The Robustness of Various Forms of Perimetry to Different Levels of Induced Intraocular Stray Light

Roger S. Anderson,¹,² Tony Redmond,¹,² D. Rodney McDowell,¹ Karen M. M. Breslin,¹ and Margarita B. Zlatkova¹

PURPOSE. To compare directly the robustness of standard automated perimetry (SAP), short-wavelength automated perimetry (SWAP), frequency-doubling perimetry (FDP), and grating-resolution perimetry (GRP) stimuli to different degrees of intraocular stray light induced by commercially available opacity-containing filters.

METHODS. Five white opacity filters of increasing density were used to simulate the typical forward light scatter and stray light values associated with age-related lens opacification and significant cataract. The individually induced intraocular stray light value for each filter was quantified with a stray light meter and plotted against individual perimetric thresholds for the right eyes of three normally sighted trained observers for SAP, SWAP, FDP, and GRP.

RESULTS. All tests were significantly but differently affected by increasing stray light. Overall average declines over a 1 log unit change in the stray light values were as follows: SAP, 4.85 dB; SWAP, 9.03 dB; FDP, 4.29 dB; and GRP, 1.36 dB. Standardized (z) scores were calculated after normalization to the spread of the normative data values for each instrument. These indicated that the standardized changes from baseline over the range of the five filters per log stray light unit were as follows: SAP, 2.177; SWAP, 1.96; FDP, 1.277; and GRP, 1.04.

CONCLUSIONS. The increased stray light values induced by cataract-simulating filters has a significant effect on all tests. However, GRP, which is known to be limited by retinal sampling rather than contrast, remains the most robust of the tests to the effects of intraocular stray light. The degree to which the normative “sensitivity” range for different types of perimetry might incorporate a component caused by individual differences in intraocular stray light is discussed and requires further research. (Invest Ophthalmol Vis Sci. 2009;50:4022–4028) DOI:10.1167/iovs.08-2934

Many types of perimeter tests are available commercially for the purpose of detecting, grading, and monitoring neural damage within the visual pathway. An important requirement of all these tests is to detect neural damage at an early stage. To be able to do this, it is important that the test be as robust as possible to aberrations and other imperfections in the optics of the eye. The need to discriminate between optical and neural visual losses while testing visual function has long been recognized.¹–⁵

The effects of optical defocus have been well established for standard automated perimetry (SAP),⁶–⁸ indicating a significant vulnerability, especially for small stimulus sizes. Short-wavelength automated perimetry (SWAP) has been shown to be resistant to the effect of blur because of the relatively large target size and the poor resolution properties of the SWS pathway.⁹ Frequency-doubling perimetry (FDP) has been found by one study to display significant robustness to the effects of optical blur⁹ and by another to be mildly vulnerable.¹¹ The typical low-pass-filtering effects of optical defocus would reasonably be expected to have little effect on a low-frequency grating stimulus such as that used in FDP (0.25 c cyc/deg). However, the aging process of the crystalline lens can have a profound effect on retinal image contrast, depending on the spatial frequency and wavelength, as a consequence of increased intraocular stray light and short wavelength light absorption, respectively.¹²–¹⁴

Cataract has been shown to have a significant effect on SAP thresholds,¹⁵,¹⁶ though some studies report little effect.¹⁷ Other studies report that such changes mask glaucomatous change.¹⁸–²¹

The effects of intrinsic cataract on SWAP have also been studied, and all those studies report a significant, diffuse effect. Kim et al.²² found that cataract predominantly causes a general reduction in sensitivity in both blue-on-yellow and white-on-white perimetry, but the reduction of blue-on-yellow sensitivity greatly exceeds that of white-on-white sensitivity by 2.4 times, presumably because it is more vulnerable to the effects of age-related lens yellowing. They advocate that caution should be observed when using blue-on-yellow perimetry to evaluate a patient who has lens opacification for glaucoma. Moss et al.²³ recommend caution in the interpretation of glaucomatous fields in the presence of cataract.

The effect of cataract removal on FDP has been investigated. Tanna et al.,²⁴ examining patients before and after cataract extraction, found a significant improvement in mean deviation (MD) but not pattern standard deviation (PSD) after surgery. Ueda et al.²⁵ reported an improvement in FDP after cataract surgery but no significant difference between eyes with an implant. Casson and James,²⁶ found a reduction in the false-positive rate postoperatively in otherwise healthy patients, especially for posterior subcapsular cataracts. Siddiqui et al.²⁷ found an improvement in MD and a worsening in PSD after cataract surgery in patients with glaucoma, indicating that cataract may mask glaucomatous loss, though other studies²⁸–³⁰ found no such change in PSD.

These studies examined patients with intrinsic cataract and, as such, did not separate the differing effects of the opacification and yellowing components of lens aging. Different perimetry tests may be affected differently by these separate components of lens aging and cataract. Several studies³¹–³³ have used diffusion filters of different density to simulate lens opacity in isolation for SAP and found a generalized reduction in sensitiv-
ity. Membrey et al.\textsuperscript{33} also found a generalized reduction in sensitivity for FDP when a single diffusing filter was used.

GRP makes use of sinusoidal gratings to test resolution acuity at different locations in the visual field. Previous studies examining grating acuity across the retina have shown that, though foveal acuity is limited by the low-pass-filtering characteristics of the eye’s optics, in peripheral vision the lower density sampling array of the retina is often unable to resolve the higher spatial frequencies that pass through the optics (for a review, see Ref. 34). Thus it is possible for the contrast in a grating stimulus to be detectable at peripheral locations, but the subject is unable to resolve the orientation because of its undersampled, aliased appearance.\textsuperscript{35–39} In this way resolution acuity, but not detection acuity, for high-contrast sinusoidal gratings has been reported to be limited by the sampling density of the coarsest array of retinal cells, the ganglion cells,\textsuperscript{30,41} at least in an eye with relatively clear media. However, the effects of different degrees of intraocular light scatter on achromatic grating resolution perimetry remain unknown.

In addition, we have little information about how these perimetric tests are affected relative to each other by age-related optical imperfections, particularly when those imperfections are of a magnitude that we would associate with normal aging rather than with significant cataract requiring surgery. In the clinical studies, with a few exceptions, visual function was generally investigated in the presence of a fairly gross optical defect (clinically significant cataract). Systems such as the Lens Opacity Classification System (LOCIS III), designed to grade cataractous lens appearance according to the degree of yellowing and opacity (observed with a slit lamp), are not suitable to measure more subtle changes in lens scatter associated with normal precataractous aging.

One important aspect of lens aging is the increase in stray light at the retina.\textsuperscript{44} Intraocular stray light causes a veiling luminance across the retina in the presence of glare and thus reduces the contrast of the retinal image. Various glare testing methods, typically involving some sort of acuity chart to which a glare source is introduced, have been used to quantify retinal stray light. The repeatability and discriminative power of such tests have been found to be poor.\textsuperscript{42–43} The psychophysical “direct compensation method” proposed by Van den Berg\textsuperscript{44} has shown better sensitivity than the glare testing methods in a wide range of studies\textsuperscript{49} and the recent development of the “compensation comparison” method and its implementation in a commercially available stray light meter (C-Quant; Oculus, Wetzlar, Germany) has permitted us to make more reliable and repeatable measures of intraocular stray light in naive subjects.\textsuperscript{35–36} The compensation comparison technique involves matching the magnitude of flicker in two halves of a central field, one of which has compensation luminance added in counterphase, to a flickering glare source added to the other.\textsuperscript{35} The instrument plots the measured stray light value for each subject against the subject’s age on top of a normative age function was generally investigated in the presence of a fairly gross optical defect (clinically significant cataract). Systems such as the Lens Opacity Classification System (LOCIS III), designed to grade cataractous lens appearance according to the degree of yellowing and opacity (observed with a slit lamp), are not suitable to measure more subtle changes in lens scatter associated with normal precataractous aging.

One important aspect of lens aging is the increase in stray light at the retina.\textsuperscript{44} Intraocular stray light causes a veiling luminance across the retina in the presence of glare and thus reduces the contrast of the retinal image. Various glare testing methods, typically involving some sort of acuity chart to which a glare source is introduced, have been used to quantify retinal stray light. The repeatability and discriminative power of such tests have been found to be poor.\textsuperscript{42–43} The psychophysical “direct compensation method” proposed by Van den Berg\textsuperscript{44} has shown better sensitivity than the glare testing methods in a wide range of studies\textsuperscript{49} and the recent development of the “compensation comparison” method and its implementation in a commercially available stray light meter (C-Quant; Oculus, Wetzlar, Germany) has permitted us to make more reliable and repeatable measures of intraocular stray light in naive subjects.\textsuperscript{35–36} The compensation comparison technique involves matching the magnitude of flicker in two halves of a central field, one of which has compensation luminance added in counterphase, to a flickering glare source added to the other.\textsuperscript{35} The instrument plots the measured stray light value for each subject against the subject’s age on top of a normative age range curve.

In this study we sought to simulate the scatter component of lens aging to different degrees to quantify the extent to which even moderate levels of opacification alone might increase the intraocular stray light value and affect different kinds of perimetry. We can thus investigate the robustness of three perimetric tests, readily available to practitioners, plus the more novel technique of GRP, to increased intraocular stray light induced by simulated lens opacity.

**Methods**

**Subjects**

Three experienced psychophysical observers were used in the study: TR (25 years old, −3.00 D myopic, blue irides), RR (45 years old, emmetropic, brown irides), and DRM (62 years old, emmetropic, brown irides). All had undergone full eye examination within the previous 3 months and had corrected logMAR visual acuities between −0.14 and −0.10. Results from a Humphrey visual field test, performed within the previous 3 months, showed normal total deviation and PSD plots for each subject. There was no evidence of any pathologic condition on slit lamp biomicroscopic examination, all lenses were clear of any grossly observable opacification, and there was no evidence of color vision anomalies on Ishihara and City University color vision tests. No subject had undergone any previous ocular surgery. None had any systemic condition or used any medication known to affect visual performance. Intraocular pressures were lower than 22 mm Hg in all subjects at the time of recruitment. Rim area parameters were shown by Heidelberg Retinal Tomograph and Moorfields’ regression analysis to be within normal limits.

Ethical approval for the study was obtained from the Faculty Research Ethics Committee, and the research protocol adhered to the tenets of the Declaration of Helsinki.

**Stimuli and Apparatus**

All testing was carried out in the Vision Science laboratory at the University of Ulster (Coleraine, Northern Ireland). The right eyes of the three observers were tested with the Humphrey Visual Field Analysers (HFA; Carl Zeiss Meditec, Dublin, CA) in white-on-white SAP and SWAP modes. Stimuli used were the standard Goldman size III (0.43°) white and Goldman size V (1.75°) blue, respectively. For SWAP, the yellow luminance background was the standard 100 cd/m², and the Goldman size V blue target had a narrow band of peak wavelength 440 nm and duration of 200 ms.

Subjects also underwent FDP (Welch Allyn, Skaneateles Falls, NY) and a form of GRP. The full-threshold central 24-2 programs for SAP and SWAP were used, as was the full-threshold C20 program for FDP. The FDP stimuli consisted of low spatial frequency (0.25 cyc/deg) sinusoidal gratings that underwent rapid (25 Hz) counterphase flicker.

For the GRP test, static achromatic sinusoidal grating patches with the same mean luminance as their surround were generated (Visual Stimulus Generator VSG2/3; Cambridge Research Systems, Rochester, UK) on a 21-inch high-resolution CRT monitor with pixel resolution of 1280 × 965 and a frame rate of 73 Hz (500PS; Sony Corp., Tokyo, Japan). The luminance of the background field was 10 cd/m². Chromaticity coordinates for the background and stimuli were $x = 0.218$ and $y = 0.328$. Gratings were circular patches of 2.5° diameter and 90% contrast encased within a Gaussian window and were presented at 20° eccentricity in each of the four oblique retinal meridians while the subject fixated a central cross. Presentation time was 1 second, including a ramp on-and-off time of 0.3 seconds. Threshold spatial frequency for resolution was determined with a two-alternative forced-choice staircase psychophysical procedure during which the subject was required to determine the orientation (horizontal or vertical) of the grating by pressing one of two response buttons. Spatial frequency increased by 0.8 dB after three correct responses and decreased by 0.8 dB after one incorrect response. Threshold was taken as the mean of the first four reversals.

Each subject was corrected for the working distance of each of the instruments (33 cm for all except FDP), as appropriate, for subject refractive error and age. An appropriate correcting lens for each subject was fitted in the standard lens holder for SAP and SWAP. A wide field spectacle frame was used for FDP and GRP to give maximum field of view, while the fellow eye was occluded with a patch. Pupil diameters were measured before the commencement of each test, with the subject seated in the observation position.

**Opacification Filters**

Different degrees of lens opacification were simulated using five white opacity-containing filters (Fog filters; Lee Filters, Andover, UK) to induce light scattering. The opacity filters, described in one of our
previous papers. Display very flat spectral transmission spectra ranging from 0.86 to 0.42 mean transmittance.

Our previous study also confirmed that the light scattered from a light source at angles between 4° and 15° varied inversely proportional to the square of the angular distance, which has been shown to be a typical relationship for the intraocular scatter caused by lens opacification. With the densest filter, high-contrast logMAR acuity was reduced for each subject by only 0.02 to 0.08 log units, whereas low-contrast (Bailey-Lovie) logMAR acuity was reduced by 0.22 to 0.45 log units for the same filter.

In addition, we used a stray light meter (C-Quant; Oculus) to measure the baseline intraocular stray light for each subject and the individual increase in intraocular stray light that resulted when each of the five filters was placed in front of the eye close to the cornea. The stray light meter plot in Figure 1 shows the baseline measurements for subjects (black dots) plotted against their ages. Subject TR displays above average stray light for age and subject DRM displays below average stray light for age, but all values are within the normal range indicated by the gray band.

Figure 1b shows how the stray light value for the 45-year-old subject (RR) changed for each filter. It can be seen that the filter (Fog 1; Lee Filters) effectively affords an average 45-year-old the stray light value of an average 58-year-old. Filters 2 and 3 (Fog 1; Lee Filters) turn the subject into an average 72- and 82-year-old, respectively (i.e., each filter adds an extra 10 to 14 years to the subject’s stray light value). Filters 4 and 5 take the subject into the realm of significant cataract. The effect was similar for the other two subjects. By way of comparison, the gray dot represents a 64-year-old patient with cataract displaying reasonable visual acuity (0.12 logMAR) and LOCSIII grading (NO2, C3, P1) but awaiting surgery in the hospital based on his stray light value (1.7).

The stray light meter (C-Quant; Oculus) thus yields a clinically meaningful indication of the degree to which each filter increases the forward light scatter for each subject, and the perimetric threshold can then be plotted against the actual stray light value for each filter condition. For analysis, the stimulus locations closest to those of the GRP test (20° eccentricity in the four oblique meridians) were extracted from the SAP, SWAP, and FDT plots. For each test condition, it was noted whether any of these locations fell outside the 95% confidence limit of the instrument.

During each testing session there was at least a 5-minute interval between each test to overcome fatigue, and observers were allowed a minimum of 2 minutes’ adaptation time to the yellow background before beginning each SWAP test.

**RESULTS**

All visual field plot results, for all observers, were within acceptable levels of reliability (<20% FP, <33% FN, <33% FL). Perimetric thresholds plotted against the induced stray light values for each subject can be seen in Figure 2. For initial comparison, all thresholds were expressed in the standard perimetry units of decibels (dB; 1 dB = 0.1 log unit change). (It should be noted that, for the FDT, contrast is normally based on stimulus amplitude rather than power, which means that 1 dB = 0.05 log unit; however we have here converted all threshold dB units to 0.1 log unit change.) For GRP, thresholds were converted from cyc/deg so that 1 dB represents a 0.1 log unit change from a 1 cyc/deg reference level. The decibel scale has been kept fixed for each plot (20 dB range). Looking at the overall sensitivities (heights of the curves), subject RR (an emmetrope with dark eyes) displayed highest sensitivity for SWAP and FDP. Subject TR (a low-medium myope with blue eyes), despite being the youngest subject, had an overall sensitivity comparable to that of the other, older, emmetropic subjects for GRP but lower than that of the others for SAP, SWAP, and FDP.

Given that these differences in the heights of the curves represented real between-subject differences in overall sensitivity, rather than measurement variability, we performed repeated-measures ANOVA in which the subject differences were removed from the error estimate to isolate the effect of stray light. All tests were progressively and significantly affected by increasing stray light values (P < 0.05) but to different degrees. SAP demonstrated an overall average decline of 4.85 dB/log stray light unit (P < 0.001), SWAP 9.05 dB/log stray light unit (P < 0.001), FDP 8.06 dB/log stray light unit (P < 0.001), and GRP 1.56 dB/log stray light unit (P = 0.014) over the range of five filter values. Fisher LSD post hoc analysis indicated that the mean sensitivity values became significantly different (P < 0.05) from the baseline at the third filter for SAP, the second filter for SWAP, and the first filter for FDP but not until the fourth filter for GRP. The filled symbols on each plot indicate a test condition in which one or more of the four analyzed locations was indicated by the instrument to fall outside the “2% probability” limit (from StatPac for Windows, available at http://www.statpac.com), i.e., close to the lower limit of the 2.5% to 97.5% (95%) confidence range.

However, a threshold change of 1 dB (or 0.1 log unit) for one instrument is not necessarily the same as a 1-dB change for another instrument. SAP and SWAP measures ΔI/I (Weber) contrast, FDP uses Michelson contrast, and GRP measures threshold spatial frequency. To yield a more clinically meaningful comparison, the normative 95% threshold confidence ranges were separately derived from previous studies for
SAP, SWAP, FDT, and GRP at this eccentricity and for this age range and are indicated by the gray bands on each plot. For SAP, despite a significant change in the averaged dB value, all points remained within the 95% confidence range for the test, in part because the subjects' baseline sensitivities lie in the upper region of the normative range. The final stray light value for subject DRM and the final two stray light values for subjects TR and RR (filled symbols) resulted in at least one of the four

**FIGURE 2.** Perimetric sensitivity/resolution (dB) versus log stray light value for (a) SAP, (b) SWAP, (c) FDP, and (d) GRP for all three subjects. For all perimeters, 1 dB here represents 0.1 log unit change in threshold.
points being highlighted by the instrument as falling outside the 2% probability. For SWAP, there was a greater range of baseline sensitivities, and only one subject (TR, with myopia) strayed outside the normative sensitivity range (at the two highest stray light values); the instrument also indicated at least one point as abnormal at 2% probability at these stray light values. For FDP, all subjects began with high baseline sensitivities and remained within the normative range. However, the instrument indicated at least one abnormal point at the higher stray light values. For GRP, all subjects remained within the normative range at all stray light values, and no individual locations fell outside the 95% probability range at any time.

We also compared the effects of stray light on different perimeters by using z-scores, based on the normal data for each perimeter. First we normalized the baseline thresholds in relation to the mean of the normative data ranges for each instrument. This allowed us to compute the standardized (z) scores for the change from baseline for each subject using the SD of the normative range. Over the range of the five filters, the decline in standardized score from baseline was 2.177 U/U log stray light for SAP, 1.96 U/U log stray light for SWAP, 1.277 U/U log stray light for FDP, and 1.04 U/U log stray light for GRP.

**DISCUSSION**

When testing patients for visual damage, it is important that any test used be capable of differentiating neural and optical losses of vision. Perimetry has as its goal the detection, grading, and monitoring of neural loss in the visual pathway and, as such, should be as robust as possible to the effects of optical imperfections be they optical aberrations, age-related lens yellowing, or opacification. Previous studies have shown declining performance with age for SAP,52,53 FDP,50,54 GRP,51,55 and especially SWAP,49,54 which has previously been shown to affect performance for SAP and SWAP,58,59 regardless of the method of optical correction,60 SWAP,61 and GRP62 but not for FDP.53 Simulated lens opacification has a significant effect (in decibel terms) to age-related deterioration in vision. Perimetry has as its goal the detection, grading, and monitoring of neural loss in the visual pathway and, as such, should be as robust as possible to the effects of optical imperfections be they optical aberrations, age-related lens yellowing, or opacification. Previous studies have shown declining performance with age for SAP,52,53 FDP,50,54 GRP,51,55 and especially SWAP,49,54 which has previously been shown to affect performance for SAP and SWAP,58,59 regardless of the method of optical correction,60 SWAP,61 and GRP62 but not for FDP.53 Simulated lens opacification has a significant effect (in decibel terms) to age-related deterioration in vision. The result for GRP builds on previous laboratory studies that have shown grating resolution in the periphery is limited by retinal ganglion cell sampling density. This sampling, rather than optical, limit to performance has been shown to allow peripheral resolution acuity to remain robust to optical defocus up to 3 D4,5 and to display no change in performance when grating contrast is reduced to as low as 20%.64 The same appears to be true for wide-angle scatter.

Although GRP appears to be more robust to the effects of forward light scatter than the other tests in decibel terms, the determination of clinically meaningful change for each type of test is still not straightforward. Any change in sensitivity resulting from increasing light scatter would have to take the patient outside the limits of the normative sensitivity value to be interpreted by the instrument as clinically significant. The problem with determining what constitutes a clinically meaningful change in this way is that the normative sensitivity range for each test includes, in different proportions, a neural and an optical component. It can be seen in Figure 2 that the normative range of sensitivity values is different for each of the perimetry tests used here, but less understood is the degree to which the variability in the instrument normative databases represents individual variations in intraocular stray light and true neural sensitivity differences. A perimetry test that displays poor robustness (in decibel terms) to increasing intraocular stray light may reasonably be expected to demonstrate a confoundingly greater variation in the normative range. This is borne out to some degree in the present data: in terms of decibels, SWAP displays a greater vulnerability than SAP to stray light but also a greater normative sensitivity range. Because of this larger normative variation (possibly as a result of greater scatter effects for short-wavelength stimuli), differences in the z-scores are smaller for SWAP than for SAP. The z-score thus masks the true vulnerability of SWAP to increased stray light. The same may be true for FDP. GRP displays the lowest vulnerability to intraocular stray light in decibels but also the narrowest normative threshold range, possibly because the interindividual scatter/stray light component of variability (which is retinal sampling limited rather than contrast limited) is lower. Although this means the z-score is not flattened by a large SD to the same degree as the other tests, it implies that any change observed in threshold, such as during testing for glaucoma, is more likely to represent true neural loss rather than optical attenuation.

Further work is required to better identify the components of perimetric sensitivity that are of optical rather than neural origin. The authors suggest that the term sensitivity be used more thoughtfully in the field of perimetry and that it be used primarily when we are confident that we are observing change in neural function rather than optical vulnerability. Clinical practice should strive to make use of stimuli that better separate optical and neural losses of vision to better detect and monitor diseases that cause damage to the neural visual pathway.

**References**

Effects of Stray Light on Different Perimetry Thresholds


47. Gardiner SK, Johnson CA, Spry PG. Normal age-related sensitivity loss for a variety of visual functions throughout the visual field. *Optom Vis Sci*. 2006;83(7):438–443.