Long-term Fluctuation in Short-Wavelength Automated Perimetry in Glaucoma Suspects and Glaucoma Patients

Natalie Hutchings, Sarah L. Hosking, John M. Wild, and John G. Flanagan

PURPOSE. To determine the magnitude of the homogenous, LF(Ho), and the heterogeneous, LF(He), components of the long-term fluctuation (LF) in glaucoma suspects and in stable primary open angle glaucoma (POAG) patients undergoing short-wavelength automated perimetry (SWAP) and to compare the magnitude of the SWAP LF components with those elicited by standard white-on-white (W-W) perimetry.

METHODS. The sample comprised 33 glaucoma suspects and 17 patients with early-to-moderate stable POAG who underwent W-W perimetry and SWAP at each of six visits over a mean period of 12.75 months (SD, 2.29). The LF(Ho), LF(He), and error components of the long-term fluctuation were determined between the third and seventh visual field examinations. The intervening visual field examinations and the optic nerve head parameters, derived both by stereo observation and by the Heidelberg Retinal Tomograph, were used to confirm stability over the follow-up period.

RESULTS. The LF(Ho) and LF(He) components were larger in the POAG patients than in the glaucoma suspects for both W-W perimetry and SWAP; the magnitude was independent of the depth of defect and of the short-term fluctuation. All three components of long-term fluctuation were greater for SWAP than for W-W perimetry, both in the glaucoma suspects and in the POAG patients.

CONCLUSIONS. SWAP exhibits greater long-term fluctuation than white-on-white perimetry. The usefulness of SWAP will be limited if the extent of this variability is not overcome in future statistical procedures developed to detect progressive visual field loss. (Invest Ophthalmol Vis Sci. 2001;42:2352-2373)

The presence of progressive visual field loss in primary open angle glaucoma (POAG) is a fundamental indicator for change in the therapeutic management of the patient. However, the separation of true progression from the physiological fluctuation in sensitivity between any two consecutive visual fields is a major clinical dilemma.

The development of short-wavelength automated perimetry (SWAP) has attracted considerable interest. In the Humphrey Field Analyzer (HFA; Humphrey Systems Inc, Dublin, CA), a 440-nm narrow band blue Goldmann size V stimulus is presented against a 100 cdm⁻² broadband (500–700 nm) yellow background. The blue stimulus is used to preferentially stimulate the short-wavelength-sensitive (SWS) pathway and the high luminance yellow broadband background to simultaneously suppress rod activity and to adapt the medium- and long-wavelength-sensitive pathways. SWAP has been shown to detect the presence of glaucomatous visual field loss and to identify progressive loss before conventional white-on-white (W-W) perimetry. Nevertheless, SWAP exhibits a greater within-examination variability (SF) in normals and in glaucoma than W-W perimetry. SWAP also exhibits a greater between-subject variability in normals compared with W-W perimetry; with the between-subject variability increasing with eccentricity.

The variation in the threshold estimate between examinations is known as the long-term fluctuation (LF). The LF is the variance additional to that of the SF and is present between two or more examinations. Classically, the LF is divided into two components: the homogeneous component, LF(Ho), and the heterogeneous component, LF(He). The LF(Ho) represents the proportion of the total LF that affects all the specified stimulus locations equally. Conversely, the LF(He) represents the proportion of the total LF that varies between the specified locations.

The magnitude of the between-examination variability, evaluated for a single stimulus location, is greater for SWAP than W-W perimetry in normals and in glaucoma patients and increases with eccentricity. However, the magnitude of the LF(Ho) and LF(He) in POAG patients and in patients who are suspect for glaucoma is unknown. The aim of the study was twofold: to determine the magnitude of the homogeneous and heterogeneous components of LF in glaucoma suspects and in stable primary open angle glaucoma patients undergoing SWAP, and to compare the magnitude of the SWAP LF components with those elicited by standard W-W perimetry.

METHODS. The sample was selected, retrospectively, from a cohort of 68 glaucoma suspects and 27 patients with glaucoma who had been followed prospectively over approximately 12 months (mean, 12.75 months; SD, 2.29 months). Informed consent was obtained from all volunteers. The study was carried out in accordance with the tenets of the Helsinki agreement and approval was obtained from the Medical Ethics Committee of Manchester Royal Eye Hospital and Manchester University.

To be classified as a glaucoma suspect in the given eye, one or more of the following inclusion criteria was met in addition to a normal W-W field: POAG in the fellow eye; a presenting IOP of 26 mm Hg and a vertical cup to disc (CD) ratio ≥ 0.6; a presenting IOP of 26 mm Hg and a positive family history for glaucoma; a presenting IOP of ≥ 21 mm Hg and a between-eye asymmetry in CD ratio ≥ 0.2; a presenting IOP of ≥ 21 mm Hg and a CD ratio of ≥ 0.8; a presenting IOP of ≥ 21 mm Hg and the presence of any focal disc abnormality, notching or disc hemorrhage; or a presenting IOP of ≥ 30 mm Hg. A normal visual field (HFA Program 30-2; Full Threshold algorithm) was defined by Total and Pattern Deviation probability analysis at...
the baseline visual fields. IOP was determined by Goldmann applanation tonometry.

For the POAG patients, a repeatable, glaucomatous W-W visual field defect (HFA Program 30-2; Full Threshold algorithm) at baseline was required to coexist with an abnormal optic disc, also consistent with a diagnosis of glaucoma (including increase in cup size, increase in cup disc ratio, disc asymmetry, changes in the lamina cribrosa, loss of neuroretinal rim, pallor, evidence of peripapillary atrophy, vessel changes, or disc margin hemorrhage),15 and a presenting IOP of >21 mm Hg. In patients where both eyes fulfilled the inclusion criteria, one eye was arbitrarily selected.

The exclusion criteria for both groups comprised a visual acuity in the designated eye of worse than 6/9; clinically significant cataract determined by slit-lamp examination with dilated pupils; a history of congenital color vision defect or optic nerve disorder not attributable to glaucoma; previous intraocular surgery, ocular trauma or inflammation; gonioscopic evidence of anterior chamber abnormality or angle closure; a history of CNS disorder; systemic medication known to affect the visual field; and ametropia of ≥ 6.00 DS and ≥ 2.50 DC.

The designated eye of each of the glaucoma suspects and each of the POAG patients was prospectively monitored for six visits over the follow-up period. Optic nerve head stability was assessed using two separate criteria. At each visit, the optic nerve head in the designated eye was evaluated by stereo observation using slit-lamp biomicroscopy and by scanning laser tomography (Heidelberg Retinal Tomograph [HRT]; Heidelberg Engineering, Heidelberg, Germany). The stereo examination at each visit evaluated features of the optic nerve head (as described above) for progressive damage.15 Twelve stereometric parameters were obtained from seven image series at each visit to describe the optic nerve head topography (disc area, mean height of contour, height variation in contour, cup volume, rim volume, volume above and below surface, mean depth inside contour, and mean retinal nerve fiber layer thickness). Stability was defined as no significant change over the time to follow-up in the parameters using repeated measures analysis of variance (P > 0.01). Although there are currently no standard criteria for defining change using the HRT, the approach adopted was intended to ensure a conservative definition of stability. The optic nerve head had to meet both criteria to be classified as stable.

The visual field was determined for W-W perimetry using the Full Threshold algorithm of the Humphrey Field Analyzer 640, with Program 30-2 (stimulus size III) and for SWAP using the Full Threshold algorithm with Program 24-2 and the default stimulus parameters of a blue size V stimulus and the 100 cd/m2 broadband yellow background. Program 24-2 was used for SWAP to provide an examination of duration similar to W-W perimetry.11 The order of the type of perimetry (i.e., W-W or SWAP) was randomly assigned between patients but remained constant for each patient at each examination throughout the period of follow-up. Each visual field was deemed to be reliable in terms of the responses to the catch trials (≤33% fixation losses, ≤33% false-positive responses, and ≤33% false-negative responses), and all visual field examinations were obtained by a single examiner (SLH).

All patients had experienced W-W perimetry before commencing the study, and the cohort was deliberately biased toward patients with early visual field loss. The first visual field examination was repeated within 2 weeks, and both examinations were designated as the baseline. These fields were excluded from further analysis. The purpose of the baseline examinations was to ensure that each patient was familiar with the measurement procedure and allowed the examiner to subjectively assess a patient’s reliability. The mean interval between the third and the seventh examinations for the sample as a whole was 11.88 months (SD, 1.32; median, 11.74 months; range, 9.84–15.24 months).

The W-W and SWAP visual fields were reviewed by one of the authors (JMW), experienced in visual field interpretation, who was masked to the outcome of any other clinical findings. The W-W visual fields were assessed by inspection of the Overview, Change Analysis and Glaucoma Change Probability Analysis print-outs of the HFA STATPAC statistical software. The SWAP visual fields were assessed by evaluation of the Overview print-out, alone, as the Change Analysis and the Glaucoma Change Probability Analysis print-outs are not commercially available. Linear regression analyses were also carried out for the W-W and SWAP summary visual field indices over time to follow-up for each of the glaucoma suspects and for each of the POAG patients. The regression analyses determined whether a significant change (P < 0.05) was evident in the magnitude of the Pattern Standard Deviation (PSD) and Corrected Pattern Standard Deviation (CPSD) indices over the time of follow-up. The PSD and CPSD indices were chosen as the most likely to indicate change consistent with glaucoma.16 A second linear regression analysis, only applicable to the POAG patients, determined if there was any significant difference in depth and area of an existing cluster of abnormal locations over the duration of follow-up. A cluster was defined as a nasal step or as two or more adjacent nonedge locations in the Pattern Deviation plot exhibiting abnormality at P ≤ 0.01 significance. The decibel depth of each pattern deviation cluster was determined from the STATPAC print-out.

Patients were excluded from the sample if one or more of the following was present: failure to complete the examinations in the follow-up period; change in topical therapy or surgical intervention during the follow-up period; change in optic nerve head topography and/or progressive visual field loss for either or both W-W perimetry and SWAP; or a learning effect lasting beyond the baseline fields for either, or both, W-W perimetry and SWAP. Fortytive individuals were excluded from the original cohort (Table 1), and the sample comprised 33 glaucoma suspects and 17 POAG patients. There was no statistically significant difference in the duration of follow-up between the two groups of patients (P = 0.752).

The mean age of the 33 suspects was 66.6 years (SD, 10.0 years; range, 42.4–77.3 years; 7 women, 26 men). The descriptive statistics for the group mean Mean Deviation (MD), PSD, and CPSD for W-W perimetry and for SWAP are shown in Table 1. Seventeen of the suspects were receiving medical therapy (topical β-blocker), which remained unchanged throughout the course of the study. The mean age of the 17 stable POAG patients was 64.9 years (SD, 11.4 years; range, 42.8–84.9 years; 8 women, 9 men). The mean age of the glaucoma suspects and the POAG patients were not significantly different (P = 0.332). Descriptive statistics for the group mean MD, PSD, and CPSD for W-W perimetry and for SWAP, and the stage of the glaucomatous visual field defect for W-W perimetry, as defined by the criteria of Hodapp et al.,17 are shown in Table 1. All POAG patients were therapeutically controlled using topical β-blocker agents, and in one patient this was combined with miotic therapy. None of the POAG patients had a change in medical therapy over the course of the study.

The LF(Ho) and LF(He) components, expressed in decibels, were calculated between the third and the seventh visual field examinations (i.e., using data from the 3rd and 7th examinations only) for W-W and for SWAP. The determination of the LF(Ho) and LF(He) components has been previously described.15 Briefly, LF(Ho) and LF(He) were derived from a two-factor ANOVA with replications,18 based on the double determinations of sensitivity at the 10 standard stimulus loca-
tions incorporated in Program 30-2 and in Program 24-2 of the HFA, using the formula:

\[ Y_{jk} = \mu + I_j + V_k + LV_{jk} + E_{jk} \]

where \( Y \) is the threshold determination at each individual location, \( \mu \) is the overall mean value of sensitivity, which is related to MD by an aggregate constant, \( I_j \) is the effect of the examination locations, \( V_k \) is the effect of each visual field examination, \( LV_{jk} \) is the interaction of \( I_j \) and \( V_k \), and \( E_{jk} \) is the experimental error. The integers \( j, k, \) and \( l \) represent, respectively, the number of considered stimulus locations with a double determination of threshold (\( j = 1, 2, \ldots, n \)), the number of examinations (\( k = 2 \)), and the number of determinations of threshold at each considered location (\( I = 2 \)).

The respective values of LF(Ho) and LF(He) were determined by partitioning the variance derived from the ANOVA that is attributable to each component of LF. The mean square estimates (MSE) for \( V_k \) attributable to LF(Ho), and \( LV_{jk} \) attributable to LF(He), were then reduced to their constituent variances to remove the accompanying error variance \( E_{jk} \) by the hypothesis of

\[ \text{MSE}(V_k) = \alpha^2 + \beta^2 + \gamma^2 \] \hfill (1)

\[ \text{MSE}(LV_{jk}) = \alpha^2 + \beta^2 + \gamma^2 \] \hfill (2)

\[ \text{MSE}(E_{jk}) = \delta^2 \] \hfill (3)

In instances where a negative component of LF (\( \sigma_{k} \) or \( \sigma_{v} \)) was obtained, the variance(s) that exhibited a negative component of LF was combined with that of the error term, \( E \). The error term was then recalculated, the new value more accurately reflecting the magnitude of both the error term and the previously negative LF component. This value was then used in all subsequent analyses.

The decibel unit of differential light sensitivity is a relative measure referenced to the maximum stimulus luminance. Any decibel value obtained using W-W perimetry is not directly comparable with that using SWAP because of the differing maximum stimulus luminance used by each procedure. Therefore, to enable a more direct comparison to be made between the components of LF derived by each perimetric technique, the log unit equivalents of the LF(Ho) and LF(He) were calculated. This was achieved by converting the decibel value to log units (i.e., relative to the maximum stimulus luminance), and this value was then expressed as a reciprocal. Thus, the reciprocal log unit would approach zero (and the apostille value approach the maximum stimulus luminance) as the decibel value of the long-term fluctuation approach zero.

**RESULTS**

The frequency distribution and descriptive statistics of the LF(Ho) and LF(He) are shown for W-W and SWAP in Figure 1. The LF(Ho), LF(He), and error component were greater in the POAG patients than in the suspects for both W-W (\( P < 0.001 \)) and SWAP (\( P < 0.002 \)).

The group mean magnitude of the error component, \( E \), in W-W perimetry was 1.35 dB (SD, 0.37; range, 0.84 –2.24) for the glaucoma suspects and 1.72 dB (SD, 0.53; range, 1.15–3.29) for the POAG patients. For SWAP, the group mean error component was 1.76 dB (SD, 0.54; range, 0.89 –3.08) for the suspects and 2.27 dB (SD, 1.02; range, 0.75–5.07) for the POAG patients.

The group mean SF of the glaucoma suspects at the third examination was 1.27 dB (SD, 0.44) for the W-W procedure and 1.80 dB (SD, 0.58) for SWAP. Similarly, the group mean SF of the stable glaucoma patients was 2.03 dB (SD, 0.69) for W-W and 1.90 dB (SD, 0.49) for SWAP. Linear regression analysis showed the SF at the third examination to have a positive relationship with the error component of the long-term fluctuation in the glaucoma suspects for W-W perimetry (coefficient of determination, \( R^2 = 0.792 \); regression coefficient, \( P < 0.001 \)) and for SWAP (\( R^2 = 0.781, P < 0.001 \)). The corresponding data for the POAG patients did not show such a relationship (W-W: \( R^2 = 0.051; P = 0.384 \); SWAP: \( R^2 = 0.178; P = 0.092 \)). The SF was unrelated to either the LF(Ho) or the LF(He) in either W-W perimetry or SWAP in either the glaucoma suspect or POAG samples (\( R^2 < 0.50; P > 0.05 \)).

Similarly, the magnitude of the MD and that of the CPSD at the third examination were unrelated to either the LF(Ho) or

**TABLE 1. Descriptive Data of the Visual Field Characteristics of the Sample**

<table>
<thead>
<tr>
<th>Study Cohort</th>
<th>W-W or SWAP Field Progression</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deterioration</td>
<td>Improvement or Unreliable</td>
</tr>
<tr>
<td></td>
<td>MD (dB)*</td>
<td>PSD (dB)*</td>
</tr>
<tr>
<td>Glaucma suspects</td>
<td>W-W 7 W-W 0</td>
<td>SWAP 5 SWAP 3</td>
</tr>
<tr>
<td></td>
<td>(−5.44 to +1.44)</td>
<td>(−6.80 to +6.15)</td>
</tr>
<tr>
<td>POAG</td>
<td>W-W 6 W-W 2</td>
<td>SWAP 5 SWAP 2</td>
</tr>
<tr>
<td></td>
<td>(−4.76 to 2.05)</td>
<td>(−5.67 to 2.04)</td>
</tr>
<tr>
<td></td>
<td>(−4.39 ± 3.94)</td>
<td>(2.13 to 7.48)</td>
</tr>
</tbody>
</table>

Thirty-five glaucoma suspects and 10 POAG patients of the study cohort were excluded from the sample for one or more of the identified reasons.

* Values are means ± SD, with range in parentheses.

† Using Hodapp criteria. 17
the LF(He) for W-W perimetry or for SWAP. The $R^2$ values associated with the regression coefficients that failed to reach the criteria for significance ranged from a minimum of 0.001 to a maximum of 0.265.

The relationship between the W-W and SWAP LF(Ho), expressed in reciprocal log units is illustrated in Figure 2 (top). The corresponding data for the LF(He) component is shown for W-W perimetry (bottom left) and for SWAP (bottom right).

**DISCUSSION**

The homogeneous, LF(Ho), and heterogeneous, LF(He), components of the long-term fluctuation were larger for SWAP than for traditional W-W perimetry, both in the glaucoma suspects and in the glaucoma patients. This corroborates the findings of previous studies that have addressed the physiological variation between visual field examinations in normal subjects and in glaucoma patients.\(^8,10\) The magnitude of the LF(Ho), LF(He), and error components, specified in reciprocal log units, was significantly larger for SWAP than for W-W perimetry (LF(Ho), $P = 0.005$; LF(He), $P = 0.004$; E, $P < 0.001$).

The difference in the magnitude of the LF reported in this study for SWAP and that reported for normal subjects arises in part from the mathematical definition of LF. The LF described by Kwon et al.\(^8\) was defined as the statistical variance exhibited by the threshold value at a given single stimulus location over a series of examinations. This latter method is of limited value in separating variability from progressive loss in the clinical environment, because it requires several examinations before the value can be determined. The LF used in the present study can be used between any pair of examinations from the outset and provides a measure of reliability between examinations,\(^13\) in a similar fashion to the use of the short-term fluctuation as an indicator of reliability within an examination.

Alternative statistical techniques for approximating interval estimates have been described,\(^23,24\) and these methods may minimize the instances in which a negative estimate of an LF component is obtained. However, the likely effect of the method used in this study is an overestimation of the magnitude of the components of the LF that, in the first instance, exhibit a negative variance. The method described in this study therefore provides a conservative standard for the definition of excessive long-term fluctuation. For the suspects, correction of a negative variance corresponding to the LF(Ho) was made in nine patients for W-W perimetry and in eight patients for SWAP (two of whom were common to both types of perimetry). A correction for a negative variance corresponding to the LF(He) was made in four patients for W-W perimetry and six patients for SWAP (two of whom were common to both types of perimetry). The equivalent figures for the glaucoma patients were four patients each for W-W and SWAP LF(Ho) and three patients each for W-W and SWAP LF(He).

Previous studies have shown that ocular media absorption adversely influences the SWAP visual field more than the white-on-white field.\(^11,25,26\) Although the extent of the ocular media absorption can be quantified,\(^26–28\) the procedures require specialized techniques,\(^26–28\) are time consuming and cannot real-
The magnitude of the loss in Mean Sensitivity due to absorption is approximately 0.5 dB per decade. It has been suggested that the increased magnitude in the between-individual variability of the threshold estimate in SWAP will increase with age due to the increased effects of absorption and loss of neuronal tissue. The greater variability of SWAP compared with W-W perimetry will undoubtedly be exacerbated by the presence of glaucoma. The use of SWAP will therefore necessitate more stringent statistical procedures to account for the increased within- and between-examination variability and to separate age-related changes from that due to progressive field loss.

The positive relationship between the SF and the error component of the long-term fluctuation in the glaucoma suspects for both W-W perimetry and SWAP is an expected finding since the error component is analogous to the mean short-term fluctuation across the two examinations. The stable glaucoma patients did not exhibit such a relationship, probably as a result of the greater between-individual variability in the SF, particularly for W-W perimetry. Perhaps more surprising is the lack of relationship of the LF(Ho) or the LF(He) with the MD or CPSD for both W-W perimetry and SWAP. The physiological variation in the sensitivity at a single stimulus location between two examinations using W-W perimetry increases as the threshold increases, and the long-term fluctuation is correlated with the mean visual field loss. It might be expected, therefore, that the magnitudes of the LF(Ho), LF(He), and error component of the LF would exhibit a concomitant increase with a decrease in the MD and an increase in the CPSD, particularly in the moderate loss range. No such finding was apparent in this study. This may be explained by the limited number of patients in the moderate-to-severe range of MD and CPSD values. Alternatively, it might be a consequence of the global nature of the visual field indices, which are based on the threshold values derived at all stimulus locations, and the “sampled” nature of the LF components, which are based on the 10 locations at which double determinations of threshold are undertaken. However, the use of the 10 doubly determined locations to derive the visual field indices does not elicit a materially different relationship with the components of LF for W-W perimetry. It might also be argued that the LF components derived from these 10 locations do not adequately reflect the physiological variation of the visual field as a whole. However, these 10 stimulus locations are the only common stimulus locations in the HFA at which double determinations of sensitivity are undertaken in successive examinations. The SF is determined from the same 10 stimulus locations and is an accepted measure of within-examination variability despite similar limitations.

The magnitudes of the group means of all three components of long-term fluctuation were higher for the stable glaucoma patients than for the glaucoma suspects for both W-W and SWAP. The difference in the group means between these two samples cannot be attributed to a difference in age and occurred irrespective of the group mean SF of the second field being slightly lower in the glaucoma group for SWAP than for W-W perimetry. The log unit magnitudes of the LF(Ho), LF(He), and error components were greater for SWAP than for W-W perimetry. The increased short-term and long-term fluctuation may result from a decreased sampling of the neuronal constituents with SWAP or, more likely, from the flatter frequency-of-seeing curves obtained with SWAP. The higher values may have been influenced by the longer duration of the learning effect with SWAP and/or a carryover fatigue effect between examinations within each visit. Every effort was made to minimize the learning effect in SWAP by ensuring that the patients were experienced in SWAP perimetry and by excluding pa-
tients who exhibited systematic improvement in overall sensitivity over time to follow-up. It is evident that the increased variability within- and between-examinations exhibited by SWAP compared with W-W perimetry limits the current utility of SWAP in the clinical situation. Such variability must be reduced if SWAP is to become a viable technique.

Acknowledgments
The authors thank Eamonn O’Donoghue, Consultant Ophthalmologist, Manchester Royal Eye Hospital, for assistance with patient recruitment.

References