The Visual Field Indices in Primary Open-Angle Glaucoma

John G. Flanagan,*§ John M. Wild,†† and Graham E. Trope§‡

Purpose. The distribution of sensitivity across the visual field, as determined by automated threshold static perimetry, can be summarized in terms of visual field indices. Such indices can be weighted for the variation in threshold at each eccentricity. The aims of this study were to determine the influence of the weighting factor, the relationship between the unweighted indices derived from Programs 30-2 and 24-2, and the relationship between the number of double determinations of threshold and the magnitude of the short-term fluctuation.

Methods. One visual field derived by Program 30-2 of the Humphrey Field Analyzer was selected from each of 60 consecutive patients with primary open-angle glaucoma. The first two fields from each individual patient were avoided. Unweighted visual field indices were calculated and compared with the Program 30-2 weighted indices using an assessment of agreement evaluated with respect to the 95% confidence limits of the population.

Results. The weighting function had no influence on the mean deviation, but it caused a slight reduction in the short-term fluctuation and an elevation in the pattern and corrected pattern standard deviations. There was little difference between the indices generated by Programs 30-2 and 24-2. The short-term fluctuation increased with an increase in the number of double determinations of threshold.

Conclusions. The weighting function had little clinical influence on the visual field indices. The indices derived from Programs 30-2 and 24-2 were similar, and the short-term fluctuation would better reflect the intratest variability if all available double determinations of threshold were used to calculate the index. Invest Ophthalmol Vis Sci. 1993;34:2266-2274.

The distribution of sensitivity across the visual field determined by automated threshold static perimetry can be described in terms of visual field indices.1,2,3 These indices reduce the data to summary measures and, as such, are designed to assist the clinician in defining the visual field loss.

For the Octopus perimeters, Flammer et al4 introduced the mean defect, a measure of the mean depression of the visual field compared with that of a normal age-matched person; short-term fluctuation, a measure of intratest variability; loss variance, a measure of the uniformity of the shape of the visual field; and corrected loss variance, a measure of the loss variance corrected for the short-term fluctuation.

The corresponding indices for the Humphrey Field Analyzer were introduced by Heijl et al.3 These were weighted to compensate for the variability of the threshold at each stimulus location out to an eccentricity of 30°. The more central stimulus locations were given a greater weighting in the calculation of the indices than were the more peripheral locations where the fluctuation of the differential light sensitivity is greater.4 The weighting function was the reciprocal of the interindividual fluctuation,
for a given stimulus location. These corresponding indices were called mean deviation, short-term fluctuation, pattern standard deviation, and corrected pattern standard deviation. The main effect of the weighting function was to minimize the short-term fluctuation and the pattern SD in normal subjects. The effect of the function on the visual field indices in glaucoma has received little attention, and it is unclear whether the weighted indices give clinically different results from those of the unweighted indices.

Standard clinical practice involves the investigation of the central field out to an eccentricity of 30° (Octopus Program G1 and Humphrey Program 30-2). In cases of paracentral loss, however, clinical experience has dictated the trend toward reducing the number of stimulus locations in the more peripheral regions of the central field (Humphrey Program 24-2). Serial visual field analysis frequently involves a comparison of the indices from Programs 30-2 and 24-2, and the indices are frequently regressed against the time of follow-up. The relationship between the indices derived from Programs 30-2 and 24-2 is unknown, and it is unclear to what extent this type of serial visual field analysis might be compromised.

The derivation of the short-term fluctuation varies between the Octopus and Humphrey systems. The Octopus Program G1 calculates the short-term fluctuation from a second determination of threshold, undertaken in Phase 2 of the program, at each of 59 stimulus locations. The Humphrey field analyzer algorithm for the calculation of short-term fluctuation is based on a double determination of threshold at each of ten specified stimulus locations within 21° eccentricity. The number of test locations involving a double determination of threshold increases intrinsically within the algorithm as a function of decreased patient reliability at a given stimulus location or between adjacent locations and/or as a function of the irregularity of the field loss between the adjacent locations. The relationship between short-term fluctuation and the number of double determinations of threshold has not been investigated. Indeed, it can be conjectured that patients with a large short-term fluctuation may exhibit a large number of double determinations of threshold. Furthermore, it can be hypothesized that, if all double determinations of threshold were used to calculate short-term fluctuation, the magnitude of the short-term fluctuation would better reflect the intratest variability.

The aims of the study were to determine the influence of the weighting factor on the visual field indices, the relationship between the unweighted indices derived from Programs 30-2 and 24-2, and the relationship between the number of stimulus locations with double determinations of threshold and the magnitude of the short-term fluctuation.

### MATERIALS AND METHODS

#### Sample

The sample comprised one visual field recorded with Program 30-2 (stimulus size III) of the Humphrey field analyzer from each of 60 consecutive patients with primary open-angle glaucoma attending the Glaucoma Clinic of the Toronto Hospital General Division. The methods of recruitment, including proper consent and approval, and the research protocol complied with the tenets of the Declaration of Helsinki. All patients had presented initially with an intraocular pressure of greater than 21 mm Hg and exhibited visual field loss. Patients receiving miotic therapy were excluded. The first two fields from each patient were not used to minimize the influence of the learning effect. Fields exhibiting greater than 20% fixation losses, 33% false-positive responses, or 33% false-negative responses were generally excluded. However, in 15 cases, the total response to one or more of the catch trials was outside of these limits, but the fields were considered to be clinically reliable and were therefore included. Generalized reductions in sensitivity were excluded if they were considered (from inspection of the total and pattern deviation plots) to be the result of media opacity. All fields exhibited field loss within the stimulus grids of both Programs 24-2 and 30-2. The mean age of the patient group was found to be 69.2 yr (SD, 6.6 yr), and the male-to-female ratio was 1:1.55.

#### Calculation of Visual Field Indices

The visual field data files for each patient, contained on the Humphrey field analyzer formatted diskettes, were accessed by personal computer and analyzed by a series of custom-designed software modules.

To investigate the effect of the weighting function, the unweighted indices for each subject were calculated, compared with the corresponding weighted indices obtained by measurement, and displayed on the Humphrey field analyzer STATPAC print out. The calculated unweighted indices were derived using the sign convention of Heijl et al. The effect of the weighting factor as a function of the peripheral angle was investigated by separately generating unweighted indices representing the field out to an eccentricity of 30° and the field between 15–30° eccentricity. The inner region contained 24 stimulus locations and the outer region, 52 stimulus locations. The resulting unweighted regional indices were then each compared with the respective unweighted and weighted global indices derived from the (Program 30-2) field out to 30° eccentricity.

To investigate the effect of the reduction in the number of peripheral stimulus locations between Programs 30-2 and 24-2, the unweighted indices were re-
FIGURE 1. Scattergrams of the Humphrey Field Analyzer Program 30-2 unweighted indices derived by calculation against the corresponding weighted indices obtained by measurement. (Top left) Mean deviation (r = 0.98). (Top right) Short-term fluctuation (r = 0.24). (Bottom left) Pattern standard deviation (r = 0.94). (Bottom right) Corrected pattern standard deviation (r = 0.91).

calculated from the Program 30-2 data for a stimulus grid corresponding to that of Program 24-2 and compared with the corresponding unweighted indices calculated from the complete Program 30-2 grid.

To investigate the relationship between the magnitude of the short-term fluctuation and the number of stimulus locations involving a double determination of the threshold, the unweighted short-term fluctuation was calculated for each subject using all available double determinations of threshold and compared with the unweighted short-term fluctuation calculated from the double determination of the threshold at the ten standard stimulus locations.

The normal age-matched data at each test location necessary to calculate the various unweighted indices were provided by Allergan Humphrey and were identical to the normal data used in the STATPAC software. They were similar to the subset of that data published previously.4

Analysis

The results for the unweighted and weighted values of a given index were first evaluated in terms of a two tailed paired t-test. Scattergrams were then plotted for the various combinations and types of indices, and the data were analyzed further using an alternative technique,10 which is more appropriate than correlational analysis for assessing the level of agreement between two methods of clinical measurement. The technique overcomes the problems associated with the dependence of the data and with the potential clustering of the data points close to the particular regression line between any two given indices, thus making it difficult to recognize any between-method differences in the unweighted and weighted calculations. The difference between the appropriate pairs of indices for each subject were plotted against the mean of the two measures and interpreted first with respect to the mean of the
Visual Field Indices in POAG

FIGURE 2. Scattergrams of the difference between the weighted and unweighted pairs of indices for each subject against the mean of the two measures derived from Program 30-2 of the Humphrey Field Analyzer. The mean of the differences is indicated by the dotted line and the designated limits of agreement, defined as the mean of the differences ± 1.96 times the SD of the differences, by the solid lines. (Top left) Mean deviation. (Top right) Short-term fluctuation. (Bottom left) Pattern standard deviation. (Bottom right) Corrected pattern standard deviation.

RESULTS

Effect of the Weighting Function
The scattergrams of the Program 30-2 unweighted indices derived by calculation against the corresponding weighted indices obtained by measurement are illustrated in Figure 1. The group mean weighted mean deviation of -8.3 dB (SD, 6.2 dB) was not significantly different from the group mean unweighted mean deviation of -8.4 dB (SD, 5.7 dB; \( P = 0.521 \)). The mean of the differences between the weighted and unweighted mean deviation for Program 30-2 was 0.10 dB, and the limits of agreement were between +2.4 dB and -2.2 dB (Fig. 2).

The group mean weighted pattern standard deviation of 6.9 dB (SD, 3.1 dB) was significantly greater than the group mean unweighted pattern standard deviation of 6.4 dB (SD, 2.2 dB; \( P = 0.01 \)), and this disparity increased as the depth of the loss increased (Fig. 2). The mean of the differences was 0.45 dB, and the limits of agreement were between +2.98 dB and -2.08 dB (Fig. 2).

The group mean weighted short-term fluctuation of 2.5 dB (SD, 1.1 dB) was significantly less than the group mean unweighted short-term fluctuation of 3.2 dB and, second, to the designated limits of agreement defined as the mean of the differences ± 1.96 times the SD of the differences. In reporting differences between weighted and unweighted mean deviations, a positive difference indicated that the magnitude of the unweighted defect was greater (ie, more negative).

TABLE 1. Mean of the Differences, Limits of Agreement, and Two-tailed Significance Levels (Paired t Test) for the Weighted and Unweighted Global Program 30-2 Indices and the Unweighted Indices out to an Eccentricity of 15°

<table>
<thead>
<tr>
<th>Corrected Pattern</th>
<th>Mean Deviation</th>
<th>Short-term Fluctuation</th>
<th>Pattern Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weighted 30-2 minus unweighted (0°–15°)</td>
<td>Correlation coefficient (r)</td>
<td>0.96</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>Mean of differences</td>
<td>-1.46</td>
<td>+0.14</td>
</tr>
<tr>
<td></td>
<td>Limits of agreement</td>
<td>+2.23</td>
<td>+2.70</td>
</tr>
<tr>
<td></td>
<td>Significance level (P)</td>
<td>&lt;0.001</td>
<td>0.4</td>
</tr>
<tr>
<td>Unweighted 30-2 minus unweighted (0°–15°)</td>
<td>Correlation coefficient (r)</td>
<td>0.92</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>Mean of differences</td>
<td>-1.56</td>
<td>+0.84</td>
</tr>
<tr>
<td></td>
<td>Limits of agreement</td>
<td>+5.67</td>
<td>+3.19</td>
</tr>
<tr>
<td></td>
<td>Significance level (P)</td>
<td>&lt;0.001</td>
<td>0.015</td>
</tr>
</tbody>
</table>

Note: a negative value for the mean of the differences in the mean deviation indicates that the magnitude of the weighted mean deviation was greater (i.e., more negative), whereas a positive value for that of the difference in each of the remaining indices indicates that the weighted defect was greater. A similar sign convention is followed in Tables 2 to 4. Note also that the Pearson product moment correlation coefficients, which merely assess the strength of a relation between two variables not the agreement between them, are included in Tables 1 to 4 only for comparative purposes.

The mean of the differences was -0.66 dB, and the limits of agreement ranged from +2.1 to -3.4 dB. The mean of the differences was 0.56 dB, and the limits of agreement were between +3.81 dB and -2.69 dB (Table 2).

The relationship between the global weighted and unweighted indices and the unweighted indices derived by calculation out to an eccentricity of 15° are illustrated in Table 1. The corresponding data for the indices between 15°–30° eccentricity are given in Table 2. The magnitude of the limits of agreement for each index was broadly similar between both the weighted

TABLE 2. Mean of the Differences, Limits of Agreement, and Two-tailed Significance Levels (Paired t Test) for the Weighted and Unweighted Global Program 30-2 Indices and the Unweighted Indices Between 15° and 30° Eccentricity

<table>
<thead>
<tr>
<th>Corrected Pattern</th>
<th>Mean Deviation</th>
<th>Short-term Fluctuation</th>
<th>Pattern Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weighted 30-2 minus unweighted (15°–30°)</td>
<td>Correlation coefficient (r)</td>
<td>0.94</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Mean of differences</td>
<td>+0.76</td>
<td>-0.87</td>
</tr>
<tr>
<td></td>
<td>Limits of agreement</td>
<td>+5.05</td>
<td>+2.23</td>
</tr>
<tr>
<td></td>
<td>Significance level (P)</td>
<td>0.01</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Unweighted 30-2 minus unweighted (15°–30°)</td>
<td>Correlation coefficient (r)</td>
<td>0.98</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>Mean of differences</td>
<td>+0.66</td>
<td>-0.16</td>
</tr>
<tr>
<td></td>
<td>Limits of agreement</td>
<td>+2.85</td>
<td>+1.00</td>
</tr>
<tr>
<td></td>
<td>Significance level (P)</td>
<td>&lt;0.001</td>
<td>0.044</td>
</tr>
</tbody>
</table>

Downloaded From: http://iovs.arvojournals.org/pdfaccess.ashx?url=/data/journals/iovs/933173/ on 04/15/2017
TABLE 3. Mean of the Differences, Limits of Agreement, and Two-tailed Significance Levels (Paired t Test) for the Unweighted Program 30-2 and 24-2 Indices

<table>
<thead>
<tr>
<th></th>
<th>Mean Deviation</th>
<th>Short-term Fluctuation</th>
<th>Pattern Standard Deviation</th>
<th>Corrected Pattern Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unweighted 30-2 minus unweighted 24-2</td>
<td>0.98</td>
<td>0.85</td>
<td>0.91</td>
<td>0.89</td>
</tr>
<tr>
<td>Correlation coefficient (r)</td>
<td>-0.40</td>
<td>+0.20</td>
<td>+0.35</td>
<td>+0.36</td>
</tr>
<tr>
<td>Mean of differences</td>
<td>+1.97</td>
<td>+1.18</td>
<td>+2.29</td>
<td>+2.91</td>
</tr>
<tr>
<td>Limits of agreement</td>
<td>-2.78</td>
<td>-0.79</td>
<td>-1.60</td>
<td>-2.20</td>
</tr>
<tr>
<td>Significance level (P)</td>
<td>0.014</td>
<td>0.004</td>
<td>0.01</td>
<td>0.042</td>
</tr>
</tbody>
</table>

The group mean unweighted short-term fluctuation calculated from the standard ten stimulus locations of 1.83 dB (SD, 1.2 dB) was significantly lower than that from the group mean unweighted short-term fluctuation calculated using all double determinations of threshold of 3.21 dB (SD, 0.9 dB; \( P < 0.001 \)). The mean of the differences was 1.37 dB, and the corresponding limits of agreement ranged from +4.12 dB to -1.38 dB.

DISCUSSION

These results suggest that the weighting function has little influence on the mean deviation index. They are in agreement with the findings of others who investigated the effect of the weighting function on the mean defect and mean deviation using data calculated from Phase 1 of the Octopus Program Gl.1112 They found that the pointwise fluctuation of sensitivity recorded with the Octopus increased minimally with increase in eccentricity compared with that obtained with the Humphrey perimeter.4 They concluded that the weighting factor derived from the Octopus data was such that two indices, mean defect and mean deviation, were virtually identical and could be used interchangeably. The current study extends these findings to the unpublished weighting factors of Heijl et al3 for the Humphrey system and reaches the same conclusion despite the differences in the weighting factors used.

The wider limits of agreement between the 15-
FIGURE 3. Scattergrams of the magnitude of the short-term fluctuation against the number of double determinations of threshold. (Top) Weighted (left) and unweighted (right) global short-term fluctuation from ten double determinations of threshold. (Middle) Unweighted short-term fluctuation from all double determinations of threshold. (Bottom) Unweighted short-term fluctuation between 0–15° eccentricity (left) and between 15–30° eccentricity (right) from all double determinations of threshold.

30° eccentricity indices and those of the weighted full field in relation to the narrower limits of the agreement for the corresponding unweighted full-field comparison illustrate the effect of the weighting function (Table 2). This function acts to reduce the importance of the more peripheral stimulus locations in favor of the more central locations, hence, ensuring that the more peripheral information is less influential to the global indices. Such a conclusion is consistent with that of Heijl et al. and would also explain the increased
Visual Field Indices in POAG

TABLE 4. Mean of the Differences, Limits of Agreement, and Two-tailed Significance Levels (Paired t Test) for the Unweighted Indices Between 0° and 15° and Between 15° and 30° Eccentricity

<table>
<thead>
<tr>
<th></th>
<th>Unweighted (0°–15°) minus unweighted (15°–30°)</th>
<th>Corrected Pattern Deviation Fluctuation Deviation Fluctuation Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation coefficient (r)</td>
<td>0.83</td>
<td>0.08</td>
</tr>
<tr>
<td>Mean of differences</td>
<td>+2.22</td>
<td>-1.01</td>
</tr>
<tr>
<td>Limits of agreement</td>
<td>-9.64</td>
<td>+2.28</td>
</tr>
<tr>
<td>Significance level (P)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

The clinical similarity between the Program 30-2 and 24-2 indices (Table 5) suggests that it is statistically acceptable to perform serial field analysis on the data reduction indices derived from both programs. However, it is conceivable that the indices describing a given focal loss and occurring solely within the Program 24-2 stimulus grid would give proportionately larger indices compared with those of Program 30-2 because the indices themselves are expressed in relation to the total number of stimulus locations. In addition, the absence of peripheral stimulus locations clearly reduces the sampling potential of the given visual field examination. Interestingly, a comparison of the unweighted 15°–30° and the 0°–15° indices (Table 4) showed that the 15°–30° field exhibited a larger (ie, more negative) mean deviation (2.2 dB), a larger short-term fluctuation (1.0 dB), and similar values for the pattern and the corrected pattern standard deviation but wide limits of agreement for all indices. This would either suggest poor agreement between the two sets of data or would indicate that the sample comprised cases with less field loss and less intratest variability within 15° eccentricity. In conclusion, the indices derived by Programs 24-2 and 30-2 are clinically similar.

The increase in the short-term fluctuation with the increase in the number of stimulus locations involving a double determination of threshold (Fig. 3) is consistent with root mean square theory and with a decrease in patient reliability. It suggests that the short-term fluctuation calculated using the ten specified stimulus locations underestimates the short-term fluctuation in patients with primary open-angle glaucoma. Indeed, others have shown that the root mean square calculated from double determinations of threshold at ten stimulus locations only provides an estimate within 44% of the true short-term fluctuation at a 95% confidence level.14 Alternatively, it has been argued that a good estimate of the true short-term fluctuation can be obtained from knowledge of the surface generated by a single threshold determination of the entire Program 24-2 grid.15 The increase in the short-term fluctuation with the increase in the number of double determinations may also be explained by the inclusion of stimulus locations that exhibit greater variability as a result of glaucomatous damage. The threshold estimation algorithm of the Humphrey Field Analyzer retests locations that deviate from the expected value by 4 dB or more.16 However, not all test locations within a glaucomatous defect are retested because the expected value is derived from the adjacent seed points. Furthermore, the magnitude of the increased variability varies between glaucomatous patients.

The short-term fluctuation increases as a function of peripheral angle; it is influenced by the position of the stimulus location in relation to the field loss and by the depth of defect,4,17 which, in turn, is influenced by the range of stimulus intensity available for the given perimeter.18 It is clear that global intratest variability would be better estimated if all double determinations of threshold, including the ten initial test locations, were incorporated into the calculation of the short-term fluctuation. This is supported by one study in which it was found that the global fluctuation was not influenced when more than two threshold determinations were undertaken at a given stimulus location within a series of locations used to estimate the short-term fluctuation.19 In simulations of the normal response, others found that the variability of the short-term fluctuation decreased as the number of locations increased.20 However, the study has not attempted to address whether or not global short-term fluctuation should be calculated from the double determination of a subset of target locations or the double determina-
tion of all locations (i.e., both phases of the Octopus G1 program). Nevertheless, there would appear to be some advantage in using the additional information already available within the standard Program 30-2 data.

In summary, we determined that the weighting function used by the Humphrey Field Analyzer had little influence on the mean deviation but caused slight increases in the pattern and corrected pattern standard deviation and a slight reduction in the short-term fluctuation. There was little difference in the indices calculated using field data generated by Program 30-2 or 24-2. The short-term fluctuation would better reflect the intratest variability if all available double determinations of threshold, in addition to the ten standard locations, were used to calculate the index.

Key Words

glaucoma, perimetry, global indices, fluctuation

Acknowledgments

The authors thank Jason Simpson for his help in processing the data.

References