Comparison of Glaucomatous Visual Field Defects Using Standard Full Threshold and Swedish Interactive Threshold Algorithms

Donald L. Budenz, MD; Paul Rhee, OD; William J. Feuer, MS; John McSoley, OD; Chris A. Johnson, PhD; Douglas R. Anderson, MD

**Objectives:** To compare the severity, size, and depth of glaucomatous visual field defects using standard full threshold (FT), Swedish interactive threshold algorithm (SITA) standard (SS), and SITA fast (SF) algorithms of the Humphrey perimeter.

**Methods:** A prospective observational case series of 77 patients with glaucoma performed FT, SS, and SF 30-2 white-on-white testing programs on the same day on 2 occasions for 1 month. The severity of defects was compared using the mean deviation, pattern standard deviation, Advanced Glaucoma Intervention Study, and Hodapp-Anderson-Parrish severity scores. The sizes of defects were compared using the total number of abnormal points on the pattern deviation plot that fit standard criteria for glaucomatous visual field defects. The depths of the defects were compared using the sum of the threshold values for points identified in the pattern deviation plot as fitting criteria for glaucomatous defects.

**Results:** The mean deviations were slightly better using the SS (−9.6±7.1 dB) or SF (−9.1±6.7 dB) algorithm compared with the FT algorithm (−10.3±7.1 dB) (P=.005). There were no significant differences in pattern standard deviations between SS (8.6±4.0, P=.08) and SF (8.1±3.6, P=.19) compared with FT (8.3±3.3), although the pattern standard deviation was higher in SS fields compared with SF fields (P<.001). Advanced Glaucoma Intervention Study scores were slightly better when the SS (7.5±5.6) or SF (7.2±5.4) algorithm was used compared with the FT algorithm (8.6±5.4) (P<.001). The sizes of glaucomatous defects were slightly larger using the SS (20.9±10.7) algorithm compared with the FT algorithm (19.2±10.9) (P=.004) but not the SF algorithm (20.0±10.6) (P=.11). The depth of defects measured by the SS (220.4±108.0 dB) and SF (219.8±101.3 dB) algorithms was significantly shallower compared with that measured by the FT algorithm (192.3±79.1 dB) (P<.001). There were no significant differences in Hodapp-Anderson-Parrish severity scores among algorithms (P=.12).

**Conclusions:** Glaucomatous defects are measured significantly shallower using the new SITA algorithms but are approximately the same size and severity compared with FT measurements. Care should be taken when using threshold values to compare glaucomatous defects in a patient when converting from FT to SITA algorithms.

Arch Ophthalmol. 2002;120:1136-1141

©2002 American Medical Association. All rights reserved.

Visual field analysis and optic nerve visualization are critical indicators used in the diagnosis and management of glaucoma. Full threshold (FT) white-on-white automated static perimetry is the gold standard for the diagnosis, grading, and detection of progression of glaucomatous visual field defects in standard practice and in glaucoma clinical trials research. However, the standard FT method for measuring the visual field is time-consuming for patients and is subject to fatigue effect, which has been shown to result in poorer results. This effect is more pronounced in patients with glaucoma. The Swedish interactive threshold algorithm (SITA) is a new computer program that was developed for the Field Analyzer II (Humphrey Systems, Dublin, Calif) that reduces test-taking time. The SITA standard (SS) program has been shown to reduce test-taking time by approximately 50%, and the SITA fast (SF) program by approximately 70%. Overall, the number of stimuli presented in the SITA algorithms is reduced by approximately 29% in normal fields and 26% in glaucomatous fields. This is accomplished using a combination of techniques, including the use of information about surrounding points, threshold values in age-matched control subjects and patients with glaucoma at each location, changes in the pacing of the test, elimination of retest trials for the 10 points used for the calculation of short-term fluctuation in the FT algorithm, changing the way in which false-positive and false-negative re-
liability factors are determined, and using a maximum likelihood procedure for estimating threshold. The difference between the SS and SF programs is that the SF program allows a higher measurement error, permitting the determination of threshold to stop sooner than in the SS program.19

Because of the time-saving benefits of SITA, it has been suggested that this testing algorithm might replace the FT algorithm,14,18 which has been the gold standard for detecting and following up glaucomatous visual field defects for more than 15 years. However, few studies3,24 have compared the severity of glaucomatous visual field defects using the different algorithms. Potential differences between results of these 3 tests have implications for setting target intraocular pressures for the treatment of glaucoma and for determining whether patients’ defects are progressing over time. The objectives of this study were to evaluate the visual field defects in patients with glaucoma using the new SS and SF algorithms vs the standard FT algorithm and to determine whether results from these procedures can be compared in a patient during follow-up.

SUBJECTS AND METHODS

The human subjects subcommittees of the institutional review boards of the University of California, Davis, approved this study. Patients older than 17 years with known glaucoma, defined as characteristic cupping of the optic nerve and glaucomatous visual field defects in at least 1 eye, regardless of intraocular pressure level, were invited to participate. Subjects were required to be experienced visual field takers, having been tested on 2 or more prior occasions using the Humphrey visual field analyzer. Subjects were excluded if the visual acuity in the eye to be tested was less than 20/40. If both eyes had glaucoma and met inclusion criteria, the eye to be tested was selected by the investigator before initiation of the study. An attempt was made to perform testing on eyes with a wide range of visual field defects based on evaluation of prior FT testing.

After obtaining written informed consent, all subjects underwent the following visual field testing protocol during 3 visits on separate days within 1 month of each other. At each visit, a standard FT test using program 30-2 and a size III white stimulus on a white background was performed using the Humphrey Field Analyzer 1 perimeter. Calculations of the total and pattern deviation plots and global indexes (mean deviation and corrected pattern standard deviation) were performed using StatPac version 9.31 (Humphrey Systems). With the Humphrey Field Analyzer II, SS and SF tests using program 30-2 and a size III white stimulus on a white background were performed. Calculations of the total and pattern deviation plots and global indexes (mean deviation and pattern standard deviation [PSD]) were performed using StatPac for SITA version A10.1 (Humphrey Systems). The order of testing was alternated between subjects to equalize fatigue effects among testing algorithms. However, the order of tests was kept constant for each patient during all testing sessions. The subject’s best-corrected distance refraction and age-appropriate near-add power were placed in the lens holder, and the same prescription was used at each session. Subjects were required to take at least a 15-minute rest between visual field tests. Tests using a particular algorithm were performed with the same visual field machine, and each subject had the same visual field technician for all tests. Patients using topical miotics were required to be tested using miotics at all visits, and every attempt was made to perform testing at a consistent time after instillation of the miotic.

Visual fields with any abnormal reliability factor (fixation losses >33%, false-positive responses >33%, or false-negative responses >33%) were excluded. If a subject failed to produce 2 complete sets of reliable fields using any 1 algorithm within 1 month, they were excluded from the study. To be included in the study, both FT fields had to meet one of the following minimal criteria for glaucomatous visual field defects: glaucoma hemifield test results outside normal limits, corrected pattern standard deviation with a probability less than 5%, or a cluster of 3 or more points in the pattern deviation plot in a single hemifield (superior or inferior) with a probability less than 5%, one of which must have a probability level less than 1%. The first 2 complete sets of fields to meet inclusion criteria were used for analysis.

The severity of visual defects was graded using 2 standard grading systems, the Advanced Glaucoma Intervention Study (AGIS) severity scale and the Hodapp-Anderson-Parrish (HAP) grading scale (Table 1). The mean deviation, a global index that reflects the overall depression in the visual field, was compared. Also, the PSD, a global index that reflects the amount of localized (rather than diffuse) depression of the visual field, was compared for the 3 algorithms.

The size of glaucomatous defects was determined by counting the number of points in the pattern deviation plot that fit the following criteria for minimal abnormality25,26: cluster of 3 or more points on the pattern deviation plot in a single hemifield (superior or inferior) with a probability less than 5%, one of which must have a probability level less than 1%. Points that were on the edge of the 30° field, except for the 2 points at the far nasal positions above and below the horizontal meridian, were excluded because of high variability in these points.26 The depth of glaucomatous defects was determined by adding the threshold values for points identified as belonging to a glaucomatous scotoma in the pattern deviation plot, as already outlined.

All values were averaged for each individual patient for the same test, except for the HAP severity score, a categorical variable, which was treated separately for the first and second fields rather than averaged. Comparisons of continuous variables were performed using repeated-measures analysis of variance for the 3 testing algorithms, followed by paired t tests for comparison of any 2 algorithms. P values in the text are derived from paired t tests, while those in the tables are derived from repeated-measures analysis of variance. Pairwise McNemar nonparametric tests for correlated samples were used to compare HAP severity scores.

Table 1. HAP Visual Field Severity Score

<table>
<thead>
<tr>
<th>Criteria for early defect</th>
<th>Criteria for moderate defect</th>
<th>Criteria for severe defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean deviation no worse than −6 dB</td>
<td>Mean deviation worse than −6 dB but no worse than −12 dB</td>
<td>Mean deviation worse than −12 dB</td>
</tr>
<tr>
<td>On pattern deviation plot, &lt;25% of points depressed below the 5% level and &lt;15% of points depressed below the 1% level</td>
<td>On pattern deviation plot, &lt;50% of points depressed below the 5% level and &lt;25% of points depressed below the 1% level</td>
<td>On pattern deviation plot, &gt;50% of points depressed below the 5% level or &gt;25% of points depressed below the 1% level</td>
</tr>
<tr>
<td>No point within central 5° with sensitivity &lt;15 dB</td>
<td>No point within central 5° with sensitivity &lt;15 dB</td>
<td>Only 1 hemifield containing a point with sensitivity &lt;15 dB within 5° of fixation</td>
</tr>
</tbody>
</table>

*HAP indicates Hodapp-Anderson-Parrish grading scale.

©2002 American Medical Association. All rights reserved.
Seventy-seven glaucomatous eyes of 77 patients met the inclusion criteria. Forty (52%) of the 77 subjects were men and 37 (48%) were women. The mean (±SD) age of the subjects was 68.0 (±10.6) years (range, 38-84 years).

Table 2 provides a summary of the overall severity of glaucomatous visual field defects for the 3 algorithms using 3 measurements: mean deviation, PSD, and AGIS score. Mean deviation was worse in the FT fields compared with those of either SITA algorithm, and S had a slightly worse mean deviation compared with SF (P<.005 for all comparisons). There were no significant differences in PSD between SS (P=.08) and SF (P=.19) compared with FT, although the PSD was higher in SS fields compared with SF (P<.001). Advanced Glaucoma Intervention Study scores were significantly lower when the SS and SF algorithms were used compared with the FT algorithm (P<.001), indicating that, using this grading scale, overall visual field severity was somewhat less with the newer algorithms.

Table 3 provides a distribution of severity classification using the HAP severity scale for the 3 algorithms. By design, all patients were required to have a visual field defect on FT but not necessarily on SS or SF testing. There were no significant differences in HAP severity scores among algorithms (FT vs SS, P=.19; FT vs SF, P=.80; and SS vs SF, P=.09).

A comparison of the size and depth of glaucomatous defects is shown in Table 4. Using the 3 algorithms, the sizes of glaucomatous defects were within 1 or 2 points of each other. The mean size of the defect on the SS algorithm was about 1.7 points larger than on the FT algorithm (P=.004), but there was no difference in size comparing SF with FT (P=.11). The defects on the SS and SF fields were significantly shallower than on the FT fields (P<.001). Figure 1 is an example of a glaucomatous visual field defect that is similar in size using the 3 algorithms but shallower when measured with the SITA algorithms vs FT.

Figures 2, 3, and 4 provide graphic representations of scotoma depths when FT, SS, and SF fields were compared with SF. These graphs show that most glaucomatous defects appear shallower when measured with the SITA algorithms compared with FT, but not when SITA fields are compared with each other. There was a statistically significant correlation between the defect depths for the FT compared with the SS fields (r²=0.47, P<.001) and for the FT compared with the SF fields (r²=0.47, P<.001), although less than half of the variance in the SITA fields was explained by the FT fields. For this reason, regression analysis would not be expected to provide an accurate conversion between field types.

Table 2. Severity of Glaucomatous Visual Field Defects*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Full Threshold (FT)</th>
<th>SITA Standard (SS)</th>
<th>SITA Fast (SF)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean deviation, dB</td>
<td>−10.3 ± 7.1</td>
<td>−9.6 ± 7.1</td>
<td>−9.1 ± 6.7</td>
<td>&lt;.001†</td>
</tr>
<tr>
<td>Pattern standard deviation</td>
<td>8.3 ± 3.3</td>
<td>8.6 ± 4.0</td>
<td>8.1 ± 3.6</td>
<td>.005§</td>
</tr>
<tr>
<td>AGIS score</td>
<td>8.6 ± 5.4</td>
<td>7.5 ± 5.6</td>
<td>7.2 ± 5.4</td>
<td>&lt;.001†</td>
</tr>
</tbody>
</table>

Table 3. Classification of Glaucomatous Visual Field Defects Using the HAP Criteria*

<table>
<thead>
<tr>
<th>Classification</th>
<th>Full Threshold (FT)</th>
<th>SITA Standard (SS)</th>
<th>SITA Fast (SF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>0</td>
<td>1 (1)</td>
<td>3 (4)</td>
</tr>
<tr>
<td>Mild</td>
<td>23 (30)</td>
<td>19 (25)</td>
<td>19 (25)</td>
</tr>
<tr>
<td>Moderate</td>
<td>17 (22)</td>
<td>14 (18)</td>
<td>16 (21)</td>
</tr>
<tr>
<td>Severe</td>
<td>37 (48)</td>
<td>43 (56)</td>
<td>39 (51)</td>
</tr>
</tbody>
</table>

Table 4. Severity of Glaucomatous Visual Field Defects

<table>
<thead>
<tr>
<th>Classification</th>
<th>Full Threshold (FT)</th>
<th>SITA Standard (SS)</th>
<th>SITA Fast (SF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>0</td>
<td>1 (1)</td>
<td>3 (4)</td>
</tr>
<tr>
<td>Mild</td>
<td>23 (30)</td>
<td>19 (25)</td>
<td>19 (25)</td>
</tr>
<tr>
<td>Moderate</td>
<td>17 (22)</td>
<td>14 (18)</td>
<td>16 (21)</td>
</tr>
<tr>
<td>Severe</td>
<td>37 (48)</td>
<td>43 (56)</td>
<td>39 (51)</td>
</tr>
</tbody>
</table>

* N = 77. Results of first of 2 tests presented. Results of second tests were the same, including McNemar χ² value. Data are given as mean ± SD. HAP indicates Hodapp-Anderson-Parrish26 grading scale; SITA, Swedish interactive threshold algorithm. For FT vs SS, P=.19; FT vs SF, P=.80; and SS vs SF, P=.09.

† Analysis of variance.
‡ All 3 methods statistically significantly different from each other.
§ SS significantly higher than SF. No significant differences found when comparing SS or SF with FT.
| FT significantly higher than SS or SF. No difference between SS and SF.

COMMENT

Two important factors that go into deciding whether a glaucomatous visual field defect is worsening are a change in the size and a change in the depth of an existing defect. In this study, the size of the defects as seen on the pattern deviation plots was similar across the 3 algorithms, with a mean difference in the number of affected points of 0.8 to 1.7. Although this was statistically significant, the small difference is not clinically meaningful. In contrast, the depth of glaucomatous defects, expressed in decibel threshold values, was significantly shallower when measured with the SITA algorithms, both statistically and clinically.

Heijl and associates24 performed a retrospective review of 31 patients with glaucoma who had performed several FT and SS tests on different occasions. The time between tests compared was not reported. The authors found no difference between the size of glaucomatous defects as measured on the total deviation or pattern deviation plots, but found that defects were shallower when measured by SS compared with FT. The authors concluded that new baseline visual fields are desirable when switching from one algorithm to another, but that "larger prospective investigations are desirable to confirm these preliminary findings."24 Bengtsson and Heijl23 performed a prospective study similar to ours and found that the number of significantly depressed points was greater for SS than for FT or SF.
Table 4. Size and Depth of Glaucomatous Visual Field Defects

<table>
<thead>
<tr>
<th>Defect</th>
<th>Full Threshold (FT)</th>
<th>SITA Standard (SS)</th>
<th>SITA Fast (SF)</th>
<th>P Value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size (No. of affected points)</td>
<td>19.2 ± 10.9</td>
<td>20.9 ± 10.7</td>
<td>20.0 ± 10.6</td>
<td>.007‡</td>
</tr>
<tr>
<td>Depth, dB</td>
<td>152.3 ± 79.1</td>
<td>220.4 ± 108.0</td>
<td>219.8 ± 101.3</td>
<td>&lt;.001§</td>
</tr>
</tbody>
</table>

*Data are given as mean ± SD. SITA indicates Swedish interactive threshold algorithm.
†Analysis of variance.
‡SS defects larger than FT. No significant difference found for other comparisons.
§SS and SF significantly shallower than FT. No significant difference found between SITA algorithms.

Figure 1. Glaucomatous visual field defect measured with full threshold (top), Swedish interactive threshold algorithm (SITA) standard (middle), and SITA fast (bottom) in the same patient on the same day. The supranasal defect outlined in the pattern deviation plot is the area under consideration. There are 15, 16, and 19 points in the area of abnormality for the 3 fields, respectively, with total threshold values of 221 dB, 317 dB, and 338 dB, respectively. This is a typical example showing similar defect sizes but with shallower defects measured with SITA compared with full threshold algorithms. GHT indicates Glaucoma Hemifield Test; FL, fixation losses; FN, false-negative responses; FP, false-positive responses; MD, mean deviation; PSD, pattern standard deviation; SF, short-term fluctuation; and CPSD, corrected pattern standard deviation.
Sharma and colleagues\textsuperscript{16} compared results of FT and SS testing and found no significant difference in defect depths between the 2 algorithms. When converting from FT to one of the newer SITA algorithms, the size of defects does not change beyond the 2 points that were found to diagnose progression in the Normal-Tension Glaucoma Study.\textsuperscript{4} However, using the newer algorithms, the depth of defects measures about 70 dB (46\%) shallower. Therefore, defects on SITA testing may appear less severe compared with FT fields if the actual threshold values are compared, a method for detecting glaucomatous progression that is highly sensitive and specific when baseline and follow-up field tests are done with the same algorithm.\textsuperscript{4}

In the present study, several methods were used to characterize the overall severity of visual field defects. The HAP severity scale is a categorical scale that factors in the size of the glaucomatous defect, depth of the defect, and proximity of the defect to fixation.\textsuperscript{27} This scale can be helpful in classifying glaucoma severity to set target intraocular pressure goals. Results of FT and SITA tests did not differ substantially using the HAP scale, indicating that the classification of defects as mild, moderate, or severe is similar between algorithms and does not require modification. The AGIS scale\textsuperscript{6} is an ordinal scale used for judging glaucomatous progression. The AGIS score was 1.1 and 1.3 points higher (worse) when patients were tested with the FT algorithm compared with the SS or SF algorithms, respectively, which means that care should be taken in judging progression based on AGIS scores when changing tests in follow-up visual fields in patients with glaucoma.

The present study showed a slight but statistically significant difference in mean deviation scores, with the SS and SF testing producing 0.7 dB and 1.2 dB, respectively, better mean deviations than FT testing. Other studies\textsuperscript{17,24} have also shown marginally (approximately 1 dB) better values for mean deviation scores using the SITA algorithms compared with FT. Using the same study design as ours, Bengtsson and Heijl\textsuperscript{23} failed to show a difference in mean deviation between algorithms in patients with glaucoma. Given the results of these studies, there appears to be little, if any, difference between the FT and SITA algorithms in mean deviation scores.

In the present study, PSD values were not significantly different, suggesting that this factor may be compared if follow-up fields are obtained with a SITA test in

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{Scattergram comparing the depth of glaucomatous defects measured with Swedish interactive threshold algorithm (SITA) standard vs full threshold testing. The higher the number of decibels, the shallower the visual field defect. Most fields have defects that are shallower on SITA standard than on full threshold testing ($r^2=0.47$).}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Scattergram comparing the depth of glaucomatous defects measured with Swedish interactive threshold algorithm (SITA) fast vs full threshold testing. The higher the number of decibels, the shallower the visual field defect. Most fields have defects that are shallower on SITA fast than on full threshold testing ($r^2=0.47$).}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Scattergram comparing the depth of glaucomatous defects measured with Swedish interactive threshold algorithm (SITA) fast vs SITA standard testing. The higher the number of decibels, the shallower the visual field defect. There is no significant difference in defect depth between these 2 algorithms ($r^2=0.35$).}
\end{figure}
a patient who has previously undergone testing with FT. This confirms results obtained by other investigators. Pattern standard deviation is one of several factors used in assessing progression of glaucoma; an increasing PSD value is a sign of increasing localized field loss, by far the most common pattern of glaucomatous progression. Pattern standard deviation may worsen through the early and middle stages of the disease, but may decrease as the entire field becomes abnormal and many locations are homogeneously close to 0 dB. Corrected pattern standard deviation, which is the PSD corrected for short-term fluctuation, is not available in the SITA algorithms and cannot be compared.

The most clinically significant difference between algorithms appears to be in the depth of glaucomatous defects, which are 46% shallower when measured with the SITA algorithms compared with FT. Shorter perimetric tests seem to result in better sensitivity of the visual field, and this difference is accentuated in areas that are abnormal, according to our results. Because progression of visual fields is often judged based on change in threshold sensitivity within or adjacent to a preexisting glaucomatous defect, care should be taken in using changes in threshold values when converting from FT to SITA perimetry. Although defective visual fields progress by enlargement, areas with reduced sensitivity commonly progress by showing deepening of the abnormality. Therefore, when practical, it may be best to compare follow-up visual fields with fields performed with the same algorithm. To benefit from the advantages of SITA, it may be valuable to establish a new baseline field using SITA performed around the same time as the follow-up field done with FT. This may provide the best follow-up information for comparison with previous visual fields and enables switching to one of the new, shorter algorithms. In the future, software may be available that permits comparison of older and newer algorithms in a patient. Until then, the clinician must use judgment and expect a certain offset in threshold values to occur when switching from a baseline series of FT fields to a fresh set of fields performed with one of the SITA algorithms. A deeper defect with SITA, or one that is the same, may be a sign of some progression and need further evaluation with a subsequent FT test to be sure that the field is not progressing.

In summary, the current study demonstrates that glaucomatous visual field defects appear similar in size with the new SITA algorithms compared with FT on pattern deviation plots, but are much shallower with the SITA algorithms. Standard severity scales yield similar results across algorithms.

Submitted January 17, 2002; final revision received April 17, 2002; accepted May 16, 2002.

This work was supported by an unrestricted grant from Research to Prevent Blindness, New York, NY. Dr Johnson received research support from Humphrey Systems.

Corresponding author: Donald L. Budenz, MD, Bascom Palmer Eye Institute, Department of Ophthalmology, University of Miami School of Medicine, 900 NW 17th St, Miami, FL 33136.

REFERENCES

19. Bengtsson B, Heijl A. SITA fast, a new rapid perimetric threshold test: descrip-
22. Olsson J, Bengtsson B, Heijl A, Roothen H. An improved method to estimate fre-
25. Katz J, Sommer A, Gaasterland DE, Anderson DR. Comparison of analytic algo-
29. Sponsel WE, Arango S, Trigo Y, Mensah J. Clinical classification of glaucoma-