Diagnostic sensitivity of fast blue-yellow and standard automated perimetry in early glaucoma: a comparison between different test programs.

Bengtsson, Boel; Heijl, Anders

Published in:
Ophthalmology

DOI:
10.1016/j.ophtha.2005.12.028

2006

Ophthalmology, 2006,
vol 113:7,p 1092-7,

Title: Diagnostic sensitivity of fast blue-yellow and standard automated perimetry in early glaucoma: A comparison between different test programs

Type of submission: Manuscript

Authors: Boel Bengtsson PhD, Anders Heijl MD, PhD

Affiliation: Department of Clinical Sciences, Ophthalmology
Malmö University Hospital, Lund University
SE 205 02 Malmö, Sweden

Financial support: Swedish Research Council (K2005-74X-10426-14A, K2005-74BI-15375-01A), Carl Zeiss Meditec, Dublin, CA, and funds administered by Malmö University Hospital

Conflict of interest: Our department receives research funding from Carl Zeiss Meditec
Both authors are consultants to Carl Zeiss Meditech, but have no financial interest in SITA SWAP

Running head: Diagnostic sensitivity of blue-yellow and standard perimetry

Abstract

Purpose: To compare the ability of fast Swedish interactive threshold algorithm (SITA) short-wavelength-automated-perimetry (SWAP), lengthier Full Threshold SWAP, and Standard Automated Perimetry (SAP) to detect early glaucomatous visual field loss.

Design: Cross-sectional, prospective study of perimetric diagnostic sensitivity as defined by reference limits determined in the same healthy subjects for all three test programs.

Participants: 101 patients with ocular hypertension, or suspect or early manifest glaucoma

Methods: One eye of each patient was tested with two blue-yellow perimetric programs, SITA SWAP and Full Threshold SWAP, and the SAP SITA Fast program.

Main outcome measures: glaucomatous visual field loss defined as number of significantly depressed test point locations, or cluster of such test points.

Results: No significant difference in number of significantly depressed test point locations between the three programs could be detected, neither at the $p<5\%$ limit nor at the $p<2\%$ limit. The difference of number of points depressed below the 5^{th} percentile was 0.5 between Full Threshold SWAP and SITA SWAP, 1.09 between Full Threshold SWAP and SAP, and 1.04 between SITA SWAP and SAP. The number of eyes showing clusters of significantly depressed points was also similar with the three test programs; Full Threshold SWAP identified clusters in 66 eyes, SITA SWAP in 67, and SITA Fast SAP in 65 eyes. Average test time was 12.0 minutes using Full Threshold SWAP, 4.1 minutes with SITA SWAP, and 3.5 with SITA Fast.

Conclusion: SITA SWAP identified at least as much glaucomatous visual field loss as the older Full Threshold SWAP although test time was considerably reduced. Conventional SAP using SITA Fast was not significantly less sensitive than either of the two SWAP programs.

Short-Wavelength Automated Perimetry (SWAP) has repeatedly been reported to be able to detect glaucomatous visual field defects earlier than conventional Standard Achromatic Perimetry (SAP) [1,2,3,4]. SWAP has never become widely used in clinical practice, however, possibly because older threshold algorithms available for SWAP were quite time-consuming and tiring for patients [5]. Large inter-subject variability was another problem with SWAP [4] resulting in wide normal limits for SWAP threshold values, which meant that shallow defects were not identified as significantly depressed in probability maps [6].

We developed and optimized the Swedish interactive threshold algorithm (SITA) threshold algorithms[7] for SWAP with the aim of constructing a rapid and clinically useful perimetric test for early detection of glaucomatous visual field loss [8]. The same test parameters as those incorporated in the SWAP test of the Humphrey Field Analyzer (Carl Zeiss Meditec, Dublin, CA) were applied, i.e., 440 nm narrow band blue Goldmann size V stimuli, exposed against a bright (100 Cd/m^2) yellow background. SITA SWAP reduced test times by approximately 70% compared to of the older Full Threshold SWAP. Threshold reproducibility was similar when analyzed in a mixed group of normal subjects and glaucoma patients. Normal limits for SITA SWAP threshold measurements were narrower than for the older SWAP test, by about 14% at the pointwise 5^{th} percentiles when calculated in the same set of normal subjects [9].

The purpose of this study was to compare diagnostic sensitivity at predefined levels of specificity as defined by empirically calculated normal limits determined in a group of healthy subjects, to

glaucomatous field loss of the new SITA SWAP, the older Full Threshold SWAP, and conventional SAP as assessed by SITA Fast in patients with suspect or early glaucoma.

Methods

Patients

All patients regularly visiting the glaucoma outpatient department at Malmö University Hospital were identified and listed alphabetically according to the first letter of the surname. Patient records were retrieved starting with the letter A. Patients no older than 80 years of age having diagnoses of suspect glaucoma including ocular hypertension or manifest glaucoma, with no more than slight cataract, all lens grading ≤ 2 according to the Lens Opacity Classification System (LOCS II)[10], and no known tritan color deficiency were eligible for inclusion in this study. Eyes with moderate to severe field defects were excluded; Mean Deviation (MD) values of the most recent field prior to enrollment had to be better than –5 dB. Patients with a history of neurological disease or intra-ocular eye surgery except for cataract extraction, diabetes or other systemic diseases, or who were using medications known to affect the field or color vision were not eligible.

Ocular hypertension patients had > 1 intraocular pressure (IOP) measurement > 24 mmHg without clear signs of glaucoma on the disk or on SAP in either eye. Patients with suspect glaucoma had questionable or glaucomatous optic disk changes, e.g., thin, notched, or absent neural rims or marked vertical optic cup asymmetry, without reproducible field loss on SAP in one or both eyes. Patients with manifest glaucoma had repeatable SAP visual field loss with the Glaucoma Hemifield Test (GHT) [11] "outside normal limits" in one or both eyes.

The first 110 patients meeting all eligibility criteria were invited to participate. All included patients gave informed consent. The Committee for Research Ethics at Lund University approved the test protocol and the tenets of the Declaration of Helsinki were followed.

Examination

All patients accepting the invitation came for two visits. At the first visit, history of ocular disease or surgery was taken, and visual acuity and refraction were determined. Patients underwent color vision testing using Standard Pseudoisochromatic Plates part 3 (SPP3; Igaku-Shoin, Tokyo, Japan), dilated fundus examination, and lens grading according to LOCS II. One eye of each patient underwent perimetry; if both eyes were eligible one was randomly chosen. Three visual field tests, SITA SWAP, Full Threshold SWAP, and SITA Fast were performed, all using the 24-2 test point pattern. Test order for the three visual field programs was randomized so that approximately one third started and ended with each of the three tests. At least 10 minutes of rest was provided between visual field examinations. Before the SWAP test, patients were placed in front of the perimeter facing the intense yellow background for 5 minutes to allow for retinal adaptation. At the second visit the same three visual field tests were performed, in the same test order as at the first visit; no other procedures were performed at that visit.

Analysis of visual field test results

All subjects had experience of SAP, but less than 50% of SWAP, prior to inclusion. Therefore visual fields examinations obtained at the first visit were considered to be untrained fields; only visual field data from the second visit were analyzed.

No visual fields were excluded due to poor reliability as expressed by frequencies of fixation losses as obtained with the blind spot method, or false positive and false negative answers. Fixation loss monitoring using the blind spot technique does not perform well in SWAP, since in SWAP a Goldmann stimulus size V is used. A large percentage of subjects report seeing the large bright stimulus when it is exposed in the blind spot, even if fixation is perfect when checked on the screen monitor (unpublished observations). High frequencies of false negative answers are rather an indicator of pathology than of patient reliability [12,13,14], particularly in test programs with large intra-test threshold variability as SWAP. High frequencies of false positive answers seriously affect test results, but were not used as an exclusion criterion for the purposes of this study in order to more completely compare the applicability of the three test programs.

All visual fields were analyzed on the basis of the Statpac pattern deviation values, and the corresponding pattern deviation probability maps. The reference limits determining the probability levels for these maps were derived from a normal database originally collected to compare inter-subject variability between SITA SWAP and Full Threshold SWAP using SITA Fast SAP as a reference test [9]. The 53 healthy subjects included in this normal database allowed calculation of empirically derived normal limits for total deviation and pattern deviation at the $p<5\%$ and $p<2\%$ levels, but not at the $p<1\%$ or $p<0.5\%$ levels found in commercially available databases. More important is the fact that the perimetric normal database included tests with all three test programs: SITA SWAP, Full Threshold SWAP, and SITA Fast that had been collected on the same day in the same healthy subjects, thereby minimizing sampling error uncertainties.

Thus, the older Statpac limits for Full Threshold SWAP and SITA Fast commercially implemented in the Humphrey Field Analyzer were not applied.

Comparison of visual field results

Full Threshold SWAP, SITA SWAP and SITA Fast SAP fields of all patients included were compared in terms of number of test point locations showing significantly depressed threshold sensitivity at the p<5% and the p<2% levels in the pattern deviation probability maps, Fig. 1. Sensitivity to glaucomatous fields loss was compared in three ways:

1. By comparing median number of significantly depressed points per eye and test program using Kruskal-Wallis test.
2. By paired comparisons of number of significantly depressed points between Full Threshold SWAP – SITA SWAP, Full Threshold SWAP – SITA Fast, and between SITA SWAP and SITA Fast. The distributions of paired differences were bell shaped and approximately Gaussian distributed, allowing statistical inference using t-tests and Bonferroni correction to adjust for multiple comparisons, and also calculation of statistical power.
3. By comparing number of eyes with clusters of significantly depressed test point locations, defined as three or more adjacent points, with at least one point depressed at the p<2% level, requiring that cluster points be located in the same upper or lower hemifield. This criterion is similar to the one suggested by Katz and co-workers 1991 [15].

In all patients, the last visual field taken before enrollment was a 30-2 SAP SITA Standard field. These fields were obtained on the average 6 months, ranging from 0 to 24, prior to the study fields. The Mean Deviation values of these pre-study fields ranged from –4.99 to +1.75 dB, with an average of –1.85 dB.

Results

Of the 110 invited patients 101 came and completed all procedures outlined in the protocol. All 101 subjects were included in the study and subsequently analyzed. No subjects were excluded due to cataract or to tritan color deficiency. Sixty-eight women and 33 men with a mean age of the 70 years, ranging from 42 to 80, were included.. Sixty-three patients had manifest glaucoma including 22 fellow eyes with no prior visual field loss. Of these 22 eyes, 15 had suspect or pathological disks. Twenty patients were glaucoma suspects without prior reproducible field loss, and 20 subjects had ocular hypertension.

There were no significant differences in number of depressed points between the three programs. The median number at the $p<5\%$ limit was 9 for both Full Threshold SWAP and SITA SWAP, and 7 for SITA Fast SAP ($p=0.27$), and 5, 5, and 4 respectively at the $p<2\%$ level ($p=0.18$). Neither were there any significant differences when comparing individual differences, Table 1. The statistical power to detect a difference of two significantly depressed points was 0.98. The number of eyes displaying clusters of significantly depressed points was also similar between the three programs. Full Threshold SWAP identified one or more cluster in 65% of all eyes (95% confidence interval: 56 –75), SITA SWAP detected clusters in 66% (95% confidence interval: 57 –76), and SITA Fast SAP in 64% (95% confidence interval: 55 –74).

All three programs identified clusters of depressed points in 46 eyes. There were 16 eyes with clusters in SITA SWAP fields but no clusters in SAP fields, and 16 - partly different - eyes with clusters in Full Threshold SWAP fields but no clusters in SAP fields. On the other hand SAP testing showed clusters in 10 eyes, that had no clusters with either of the two SWAP programs, Fig. 2.

Among the fields collected for the purpose of this study SITA SWAP detected clusters in 15 eyes with previously normal SAP fields, defined as no cluster of significantly depressed points, Full Threshold SWAP in 16 and SITA Fast SAP in 13 such eyes.

SITA SWAP was considerably faster than the older Full Threshold SWAP. Average test time including all 101 patients was 4.1 minutes with SITA SWAP, 12.0 minutes using Full Threshold SWAP, and 3.5 with SITA Fast.

Frequencies of false positive answers were small for all three perimetric test programs. The median false positive frequency for was 1% for SITA SWAP, 0% for Full Threshold SWAP and 3% for SITA Fast SAP. Frequencies of false negative answers were highest with Full Threshold SWAP with a median of 8%, the median for SITA SWAP and SITA Fast was 0%.

LOCS grading of the lens was performed in all but those 21 eyes, which were pseudophakic. Grading of nuclear cataract ranged from 0 to 2, with a mode value of 1. The corresponding grading for cortical and subcapsular cataract also ranged from 0 to 2, with mode values of 0.

Discussion

The new SITA SWAP program was considerably more rapid and at least as sensitive to early glaucomatous field loss as the older Full Threshold SWAP. Surprisingly, there was no significant difference between conventional SAP and the two SWAP programs in diagnostic sensitivity. SAP identified slightly and non-significantly fewer depressed points than SWAP, and almost the same number of eyes with clusters of such depressed points. This applies to the comparison between SAP and SWAP in general, since SITA Fast has previously been reported to be able to identify at least as much significant glaucomatous field loss as SITA Standard and the older threshold algorithms [16,17] and to have a diagnostic sensitivity larger than 90% [18,19]. The Statpac normal limits were actually somewhat narrower for SITA Fast compared to SITA Standard[20], implying that SITA Fast might be a very sensitive SAP algorithm for detection of glaucomatous visual field loss. As SWAP is meant for detection, we believed that it should be compared with that SAP method that might possible be most sensitive, and also least time-consuming. Further, the test algorithm of SITA SWAP is more similar to SITA Fast than those applied in SITA Standard.

A number of previous studies have reported SWAP to be more sensitive than SAP to early glaucomatous visual field loss [1,2,3,4, 21], and to progression of such loss [22,23,24]. The conclusions of many of these studies, however, were based on relatively few subjects reaching the outcome [1,2,3,21,23,24]. Some studies applied SWAP criteria for abnormality that had been designed for SAP in clinical trials, e.g., by requiring the same depth of depressions in dB values for SWAP as for SAP [1,24]. This is not at all an optimal method for comparison between SAP and SWAP, since SWAP will satisfy such a criterion by chance more often than SAP because of the larger normal inter-subject variability of threshold values in SWAP testing [5]. Others used criteria taking the normal

inter-subject variability specific to each test modality into account as in the different Statpac probability maps [2,3,4,22]. This approach is better, since it includes the larger normal threshold fluctuations in SWAP, but could be misleading if normal limits are derived from different normal databases [3,21,25,26]. A database including highly motivated well-trained super-normal subjects will yield narrower normal limits and more sensitive probability maps, but will also have poorer specificity, than a normal database including a random selection of the population. These problems are, of course, well known and have been addressed in the SAFE papers [27, 28], where common criteria for visual field defects, both for SAP and SWAP, were defined and tested in a longitudinal prospective study.

A strength of the present comparison of test programs is that the same normal subjects were used to establish normal limits for all three test programs. The present study also includes a reasonably large number of patients. This makes us interpret our results, as indicating that there might be some differences in sensitivity between SWAP and SAP, but that those differences are probably smaller than what has been previously believed. Since there were eyes that showed defects with SWAP and no defects with SAP, and other eyes where the opposite was true, an (obvious) conclusion is that when more tests are done, more abnormal results will be found.

Selection and inclusion criteria are issues that could bias evaluations of diagnostic tests. For example, by including only eyes having normal SAP, as in patients with ocular hypertension, one would expect a few positive SWAP tests, just by chance. Therefore we included also patients with very mild glaucomatous field loss, and chose not to perform any analyses in subgroups. A longitudinal prospective study including patients with one or more risk factor for glaucoma

without any structural or functional damage, i.e., where all such diagnostic baseline measurements are normal, would of course be a better approach, but such studies would take many years and could require large numbers of patients to yield a sufficient number of positive outcomes. Results from a study with such a design would provide a high quality of evidence. The study design of the Ocular Hypertension Treatment Study (OHTS) [29] would be suitable for an evidence-based evaluation of SWAP. The baseline paper of OHTS [30] has reported that yearly SWAP data was collected in 7 of the 22 OHTS clinical centers. Analysis of those results may provide longitudinal results satisfying criteria for high quality of evidence.

In summary, we found the new rapid SITA SWAP testing strategy to be at least as diagnostically sensitive as the older lengthier Full Threshold SWAP program. The diagnostic sensitivity of SAP was not significantly lower than that of the two SWAP programs. Performance comparisons were based upon strategy-specific significance limits obtained from a common group of normal subjects. Results from prospective longitudinal studies are needed to allow definite conclusions regarding the temporal relationship of SWAP and SAP visual field defects over the course of glaucomatous disease. If the results of such studies showed SWAP to have distinct advantages, then the new rapid SITA SWAP could become a very useful clinical test for the detection of early glaucoma.

References

1. Sample PA, Weinreb RN. Progressive color visual field loss in glaucoma. Invest Ophthalmol Vis Sci 1992; 33:2068-71.

2. Johnson CA, Adams AJ, Casson EJ, Brandt JD. Blue-on-yellow perimetry can predict the development of glaucomatous visual field loss. Arch Ophthalmol 1993; 111:645-50.

3. Sample PA, Taylor JDN, Martinez GA, Lusky M, Weinreb RN. Short wavelength color visual fields in glaucoma suspects at risk. Am J Ophthalmol 1993; 115:225-33.

4. Wild JM, Moss ID, Whitaker D, O'Neill EC. The statistical interpretation of blue-on-yellow visual field loss. Invest Ophthalmol Vis Sci 1995; 36:1398-1410.

5. Wild JM, Cubbidge RP, Pacey IE, Robinson R. Statistical aspects of the normal visual fielding short-wavelength automated perimetry. Invest Ophtahlmol Vis Sci 1998; 39:54-64.

6. Soliman MAE, de Jong LAMS, Ismaeil AA, van den Berg TJTP, de Smet MD. Standard achromatic perimetry, short wavelength automated perimetry, and frequency doubling technology for detection of glaucoma damage. Ophthalmology 2002; 109:444-54.

7. Bengtsson B, Olsson J, Heijl A, Rootzén H. A new generation of algorithms for computerized threshold perimetry, SITA. Acta Ophthalmol Scand. 1997; 75:368-75.

8. Bengtsson B. A new rapid threshold algorithm for short-wavelength automated perimetry. Invest OphthalmolVis Sci 2003; 44:1388-94.

9. Bengtsson B, Heijl A. Normal intersubject threshold variability and normal limits of the SITA SWAP and Full Threshold SWAP perimetric programs. Invest Ophthalmol Vis Sci 2003; 44:5029-34.

10. Chylack LT, Leske MC, McCarthy D, Khu P, Kashiwagi T, Sperduto R. Lens opacities classification system II (LOCS II). Arch Ophthalmol 1989; 107:991-7.

11. Asman P, Heijl A. Glaucoma hemifield test. Automated visual field evaluation. Arch Ophthalmol 1992; 110:820-6.

12. Heijl A, Lindgren G, Olsson J. Reliability parameters in computerized perimetry. Doc Ophthalmol. 1987; 49:593-600.

13. Katz J, Sommer A. Reliability indexes of automated perimetry tests. Arch Ophthalmol 1988; 106:1252-4.

14. Bengtsson B, Heijl A. False-negative responses in glaucoma perimetry: indicators of patient performance or test reliability? Invest Ophthalmol Vis Sci 2000; 41:2201-4.

15. Katz J, Sommer A, Gaasterland DE, Anderson DR. Comparison of analytic algorithms for detecting glaucomatous visual field loss. Arch Ophthalmol 1991; 109:1684-9.

16. Bengtsson B, Heijl A. Comparing significance and magnitude of glaucomatous visual field defects using the SITA and Full Threshold strategies. Acta Ophthalmol Scand 1999; 77:143-6.

17. Wild JM, Pacey IE, O'Neill EC, Cunliffe IA. The SITA perimetric threshold algorithms in glaucoma. Invest Ophthalmol Vis Sci. 1999; 40:1998-2009.

18. Sekhar GC, Naduvilath TJ, Lakkai M, Jayakumar AJ, Pandi GT, Mandal AK, Honovar SG. Sensitivity of Swedish interactive threshold algorithm compared with standard full threshold algorotihm in Humphrey visual field testing. Ophthalmology 2000; 107:1303-8.

19. Budenz DL, Rhee P, Feuer WJ, McSoley J, Johnson CA, Anderson DR. Sensitivity and specificity of the Swedish interactive threshold algorithm for glaucomatous visual field defects. Ophthalmology 2002; 109:1052-8.

20. Bengtsson B, Heijl A. Inter-subject variability and normal limits of the SITA Standard, SITA Fast, and the Humphrey Full Threshold computerized perimetry strategies, SITA Statpac. Acta Ophthalmol Scand 1999; 77:125-9.

21. Landers JA, Goldberg I, Graham SL. Detection of early visual field loss in glaucoma using frequency-doubling perimetry and short wavelength automated perimetry. Arch Ophthalmol 2003; 121:1705-10.

22. Johnson CA, Adams AJ, Casson EJ, Brandt JD. Progression of early glaucomatous visual field loss as detected by blue-on-yellow and standard white-on-white perimetry. Arch Ophthalmol 1993; 111:651-6

23. Bayer AU, Erb C. Short wavelength automated perimetry, frequency doubling technology perimetry, and pattern electroretinography for prediction of progressive glaucomatous standard visual field defects. Ophthalmology 2002; 109:1009-17.

24.Girkin CA, Emdadi A, Sample PA, Blumenthal EZ, Lee AC, Zangwill LM, Weinreb RN. Short-wavelength automated perimetry and standard perimetry in detection of progressive optic disc cupping. Arch Ophthalmol 2000; 118:1231-6.

25. Manberger SL, Sample PA, Zangwill L, Weinreb RN. Achromatic short wavelength automated perimetry in patients with glaucomatous large cups. Arch Ophthalmol 1999; 117:1473-7.

26. Polo V, Larrosa JM, Pinilla I, Perez S, Gonzalvo F, Honrubia FM. Predictive value of short-wavelength automated perimetry. Ophthalmology 2002; 109:761-5.

27. Johnson CA, Sample PA, Cioffi GA, Liebmann JR, Weinreb RN. Structure and function evaluation (SAFE): I. Criteria for glaucomatous visual field loss using standard automated

perimetry (SAP) and short wavelength automated perimetry (SWAP). Am J Ophthalmol 2002; 134:177-85.

28. Johnson CA, Sample PA, Zangwill LM, Vasile CG, Cioffi GA, Liebmann JR, Weinreb RN. Structure and function evaluation (SAFE): II. Comparison of optic disk and visual field characteristics. Am J Ophthalmol 2003; 135:148-54.

29. Kass MA, Heuer DJ, Higginbottam EJ et al. The ocular hypertension treatment study: a randomized trial determines that topical ocular hypotensive medication delays or prevents the onset of primary open-angle glaucoma. Arch Ophthalmol 2003; 120:701-13.

30. Gordon MA, Kass MA. The ocular hypertension treatment study. Design and baseline description of the participants. Arch Ophthalmol 1999; 117:573-83.

Table 1. Mean individual differences in number of significantly depressed between the three different test programs

	Full Threshold SWAP - SITA SWAP	*Full Threshold SWAP - SITA Fast SAP*	*SITA SWAP - SITA Fast SAP*
P<5% limit	0.05 (p=0.91*)	1.09 (p=0.08*)	1.04 (p=0.07*)
P<2% limit	0.15 (p=0.71)	1.09 (p=0.06*)	0.94 (p=0.05*)

SWAP = short wavelength automated perimetry, SITA = Swedish interactive threshold algorithm, SAP = standard automated perimetry, p= probability value, n =number of patients
*only probability values <0.017 should be considered as significant a according to the Bonferroni correction for multiple comparisons

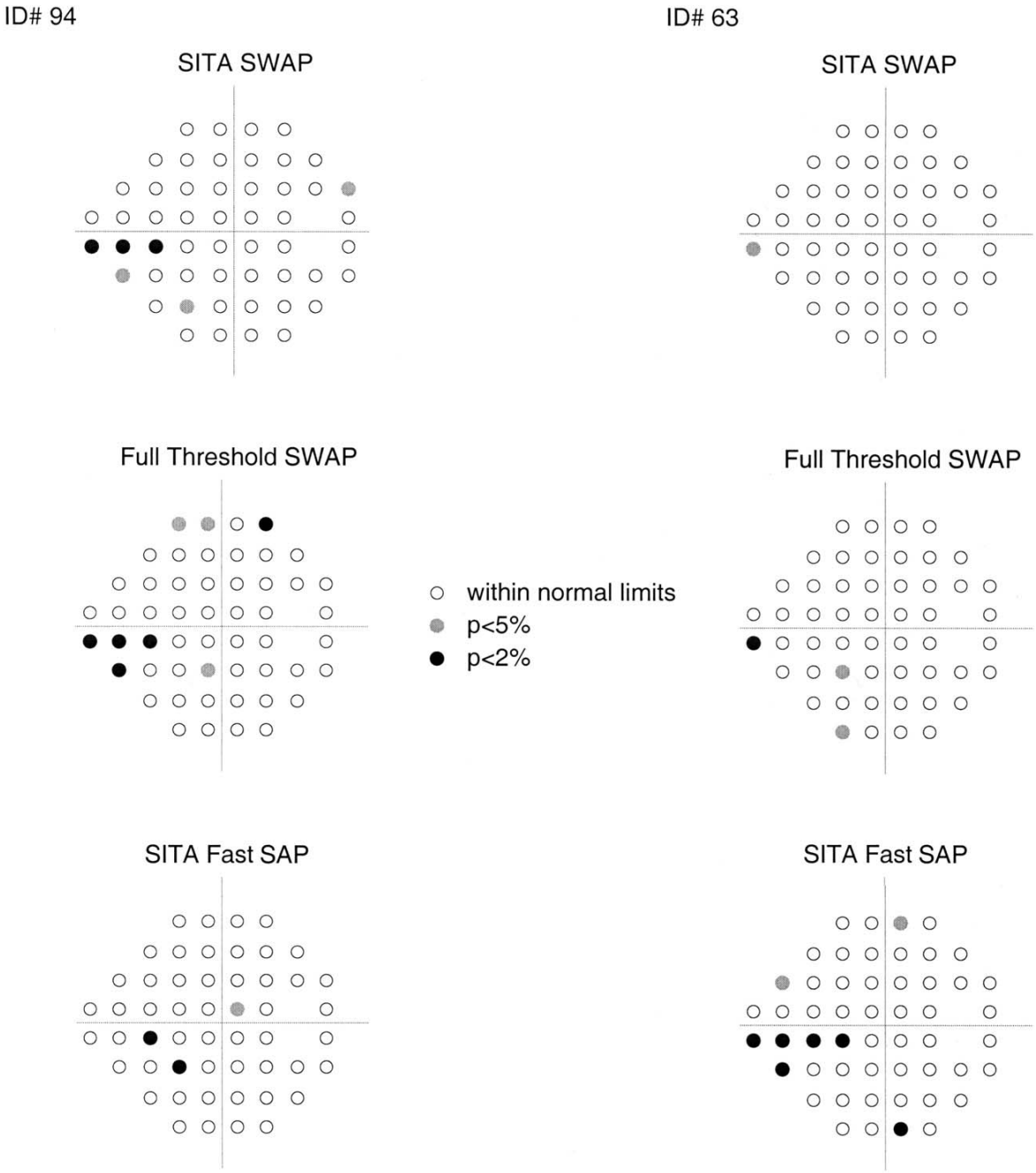

Fig.1 Visual fields of two eyes, each tested with Swedish Interactive threshold algorithm (SITA) short wavelength automated perimetry (SWAP), Full Threshold SWAP and SITA Fast standard automated perimetry (SAP). Significantly depressed points are shown are marked in grey or

black. Left: In this eye both SWAP tests detected more significant loss than SAP. Right: Here SITA Fast SAP identified more significantly depressed points than either SWAP test.

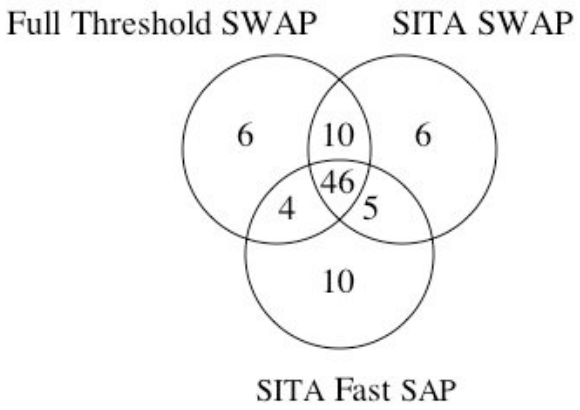

Fig. 2 Venn diagram showing number of eyes with clusters of at least 3 significantly depressed test point locations for each test program.