COMPARING SITA AND FULL THRESHOLD STRATEGIES

PAOLO CAPRIS, GIUSEPPE GATTI, GUIDO CORALLO, SILVIA ROMITI, ENRICO GANDOLFO and MARIO ZINGIRIAN

Department of Neurological Sciences and Vision, Ophthalmology B, University of Genoa, Genoa, Italy

Abstract

Purpose: The accuracy and reproducibility of the new perimetric strategy Swedish Interactive Thresholding Algorithm (SITA) in comparison with the full threshold strategy of the Humphrey Field Analyzer (HFA) were evaluated in glaucoma patients. Between-algorithm point-wise variability was studied according to defect depth and test location.

Method: One eye each of 38 glaucoma patients, with a mean age of 51 years (range 24-70 years), was tested in random order with SITA and full threshold strategies (HFA Program Central 30-2) during a first session. At a second session (about 15 days later), the two examinations were repeated in reverse order.

Results: The average mean sensitivity was significantly different with each strategy: full threshold (23.35dB) and SITA (24.11dB) (p<0.0001). The average between-algorithms point-wise difference was 3.23 ± 1.10dB (range 0-29dB). These differences were correlated with defect depth (p<0.0001) and MD (p=0.008). Between-algorithm differences were more important in the upper hemifield (p=0.007) and correlated with eccentricity (p=0.004).

Conclusions: SITA and full threshold strategies had similar levels of performance with SITA giving a 47% reduction of time. Inter-test variability was greater where sensitivity was lower and in peripheral areas.

Introduction

The standard bracketing strategy is currently the most sensitive, accurate and reproducible perimetric method for to measure the differential light sensitivity in clinical populations. These positive features of the standard bracketing strategy are unfortunately accompanied by long examination times and fatigue effects.

Swedish Interactive Thresholding Algorithm (SITA), the new threshold strategy, was developed for the Humphrey Visual Field Analyzer (HFA) with the purpose of maintaining the same accuracy, reproducibility and sensitivity of the standard bracketing strategy with a considerable saving of time.

SITA works by the continuous estimation of measurement errors of threshold values with interruption of the staircase procedures when mathematical and statistical evaluations reach predetermined error levels. The continuous adaptation of the visual

Address for correspondence: Paolo Capris, MD, Clinica Oculistica Università di Genova, ‘Ospedale S. Martino’, Largo Rosanna Benzi 10, 16132 Genova, Italy

Perimetry Update 1998/1999, pp. 11–15
Proceedings of the XIIIth International Perimetric Society Meeting, Gardone Riviera (BS), Italy, September 6–9, 1998
edited by M. Wall and J.M. Wild
field model, starting from a prior defined one, allows further time-saving by reducing the number of presentations. The elimination of false-positive catch trials and improvement in time-pacing during the examination, are other sources of reduced test duration.4,5

The reliability and reproducibility of SITA-Standard (SITA-S) strategy compared to the full threshold strategy have been studied in normal subjects and glaucoma patients. In normal subjects, test time with SITA was 50% compared with the full threshold strategy, and a comparable test-retest variability was demonstrated.

Also, a higher mean sensitivity (MS) was found with SITA. An approximately 10% better point-wise test-retest variability has been reported for SITA.9

The aim of this study was to evaluate the point-wise differences in threshold measurement and the test-retest variability between SITA-S and the full threshold strategy, according to test location and defect depth.

Methods

Thirty-eight eyes of 38 glaucoma patients with prior experience in performing automated visual field examinations (at least three tests) were included in the study. We tested 21 females and 17 males with a mean age of 51, ranging from 24 to 70 years.

The visual field mean deviation index (MD) ranged from −2 to −16dB (average MD = −5.16 ± 5.37dB) (Brusini’s glaucoma staging system: stages I-IV localized and mixed).10

Visual fields were performed in two sessions with the Humphrey 750 II Visual Field Analyzer, Central 30-2 Program. All the tests were performed by the same perimetrist and with the same instrument. The foveal threshold for both strategies and the short-term fluctuation in the full threshold test were measured. At the first session, each subject underwent the visual field test with the full threshold strategy and the SITA-S strategy with a 30-minute interval between tests. The order of test strategies within the first session was randomized. At the second session, at least 15 days later, the same examinations were performed in reverse order.

Analysis of data

We studied the differences in sensitivity for all the test points with the two strategies. We used the results of the full threshold strategy as the gold standard. We defined the differences in sensitivity (in absolute values) for all test points with the same strategy in the two sessions as the reproducibility. In order to evaluate the difference in sensitivity measured between the two tests, the average sensitivity of all the test points for all subjects was compared. The same comparison was carried out for test time (Mann-Whitney U Test) and reliability indices.

The between-algorithm point-wise difference was correlated with the absolute sensitivity in order to evaluate accuracy according to loss of sensitivity. Between-algorithm individual average differences were also correlated with MD (Spearman Rank correlation). Between-algorithm point-wise differences were evaluated according to test location. The results of the upper and lower hemifield and of the central (inside 9°), paracentral (10°-16° of eccentricity), pericentral (17°-23° of eccentricity) and peripheral (outside 20°) areas were compared.
Comparing SITA and full threshold strategies

Results

The average MS of all the full threshold and all the SITA-S examinations was 23.35 ± 8.90dB and 24.11 ± 9.22dB, respectively. This between-algorithm average difference (0.76dB) was statistically significant (p<0.0001).

The average MS of all the full threshold examinations from the first session (23.27 ± 8.72dB) was not significantly different from the corresponding value obtained with the same strategy at the second session (23.56 ± 8.94dB). On the contrary, a more important (but not significant) difference was found in the average MS of all the SITA-S examinations between the first (23.70 ± 9.21dB) and second (24.49 ± 9.24dB) sessions.

The full threshold test results showed a slightly lower, but not significant, average point-wise test-retest difference (absolute values) than SITA-S. The test-retest variability was 3.13 ± 1.44dB for full threshold and 3.96 ± 1.83dB for SITA-S.

The average difference (absolute values) in the between-algorithm point-wise sensitivity was 3.23 ± 1.10dB (range 0-29). The range of these point-wise differences was frequently very wide, as indicated by the SD values (3.47 ± 1.51dB). The point-wise absolute sensitivity of all the 3060 threshold measurements and the corresponding between-algorithm differences were inversely correlated (rs=−0.238; p<0.0001). The mean individual absolute between-algorithm difference was significantly correlated with MD (rs=−0.62; p=0.008).

Fixation losses, and false-positive and negative answers, were not significantly different between the two algorithms. The point-wise between-algorithm differences were also correlated to test location. Lower between-algorithm point-wise difference values were found in the lower hemifield. This average difference (0.36dB) was significant (p=0.007). A significant correlation (p=0.004, r²=0.91) between eccentricity areas and inter-test variability was found. The average values of the central, paracentral, pericentral and peripheral areas were 2.7 ± 0.8, 2.93 ± 0.7, 3.22 ± 0.7 and 3.4 ± 0.5, respectively. The more peripheral the test location, the greater the variability between algorithms.

All SITA tests took less time than the full threshold tests. Average test duration for SITA was 487.02 ± 108.8 seconds. This was significantly (p<0.0001) shorter than for the full threshold (920.14 ± 163.62 seconds). SITA test times were reduced by approximately 47% compared to the full threshold.

Discussion

The average MS of all examinations was slightly better with the SITA-S strategy. This difference (0.76dB; 3.3%) was lower than values obtained in normal and glaucomatous subjects by other authors.

The MS difference between algorithms was greater during the second session (0.93dB), improving the sensitivity of the SITA-S examination, possibly due to a specific learning effect for this algorithm and the shorter test duration. This better sensitivity for SITA during the second session is in agreement with the theoretically expected value of 1dB, according to the different estimation and calculation of threshold values utilized by SITA-S and the full threshold algorithm.

Inter-test variability of threshold values obtained with the full threshold strategy was smaller than, but not significantly different from, SITA.

The improvement in MS during the second session of SITA was not accompanied by a
significant modification of point-wise test-retest variability. This confirmed the good
reproducibility of SITA even though we did not find better results compared with the full
threshold strategy, as shown by other authors7,9.

The average point-wise between-algorithm difference was 3.23 ± 1.10dB. This value
was similar to the test-retest variability of both SITA and full threshold strategies. The
between-algorithm differences were correlated with the point-wise light sensitivity values;
variability increased as sensitivity decreased. This phenomenon is explained by the greater
fluctuation in areas of lower sensitivity.

Between-algorithm variability was more pronounced in more damaged visual fields, as
shown by correlation with MD. The lower between-algorithm variability in the lower
hemifield is explained by the better sensitivity of these areas in glaucoma patients. As
expected, the between-algorithm variability was related to visual field eccentricity; greater
values were present in the periphery where the fatigue effect is more influential, as these
examination points are tested last.

The best values were found in the macular area (10°), and the greatest variability in the
area outside 23° (which is excluded by the Central 24-2 Program).

The shorter examination times of SITA appeared to influence the accuracy of threshold
measurement. The analysis of reliability indices revealed good comparability between FP
and FN evaluation, even though a new method of estimation of false responses was adopted
by SITA.

The comparison of SITA-S and full threshold strategies allowed us to confirm the
acceptable accuracy and similar reproducibility of the new algorithm, and the unquestionable
test time-saving of about 47% in our study. Between-algorithm differences were more
important where loss of sensitivity was present and related to eccentricity. We also observed
that patients’ favorable comments of a short examination, when first tested, were not always
repeated after the second session. This is in agreement with a supposed specific learning
effect, which can explain the improvement in MS between the two SITA sessions.

We found that SITA had an acceptable accuracy and reproducibility. There was a wide
range of between-algorithm variability, which was more consistent in the periphery. SITA
appears acceptable for use in the detection of glaucomatous abnormalities. Its use in the
follow-up of glaucoma patients has not yet been established.

References

2. Hudson C, Wild JM, O’Neill EC: Fatigue effect during a single session of automated perimetry. In-
1995.
4. Olsson J, Rootzén H, Heijl A: Maximum likelihood estimation of the frequency of false positive and
false negative answers from the up-and-down staircases of computerized threshold perimetry. In:
Ghedini Publ 1989
5. Olsson J, Bengtsson B, Heijl A, Rootzén H: Improving estimation of false-positive and false-negative
Comparing SITA and full threshold strategies


