Relationship of SITA and Full-Threshold Standard Perimetry to Frequency-Doubling Technology Perimetry in Glaucoma

Catherine Boden,1 John Pascual,1 Felipe A. Medeiros,1 Makoto Aihara,1,2 Robert N. Weinreb,1 and Pamela A. Sample1

PURPOSE. To compare full-threshold (FT) and SITA algorithms for standard automated perimetry (SAP) with frequency-doubling technology perimetry (FDT) in glaucoma, to help clinicians to relate results in patients who have had two or more of these tests during follow-up.

METHODS. This study was a retrospective analysis of data from a longitudinal prospective study at the University of California, San Diego. One hundred four eyes of 104 patients with glaucomatous optic neuropathy detected by optic disc stereophotographs were included. All patients had standard perimetry (SITA and FT) and FDT within 3 months of each other. Global indices, abnormality and severity using two threshold algorithms of standard perimetry were compared with FDT.

RESULTS. More eyes had normal visual fields by SAP-FT (57 eyes) than by either SAP-SITA (42 eyes) or FDT (45 eyes), although SAP-FT agreed more closely with FDT (κ = 0.54 ± 0.08) on the presence of a visual field defect than did SAP-SITA (κ = 0.34 ± 0.08). Correlations of FDT to standard perimetry global indices were similar regardless of the threshold strategy used for standard perimetry, yielding r² = 0.38–0.57 for SAP-FT with FDT, 0.36–0.54 for SAP-SITA with FDT.

CONCLUSIONS. Despite many similarities of SAP-SITA and SAP-FT, switching the standard of comparison from SAP to SAP-SITA changes the relationship to FDT with regard to visual field abnormality, but not correlations of global indices. FDT detected abnormal fields in more eyes than SAP-FT. SAP-FT tended to detect a subset of eyes found abnormal by FDT. Visual field defects may be detected more often by FDT and SAP-SITA in eyes with early visual field loss, but these two tests may not agree on which eyes show field loss in patients who undergo both tests at follow-up. (Invest Ophthalmol Vis Sci. 2005;46:2433–2439) DOI:10.1167/iovs.04-1108

Before being widely accepted in clinical practice, a new technology typically is validated against an existing standard. However, subtle changes in the existing standard may have a substantial impact on our understanding of the new technology. In this regard, a considerable amount of accumulated knowledge about standard perimetry (SAP) is based on measuring sensitivity with a full-threshold (FT) algorithm.

Validation and evaluation of new perimetric tests is conducted against standard automated perimetry (SAP). The strengths and weaknesses of these new tests can be compared with an already extensive knowledge base about SAP. More recently, the Swedish Interactive Threshold Algorithm (SITA) has been developed and tested.1–10 Both SAP-FT and SAP-SITA use the same stimulus, test the same locations, and measure increment thresholds.1–11 However, SAP-FT uses a 4-2-2 FT algorithm to acquire thresholds,1 whereas SITA uses prior knowledge to help determine thresholds.1,2 The SITA strategy capitalizes on complex mathematical modeling of the visual field to choose the initial stimulus intensity more efficiently at each location.2 In conjunction with other features of the SITA strategy, this has meant a reduction in test time1 and tighter confidence limits1 with no practical clinical loss of accuracy.1,2,4

The properties of SITA versus the original algorithm have been well-documented.1–10 Standard perimetry test results are generally similar whether the or SITA-standard algorithm is used, though some have reported minor differences.4,13 In light of the similarities between and SITA results it should not matter which threshold algorithm is used when comparing outcomes from SAP with other visual function tests. However, there are sufficient differences between algorithms that such a comparison is warranted.

The advantages of FDT for early detection are well documented.14–20 In mode, FDT shows good specificity (86%‐100%) at high sensitivity (93%).21,22 It can be sensitive for detecting glaucoma in the earlier stages of the disease.14,21–23 FDT global indices correlate highly with those of SAP-FT.24 Many of the early studies evaluating FDT made comparisons to SAP-FT.14,16,24–26 FDT differs from SAP in several ways. The N-50 program used in the present study employs a high-temporal, low-spatial-frequency grating stimulus measures a contrast threshold with a Modified Binary Search (MOBS) algorithm and tests fewer locations than SAP. Frequency-doubling perimetry is believed to measure the integrity of the parasympathetic ganglia cells that project to the magnocellular layers of the lateral geniculate nucleus,27–29 whereas white-on-white perimetry is not specific to any type of retinal ganglion cell.26,30 Therefore, FDT is not expected to agree completely with standard perimetry, but the relationship between standard perimetry and FDT should be the same for the and SITA algorithms.

To our knowledge, the results of all three tests have not been compared in the same patients. In the present study, we sought to replicate previous studies evaluating FDT perimetry.
TABLE 1. Criteria for Defining Visual Field Severity Based on HAP Criteria for SAP31 and Modified HAP Criteria for FDT33

<table>
<thead>
<tr>
<th>Severity</th>
<th>SAP Criteria*</th>
<th>FDT Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early</td>
<td>MD &lt; -6 dB</td>
<td>Not moderate</td>
</tr>
<tr>
<td></td>
<td>Fewer than 25% of PD points P &lt; 5% and fewer than seven points below the 1% level</td>
<td>Four sectors at 1% to 5%</td>
</tr>
<tr>
<td></td>
<td>No points in the central 5° with sensitivity &lt;15 dB</td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>MD &lt; -12 dB</td>
<td>One sector at 0.5% (but no more) or more than 13 sectors at 1% to 5%</td>
</tr>
<tr>
<td></td>
<td>Fewer than 50% of PD points P &lt; 5% and fewer than 14 points below the 1% level</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No points in the central 5° with sensitivity of 0 dB</td>
<td></td>
</tr>
<tr>
<td>Advanced</td>
<td>MD &gt; -12 dB</td>
<td>More than one sector at 0.5%</td>
</tr>
<tr>
<td></td>
<td>More than 50% of PD points P &lt; 5% or more than 14 points below the 1% level</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Any points in the central 5° with sensitivity of 0 dB</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Points within the central 5° with sensitivity of &lt;15 dB in both hemifields</td>
<td></td>
</tr>
</tbody>
</table>

* SAP refers to both SAP-SITA and SAP-FT.

against SAP-FT and then to determine whether the outcome is unchanged when the standard of comparison is SAP-SITA. Patients may have a combination of FDT, SAP-FT, and SAP-SITA over time. The results of the present study can guide clinicians managing these patients to relate results from two or more of these tests. The following characteristics of the visual fields were compared: the presence of visual field abnormality, the severity of visual field loss, and the global indices.

METHODS

Patients in this study were part of an ongoing prospective longitudinal study of patients with primary open-angle glaucoma at the University of California, San Diego, Hamilton Glaucoma Center: the Diagnostic Innovations in Glaucoma Study (DIGS). This study was approved by the University of California, San Diego, Human Subjects Committee and conforms to the Declaration of Helsinki. All patients gave written informed consent to participate in the research.

Patients

Enrollment Criteria for DIGS. Each subject underwent a complete ophthalmic examination that included review of relevant medical history, best corrected visual acuity, slit lamp biomicroscopy (including gonioscopy), applanation tonometry, dilated funduscopy, medical history, best corrected visual acuity, slit lamp biomicroscopy (including gonioscopy), and fundus photography. Patients had to have a best-corrected acuity of 20/40 or better, a spherical refraction within ±5.0 D, and a cylinder within ±3.0 D. Patients were excluded if they had a history of intraocular surgery (except uncomplicated cataract surgery and glaucoma surgery), other intraocular diseases, other diseases affecting the visual field (pituitary lesions, demyelinating diseases, autoimmune disease, or diabetes), or problems other than glaucoma affecting color vision. One eye was selected randomly as the study eye, unless only one eye was eligible, in which case the eligible eye was the study eye.

Selection Criteria. The selection criteria for the study were met by 104 of 104 patients. DIGS patients who had glaucomatous optic neuropathy detected on masked stereophotograph review and had reliable SAP-FT, SAP-SITA, and FDT N-30 tests all conducted within a 3-month time window were included. Although each patient had to have all three visual field tests to be included, the visual field results were not considered for inclusion purposes.

Procedures

Visual Field Testing. SAP was conducted with a Goldmann size III (0.43°) stimulus on a 51.5- apostilb background, with both the full-threshold algorithm (SAP-FT) and the Swedish Interactive Testing Algorithm (SAP-SITA). The Humphrey 24-2 test pattern on the Humphrey Visual Field Analyzer (Carl Zeiss Meditec, Dublin, CA) was used.

The two locations nearest the blind spot were excluded from the analysis, leaving 52 test locations. All SAP fields were analyzed with Statpac software (Carl Zeiss Meditec, Inc.).

Frequency-doubling technology perimetry presents a 0.25 cyc/deg grating stimulus that counterphase flickers at 25 Hz. The N30 program of the Humphrey Instruments perimeter (Carl Zeiss Meditec) with FDT (Welch-Allyn, Skaneateles Falls, NY) was used. The 10° square stimulus is presented at 18 locations, testing to 30° nasally and 20° temporally. A circular stimulus encompassing the central 10° was also presented, although we did not include that location in the analysis.

In most eyes, SAP-FT and FDT were performed on the same day (except in 5 of 104 eyes in which FDT was performed within 3 days of SAP-FT). The order in which the tests were conducted within a session was randomized, to equalize fatigue effects across tests. SAP-SITA testing was conducted as part of a separate clinic visit. The mean ± SD number of days between SAP-FT and SAP-SITA was 22 ± 40 days (a positive value indicates SAP-FT was conducted first; 19 eyes were tested first with SAP-SITA). All patients were experienced with visual field testing, and so learning was not a factor. Only reliable (fixation losses, false-positives, and false-negatives all ≤25%) visual fields were included. The only exception to this rule was in the case of advanced visual field loss in which a false-negative rate in excess of 25% could be explained by the visual field loss alone.

A visual field was considered to be abnormal if it met one or more of the following criteria: (1) PSD with a probability of P < 5% or worse; (2) three contiguous abnormal PD points (P < 5%) in non-edge locations with at least one point having P < 1% (for FDT, the criteria was four points with P < 5%24); and/or (3) a Glaucoma Hemifield Test (GHT) result outside normal limits (except FDT).9 The GHT was not included in the criteria for abnormality for FDT, as it is not available but was included for SAP, as it is used clinically.

The severity of visual field loss was determined by Hodapp, Anderson and Parrish (HAP) criteria for standard perimetry (Table 1).31 A modification of the HAP criteria by Sponsel et al.24 was used for FDT perimetry (Table 1). In addition to abnormality and severity ratings, we compared mean defect (MD), pattern standard deviation (PSD) and the total number of abnormal PD points for FDT with SAP-FT and then SAP-SITA.

Optic Disc Stereoscopic Photographs. Simultaneous stereoscopic photographs (Simultaneous Stereo Camera TRC SS; Topcon Instrument Corporation of America, Paramus, NJ) were obtained for all patients. Simultaneous stereophotographs closest to the visual field date were examined for glaucomatous optic neuropathy. Stereophographs were taken a mean ± SD of 0.22 ± 0.78 years from the date of the SAP-FT. Each masked stereophotograph was graded independently by two experienced reviewers. Glaucomatous optic neuropathy was diagnosed if evidence of excavation, focal or diffuse rim thinning, or...
The agreement of SAP-FT with SAP-SITA was evaluated (Table 4). Eighty-nine percent (41/46) of eyes flagged as abnormal by SAP-FT were also abnormal by FDT, but only 70% (45/64) of eyes called abnormal by SAP-SITA were also abnormal by FDT. To evaluate whether there were differences in agreement depending on the criteria used to identify visual field defects, the PSD and cluster criteria were analyzed separately. Agreement on field defects by PSD and cluster analysis was moderate (Table 3). Agreement with SAP-FT was 0.61 ± 0.08 with the PSD and 0.49 ± 0.08 with cluster analysis. SAP-SITA agreement with FDT was lower but did not differ significantly from the SAP-FT results (0.53 ± 0.08, PSD; 0.40 ± 0.09, cluster criteria).

Categorization of Visual Field Severity

Fair agreement (0.37 ± 0.06; \( \kappa \) = 0.6) between FDT and SAP-FT was found with respect to severity of visual loss. Agreement was slightly lower for SAP-SITA (0.24 ± 0.06) but this difference was not statistically significant. Figure 1 shows the breakdown of FDT severity for each category of SAP severity. The difference between SAP-SITA and SAP-FT occurred in eyes with normal and early visual field loss. FDT agreed better on which eyes were normal (40 eyes) with SAP-FT than with SAP-SITA (26 eyes). Fifteen eyes that were normal by FDT were found to have early visual field loss by SAP-SITA.

Correlation of Global Indices

Mean global indices are presented in Table 5. Correlation of FDT with SAP is presented in Figure 2. FDT showed \( r^2 \) of 0.57 with SAP-FT MD, 0.55 with SAP-FT PSD, and 0.38 with SAP-FT total number of abnormal PD points. Similar \( r^2 \) values were found for the relationship of FDT with SAP-SITA: 0.48 with MD, 0.54 for PSD and 0.36 with total number of abnormal PD points.

### Table 2. Severity of Glaucomatous Visual Field Defects According to HAP Criteria for SAP and Modified HAP Criteria for FDT

<table>
<thead>
<tr>
<th>HAP Severity</th>
<th>SAP-FT</th>
<th>SAP-SITA</th>
<th>FDT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>57 (55)</td>
<td>42 (40)</td>
<td>45  (43)</td>
</tr>
<tr>
<td>Early</td>
<td>14 (13)</td>
<td>24 (23)</td>
<td>16  (15)</td>
</tr>
<tr>
<td>Moderate</td>
<td>16 (15)</td>
<td>21 (20)</td>
<td>10  (10)</td>
</tr>
<tr>
<td>Severe</td>
<td>17 (16)</td>
<td>17 (16)</td>
<td>33  (32)</td>
</tr>
</tbody>
</table>

Data are the number of patients tested (percentage of the total group). \( N = 104 \).

### Table 3. Agreement with FDT

<table>
<thead>
<tr>
<th></th>
<th>SAP-FT</th>
<th>SAP-SITA</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSD</td>
<td>0.61 ± 0.08</td>
<td>0.53 ± 0.08</td>
</tr>
<tr>
<td>Cluster</td>
<td>0.49 ± 0.08</td>
<td>0.40 ± 0.09</td>
</tr>
<tr>
<td>Any of the above</td>
<td>0.54 ± 0.08*</td>
<td>0.34 ± 0.09*</td>
</tr>
</tbody>
</table>

Visual field defect according to the PSD (\( P < 0.05 \)) and cluster of points (at least three abnormal PD points) (\( N = 104 \)). Data are \( \kappa \ ± SE. * SAP-FT and SAP-SITA agreement with FDT significantly different (\( P < 0.05 \)).

### Table 4. Comparison of SAP-FT and SAP-SITA Visual Field Abnormality Stratified by FDT Visual Field Abnormality

<table>
<thead>
<tr>
<th></th>
<th>SAP-FT Abnormal</th>
<th>SAP-FT Normal</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abnormal eyes by FDT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAP-SITA abnormal</td>
<td>39</td>
<td>6</td>
<td>45</td>
</tr>
<tr>
<td>SAP-SITA normal</td>
<td>2</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td>18</td>
<td>59</td>
</tr>
<tr>
<td>Normal eyes by FDT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAP-SITA abnormal</td>
<td>3</td>
<td>16</td>
<td>19</td>
</tr>
<tr>
<td>SAP-SITA normal</td>
<td>2</td>
<td>24</td>
<td>26</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>40</td>
<td>45</td>
</tr>
</tbody>
</table>

Abnormality criteria for FDT was cluster analysis and PSD; for SAP, it was cluster analysis, PSD, and GHT. \( \kappa = 0.65 ± 0.11 \) for abnormal eyes; \( \kappa = 0.09 ± 0.11 \) for normal eyes. Data are the number of eyes.
Post Hoc Analysis of SAP-SITA and SAP-FT in Healthy Eyes

TSAP-SITA identified a greater number of abnormal eyes than did SAP-FT and an approximately equivalent number of eyes than did FDT. To explore this result further, specificities of SAP-SITA and SAP-FT were estimated from a post hoc analysis in a group of 39 eyes from 39 normal control subjects who had both tests within 6 months. Our normal group tended to be younger (59.8 ± 13.6 years) than the patients with glaucoma (63.4 ± 11.9 years), but this trend was not significant (Student’s t-test, $P = 0.12$). Visual field abnormality was based on

**FIGURE 1.** Severity of visual field loss according to HAP criteria.$^{31}$ The four categories of SAP severity—normal (□), early (●), moderate (■), and severe (■)—are represented. (A) SAP-FT; (B) SAP-SITA.
the same criteria as for the main analysis (GHT, PSD, and cluster criteria, as described earlier). The specificity of SAP-FT and SAP-SITA, respectively, was 77% and 72%.

### DISCUSSION

In our study, FDT detected more eyes with glaucomatous optic neuropathy as having visual defects than did SAP-FT. The two tests agreed moderately well because those deemed abnormal by SAP-FT were, for the most part, a subset of those deemed abnormal by FDT. The global indices of the two tests correlated moderately well. To put the present results into perspective, our analyses with SAP-FT as the standard of comparison support the literature, which shows the FDT finds defects earlier than SAP-FT.

Previous studies have shown a strong correlation between FT and SITA MD and PSD. Test-retest variability of FT and SITA tests are similar in patients with glaucoma and in normal control subjects. One study found differences in average MD between the threshold strategies, but these differences are probably not clinically significant. Another study did not find any differences. Heijl et al. have demonstrated that it is generally safe to switch strategies, although it is preferable to establish new baselines.

Some differences have been reported, however. The average light sensitivity is approximately 1 dB better for SITA. The normal confidence bands are slightly tighter, and the normal decrease in visual sensitivity with age may be smaller with SITA. Finally, more total and PD points were abnormal with SITA in one study, but in not another. The size and severity (by both the HAP criteria and AGIS (Advanced Glaucoma Intervention Study) scores) of glaucomatous defects with SITA are similar to SAP-FT, but defects are significantly shallower on SAP-SITA.

When SAP-SITA was used as the standard of comparison, FDT found a similar number of eyes abnormal. However, the agreement was poor between SAP-SITA and FDT, in that different eyes were deemed normal and different ones abnormal by the two tests. That is, these two tests did not agree, as well, on which eyes had field loss. SAP-SITA and FDT global indices correlated moderately well.

There are several explanations for the differences found. First, the SITA threshold algorithm could be more sensitive to early visual field loss than SAP-FT. The properties of the algorithm (e.g., tighter confidence intervals) may have increased detection in some eyes. This would explain the higher number of eyes identified with glaucomatous optic neuropathy, but it does not explain why the eyes identified differed from those found by FDT. Also, SITA uses prior knowledge in determining thresholds, which is not used by SAP-FT or by the MOBS algorithm of the FDT N-30. This knowledge used by SITA is based on models of normal and glaucomatous visual fields, relying on information about age-corrected normal values at each test point, frequency-of-seeing curves, and correlations between thresholds at different test locations. These assumptions may influence threshold results at some values, possibly introducing differences between SAP-SITA and SAP-FT and could also affect comparison with FDT. Second, SAP-SITA could show more false positives. However, a similar percentage of eyes were identified as normal using the SITA and FT algorithms in the post hoc analysis of healthy eyes, indicating that the criteria for abnormality were comparable between the two tests. This suggests that an elevated number of false positives by the SITA algorithm is not a likely explanation, although our sample for the post hoc analysis was small (n = 39). Finally, SAP-SITA could detect different types of patients than FDT. FDT and SAP measure different visual functions. SAP is nonspecific, in that most retinal ganglion cell (RGC) subtypes can respond to the stimulus. FDT is a visual-function-specific test, in that the low-spatial-frequency, high-temporal-frequency stimulus is thought to target a subset of RGCs, the magnocellular cells. In a previous study comparing SAP with a variety of visual-function-specific tests, including FDT, Sample et al. found individual differences in which perimetric test detected field loss earliest. Longitudinal data were not available for all participants in this study, but will be helpful in answering these questions when they become available.

The value of a new technology often relies on comparison with a standard. For instance, FDT is thought to detect more eyes with visual field loss in early glaucoma than the standard against which it has been most often compared—SAP-FT. If the standard changes, then the value of the new technology might be interpreted slightly differently. The present results support the utility of FDT, but perhaps in a different light when the evaluation is made with SAP-SITA. Assuming the present results accurately reflect the relationship of FDT and SAP-SITA, then FDT’s major contribution may be as a complement to SAP-SITA for detecting visual loss in eyes missed by SAP-SITA.

Updating the standard of comparison when new algorithms and/or software become available is unavoidable. Newer studies are likely to use the SAP-SITA algorithm as the reference, and it is not sufficient to know how the two standards differ from one another (thorough comparisons of SAP-SITA and SAP-FT have been done). The differences should also be taken into consideration when making decisions about the value of other technologies and when making clinical decisions about patients who have a mix of new and old technologies during follow-up.

In summary, this study shows how shifting from SAP-FT to SAP-SITA as the standard of comparison changes the relationship to FDT results. Although correlational analyses were unaffected, SAP-FT tended to flag a subset of the eyes found abnormal by FDT, whereas SAP-SITA classified a different group of eyes as abnormal than did FDT, even though the number of eyes flagged by each was similar. This means that a patient with early visual field loss is less likely to show a visual field defect on SAP-FT than on either of the other two tests (by our criteria). SAP-FT tends to flag a visual field defect in a subset of the same eyes as FDT. In contrast, SAP-SITA may tend to identify different eyes. Thus, if a patient with early visual field loss has both FDT and SAP-SITA during follow-up, the results of the two tests may not concur. If the defects are repeatable, it could mean that more eyes will be identified by using a combination of the two tests, but a longitudinal study is needed to confirm this.
References


3. Bengtsson B, Heijl A. Inter-subject variability and normal limits of the SITA standard, SITA fast, and the Humphrey full threshold

**Figure 2.** Scatterplots of FDT relative to SAP-FT and SAP-SITA global indices. (A) MD: SAP-FT (○; $r^2 = 0.57; P < 0.001$) and SAP-SITA (▲; $r^2 = 0.48; P < 0.001$); (B) PSD: SAP-FT (○; $r^2 = 0.55; P < 0.001$) and SAP-SITA (▲; $r^2 = 0.54; P < 0.001$); and (C) the number of abnormal PD points: SAP-FT (○; $r^2 = 0.38; P < 0.0001$) and SAP-SITA (▲; $r^2 = 0.36; P < 0.0001$). Linear fits are represented by solid (SAP-FT) and dotted (SAP-SITA) lines.