 79–97.
44. Koenderink JJ, Bouman MA, Bueno de Mesquita AE & Slappendel S. Perimetry of contrast detection thresholds of moving spatial sine patterns. II. The far peripheral visual field (eccentricity 0 degrees-50 degrees). *J Opt Soc Am* 1978; 68: 850–854.
45. Koenderink JJ, Bouman MA, Bueno de Mesquita AE & Slappendel S. Perimetry of contrast detection thresholds of moving spatial sine wave patterns. I. The near peripheral visual field (eccentricity 0 degrees-8 degrees). *J Opt Soc Am* 1978; 68: 845–849.
46. Heijl A, Lindgren G & Olsson J. Normal variability of static perimetric threshold values across the central visual field. *Arch Ophthalmol* 1987; 105: 1544–1549.
47. Swanson WH, Dul MW, Horner DG & Malinovsky VE. Individual differences in the shape of the nasal visual field. *Vision Res* 2017; 141: 23–29.