ANALYTIC APPROACHES TO THE INTERPRETATION OF AUTOMATED THRESHOLD PERIMETRIC DATA FOR THE DIAGNOSIS OF EARLY GLAUCOMA*

BY *Alfred Sommer,* MD, *Christopher Duggan,* BA (BY INVITATION), *Cheryl Auer,* MS (BY INVITATION), AND (BY INVITATION) *Helen Abbey,* ScD

INTRODUCTION

THE POPULARITY OF AUTOMATED PERIMETERS HAS FLOURISHED, LARGELY BECAUSE they provide the practicing ophthalmologist with a reproducible, standardized tool for examining the visual field without need of a highly skilled and experienced technician. First generation machines generally tested the visual field with stimuli of one or more fixed intensity, duplicating in many ways the isopter and suprathreshold static strategies employed in manual perimetry.[1-5] Criteria for distinguishing normal from abnormal fields were largely empiric, machine, technique, and investigator specific, and constantly being revised.[1-6] Seemingly excellent reports,[2] in terms of sensitivity (proportion of abnormals correctly identified as abnormal) and specificity (proportion of normals correctly identified as normal), often could not be confirmed.[7] Much of this variation was undoubtedly due to: differences in the severity of field loss among "abnormal" subjects, the adequacy with which alleged "normals" were scrutinized for other conditions that might affect their visual field, the age, responsiveness, prior training, clarity of the media of those being examined, the diligence with which manual perimetry was carried out by the "reference" technique, and the basic study design.[1,7-10]

*From the Dana Center for Preventive Ophthalmology of the Wilmer Institute and the Johns Hopkins School of Public Health, Baltimore, Maryland. Supported by research grant EY03605 and core grant EY01765 from the National Eye Institute. The Humphrey Instrument Company provided the Field Analyzer used in this study.

More recently, relatively low cost machines which measure retinal "threshold" sensitivity have become popular. In theory at least, threshold testing should be more efficient at quantifying abnormalities in the visual field,[11] and therefore more effective at monitoring subtle changes in retinal and optic nerve function. As a number of investigators have clearly perceived however, "thresholds" are not absolute values.[12,13] In theoretical terms, a threshold stimulus will be detected 50% of the time.[12] Since most testing strategies arrive at the threshold value after only two or three reversals (seeing to nonseeing and vice versa) rather than hundreds, threshold values will fluctuate with repeated testing. At least three factors, in addition to chance, contribute to this variation: a learning phenomenon (as one gains familiarity with the testing procedure threshold values improve by an average of 2 decibals (dB) [0.2 log units]),[14] and short-term (minutes to hours) and long-term (days to months) fluctuations in retinal sensitivity.[15-18] Threshold values provided by routine tests are better thought of as an estimate, the true value having a reasonable probability of lying within a certain range (roughly, ± 2 to 4 dB) of the measured value.[12]

A major problem facing automated perimetry is the development of practical and useful guidelines for identifying pathologic depressions in retinal sensitivity. A number of approaches have been tried, none with consistant or uniform success. The simplest consider 1 (or more) points depressed, 5 dB (or more) as significant.[8,10,19,20] Others utilize more complex, theoretical, and empirically derived criteria, sometimes based on outside reference standards.[9,12,14,21-23]

Practicing clinicians commonly employ a more intuitive technique, accounting for the popularity of grey-scale printouts of threshold values in place of raw numerical scores. Grey-scales simplify the search for characteristic changes familiar from kinetic and suprathreshold perimetry. Unfortunately, grey-scales often impute values to unmeasured areas (by interpolation from surrounding test locations), provide no information on the range of normal variation to be expected, and fail to make maximal use of the large amount of quantitative data gathered.

Given the present status of threshold analysis, and our own group's need for better diagnostic tests of early glaucomatous field loss for investigational and clinical purposes, we undertook assessment of threshold perimetric results with three primary objectives: to avoid, as much as possible, potentially biased outside reference standards; to take maximal advantage of the rich load of quantitative data available; and to enroll the analytic capabilities of the perimeter's on-board computer in the decision-making process, using strictly defined criteria.

We approached the problem from two directions, each based on a different aspect of early glaucomatous field loss: mild, if widespread alterations occurred in neuroretinal integrity and visual processing,[24-30] and pronounced asymmetry, in the severity of visual field disturbance, particularly about the horizontal meridian.[31-36] To capitalize on the former, we explored measures of global variability (cv) of threshold measurements. To capitalize on the latter, we compared the cumulative threshold sensitivity of discrete, corresponding locations, above and below the horizontal meridian.

SUBJECTS AND METHOD

All subjects were already enrolled in our prospective nerve fiber layer (NFL) study.[28] As such, they had undergone annual clinical examinations, fundus and NFL photography, threshold-related, suprathreshold kinetic, and static manual perimetry on a Goldmann perimeter. These methods are detailed elsewhere.[28] Manual perimetry took between 45 and 60 minutes per subject, and as with all other procedures, was carried out in rigorously standardized fashion and masked to the patient's clinical status.

Threshold perimetry was carried out on the Humphrey Field Analyzer, a computer-driven, automated projection perimeter, using stimulus size III and the "Central 30-2" program on the microchip current for March through October 1984. Stimuli are presented in a random sequence with variable delays. All subjects were tested with their full refractive correction and appropriate near-add.

After locating the blind spot, the program determines threshold values for 4 "primary" points, one in each quadrant, using a staircase technique (4 dB stepwise reductions in intensity until the stimulus is no longer seen, followed by 2 dB stepwise increases until first seen).[37] Threshold values of the primary points are used to create expected values for neighboring "secondary" points, providing a time-saving starting point for establishing their thresholds. If measured threshold is 4 dB from expected, the point is retested. In our analysis, we used the average value of any retested point. A total of 76 points, equally spaced 6 degrees apart throughout the central 30 degrees of the visual field, are tested.

The Field Analyzer also generates an expected "normal" visual field contour from the most sensitive of the four primary points, decreasing expected sensitivity by 3 dB for every 10 degrees of eccentricity.

Threshold values were entered into a PDP 11/45 computer, and a variety of analyses conducted, in the search for parameters and criteria that best distinguished visual fields of normal control subjects, from visual

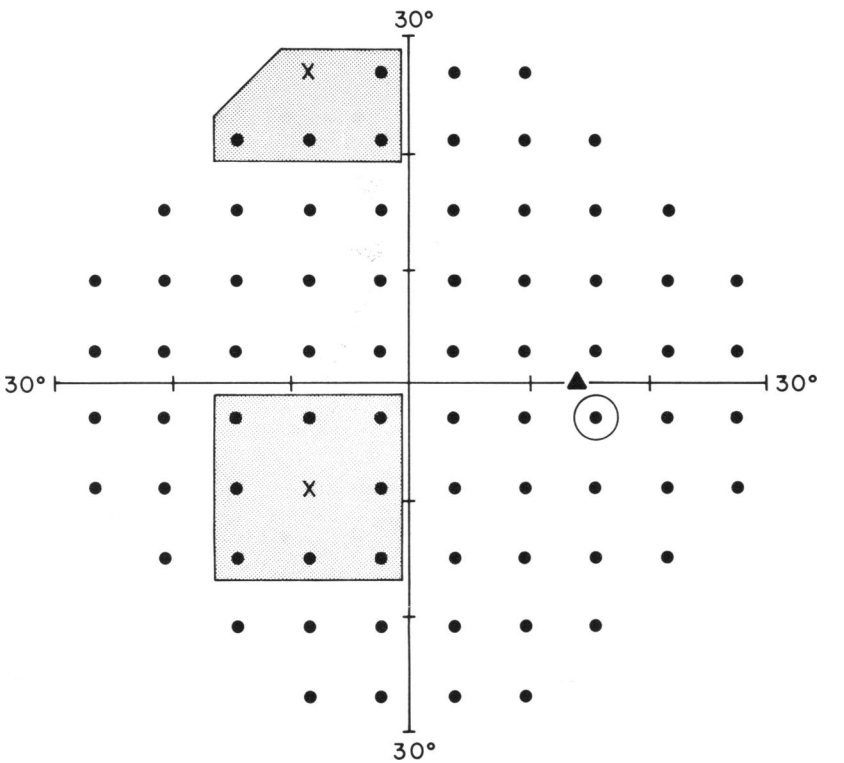

FIGURE 1
The C-30-2 program of Humphrey Field Analyzer thresholds 76 points, 6 degrees apart, in central 30 degrees. For assessing cv, threshold value of each point (excluding the single, highly variable one circled) was divided by mean threshold of 4 to 8 contiguous points (*shaded areas*). Coefficient of variation of 75 resulting ratios was calculated.

fields of patients with early glaucomatous field loss. One strategy sought to capitalize on the diffuse if subtle field disturbance in early glaucoma by examining measures of "global" variation, under the assumption, variability in retinal sensitivity is increased. The threshold of each point was divided by the average of the threshold values of all contiguous points (Fig 1). Contiguous points with 0 dB sensitivity were arbitrarily assigned a value of 1.0 dB to avoid division by 0. The mean and standard deviation of all the ratios in a given eye were calculated, and the variability for that eye expressed in several ways. The coefficient of variation (standard deviation divided by the mean ratio) provided as good a measure of

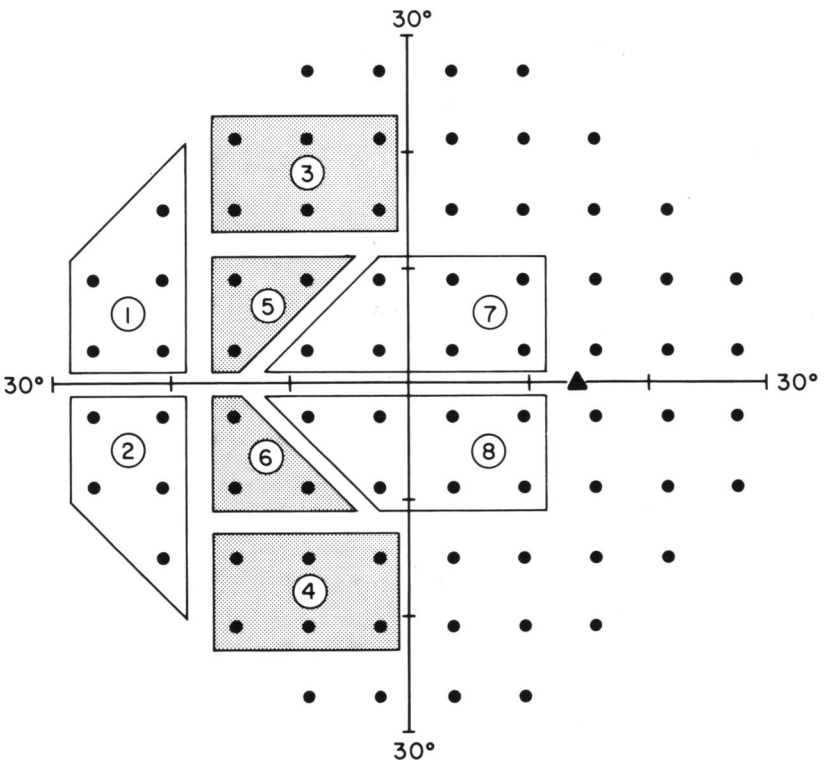

FIGURE 2
In "asymmetric strategy," total threshold values of localized groups of points were compared with corresponding areas in opposite altitudinal hemifield. After examining second set of subjects, original grid pattern was revised by removing stippled areas.

variability as any we explored. Because one point located very near the blind spot was frequently missed even by control subjects, it was ignored. The cv of each eye is therefore based on 75 ratios.

The second strategy emphasized the asymmetric nature of early field loss, especially across the horizontal meridian as exemplified by a nasal step. A variety of grids, which grouped points into localized areas of high susceptibility to early glaucomatous damage, were devised. Threshold values of corresponding areas above and below the horizontal meridian were compared. Optimal patterns (Fig 2), and corresponding differential threshold criteria for the various point-groupings, were sought.

Subjects, both abnormal and control, were studied in two sequential sets. Parameters and criteria for both strategies were developed from the

first set and "validated" on the second. The first set contained 36 control and 25 glaucomatous eyes from 45 subjects; the second set contained 105 control and 27 glaucomatous eyes from 76 subjects. Normal "controls" had intraocular pressures below 22 mm Hg, absence of gross ocular pathology, normal angles, and an entirely normal visual field by manual perimetry. "Abnormals" were patients with clinically documented elevations in intraocular pressure ($\geq$ 22 mm Hg), open angles, and mild to moderate glaucomatous field loss on manual perimetry (paracentral or full arcuate scotoma at least 0.4 log units in depth and or a nasal step of at least 10 degrees in width present to at least two isopters; patients with more advanced field loss were excluded), in the absence of other optic nerve of chorioretinal disease to account for their field disturbance (markedly tilted discs, large temporal crescents of severe peripapillary atrophy of any kind).[28]

The average age of abnormals and controls in the first group was 63 $\pm$ 12 and 57 $\pm$ 15 years, respectively, and in the second group 58 $\pm$ 14 and 51 $\pm$ 15 years.

RESULTS

Isolated or contiguous points of depressed sensitivity were common among controls (Fig 3). Even after excluding 9 points in proximity to the blind spot, 27% of all control eyes contained at least 11 points depressed 6 or more dB below the sensitivity expected from the extrapolated, normal contour; 16% contained 11 or more contiguous points depressed at least 5 dB below the expected threshold.

For the first set of subjects, a coefficient of variaion of 15 produced the best separation of glaucomatous eyes from controls (Table I). Sensitivity was over 95% and specificity 81%. Applying the same criterion to the second set of subjects produced comparable results (Table I); these could not be improved by altering the criterion.

The grid pattern developed from the first set of subjects, for detecting asymmetries in retinal sensitivity between altitudinal hemifields, is shown in Fig 2. The best separation of normals from abnormals was achieved with the criteria given in Table II, approximately 5 dB/point. Sensitivity was 96% and specificity 86% (Table II). When applied to the second set of subjects, sensitivity was only 89%. Minor modification of the grid pattern and criteria, improved results in both sets of subjects.

The cv strategy yielded a total of three false-negative eyes. All had early, mild, optic nerve damage. One was the right eye of a 66-year-old black man. In May 1983, he had a peripheral and central inferior nasal

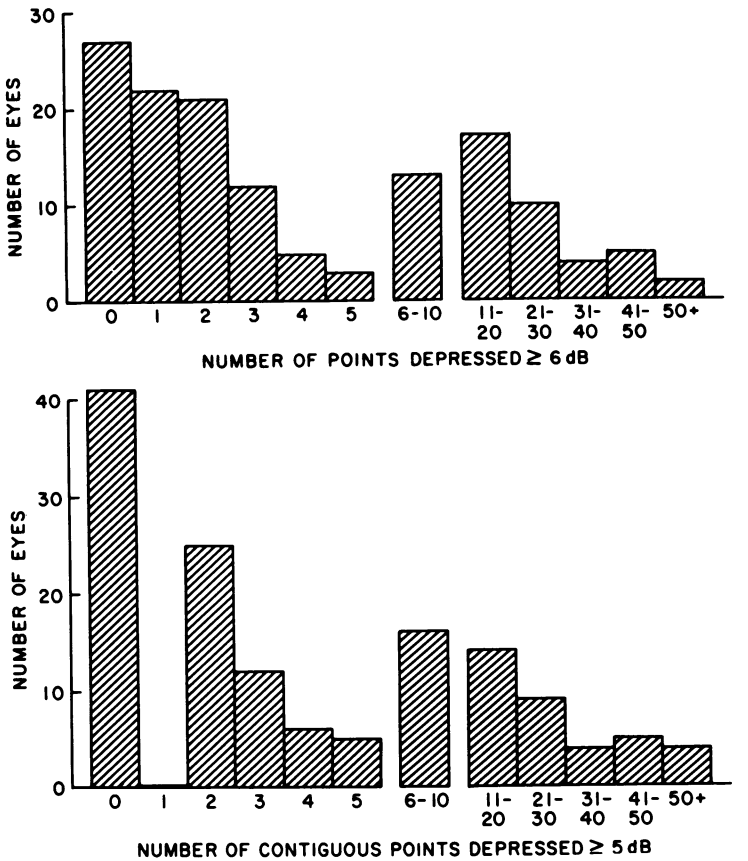

FIGURE 3
Frequency distribution of number of isolated or contiguous points of depressed sensitivity in normal control eyes. Upper graph, number of points depressed 6 dB or more below sensitivity expected from the extrapolated normal contour. Lower graph, largest number of contiguous points depressed at least 5 dB below expected threshold. The nine points adjacent to the blind spot (found to have the greatest, presumably artifactual variability) are excluded from this analysis, which is therefore based on 67 points per eye.

step; repeat manual perimetry 1 month later revealed two shallow superior paracentral scotomas. In July 1984 (the same month of his automated Humphrey examination), manual perimetry revealed a central nasal step to only 1 isopter, which no longer met our diagnostic criteria. Two independent, masked reviewers estimated his cup-disc ratio was 0.2 vertically

TABLE I: SCREENING FOR EARLY GLAUCOMA USING A GLOBAL VARIATION STRATEGY FOR INTERPRETING THRESHOLD DATA*

	GROUP I	GROUP II	TOTAL
Glaucoma eyes			
Total	25	27	52
Test positive	24	25	49
Sensitivity (%)	96	93	94
Control eyes			
Total	36	105	141
Test negative	29	89	118
Specificity (%)	81	85	84

*The coefficient of variation for the 75 ratios of the threshold value of each point divided by the mean threshold of all its (4 to 8) contiguous points. A cv $\geq$ 15 was considered abnormal.

and horizontally, his narrowest remaining neuroretinal rim as 0.4 disc diameters, and his NFL as only questionably abnormal.

The second false-negative was the right eye of a 66-year-old white man. In April 1983, manual perimetry showed only a flattening of the superior isopters; 1 month later, a repeat field revealed an inconstant arcuate defect of 0.1 log units superiorly. One year later, manual perimetry revealed a shallow (0.3 log unit) central nasal step superiorly, and 2 weeks later, a definite nasal step in the same location. Threshold fields were performed at this time. The cup-disc ratio was 0.6 horizontally and vertically; the narrowest remaining neuroretinal rim 0.2 disc diameters; and the NFL was considered normal.

The third false-negative was the right eye of a 54-year-old white man. In March 1984, manual perimetry revealed an inferior nasal step to multiple isopters between 30 degrees and 50 degrees eccentricity, but none were truly 10 degrees wide. The cup-disc ratio was 0.7 horizontally and vertically, and the narrowest rim 0.2 disc diameters. There was diffuse atrophy of the NFL.

The hemifield comparison strategy yielded the same number of false-negatives: two were false-negatives in the global strategy and one was not. The different one was the left eye of a 53-year-old white man, with pressures in the low 30s in both eyes when first seen. Manual perimetry revealed a superior arcuate elongation of the blind spot, to the vertical midline, of 0.3 log units. The abnormality was less striking on repeat field examination. The cup-disc ratio was 0.8 vertically and horizontally, with a 0.1 disc diameter of the narrowest remaining neuroretinal rim. The NFL was considered normal by one masked observer and mildly atrophied by another. Interestingly, there was a flame-shaped hemorrhage on the disc

TABLE II: SCREENING FOR EARLY GLAUCOMA USING ALTITUDINAL HEMIFIELD ASYMMETRY FOR INTERPRETING THESHOLD DATA

	ORIGINAL*	REVISED†
Criteria‡		
1-2	> 30	> 25
3-4	> 30	—
5-6	> 15	—
7-8	> 35	> 35
Sum of 1 to 8	< 630	< 630
Group I		
Sensitivity	24/25 = 96%	24/25 = 96%
Specificity	31/36 = 86%	34/36 = 94%
Group II		
Sensitivity	24/27 = 89%	25/27 = 93%
Specificity	90/105 = 86%	96/105 = 91%
Total		
Sensitivity	48/52 = 92%	49/52 = 94%
Specificity	121/141 = 86%	130/141 = 92%

*Developed from evaluation of threshold values of first set of subjects.
†Revised after evaluating threshold values of second set of subjects.
‡Difference in total threshold values of corresponding groups of points across the horizontal meridian. See Fig 2.

margin extending along the inferior arcade, without evidence of corresponding focal nerve fiber loss. The visual field defect detected by manual perimetry, quite possibly represented direct blockage of the visual stimulus by the hemorrhage and not optic nerve disease.

False-positive subjects were a mixed group: some had poor vision or nonglaucomatous pathology, while most simply seemed untestable or yielded peculiar defects of uncertain etiology (eg, dense ring scotomas broader than can be explained as lens-edge artifact). For example, a 76-year-old white man was noted to have slow responses and "low lids" in manual testing of both eyes. The left eye of a 71-year-old black man was abnormal by both threshold strategies. The cup-disc ratio was 0.5 vertically and horizontally, the narrowest remaining rim 0.3, and the NFL was normal. The technician conducting the Humphrey examination noted that the patient was incapable of taking the test: he started pressing the button before the test started and continued long after it had been completed. The left eye of a 76-year-old white woman was considered abnormal. Although she was "reliable" (3/41 losses of fixation), vision was only 20/50 in her right eye and 20/40 in her left eye. The left eye had "dry" macular degeneration and severe peripapillary atrophy (changes in the

right eye were far milder). In retrospect, manual perimetry suggested a questionable depression in sensitivity at the superior pole of the blind spot in the left eye, compared with the right eye. The field abnormalities detected by the threshold analytic strategy might well be real.

The left eye of a 51-year-old black man, with 20/25 vision and normal manual fields, is another of several instances suggesting that automated threshold examination evaluated with our analytic strategies may sometimes be more sensitive than manual perimetry. The Humphrey examination revealed a shallow superior arcuate-type depression beginning at 10 degrees, corresponding to an inferior choroidal rupture apparently related to trauma suffered in 1943. Although an accurate history had been recorded, the eye was inadvertently coded as a "normal" control; after review, it was excluded from the analysis.

DISCUSSION

No completely satisfactory method as yet exists for distinguishing visual fields of normal individuals from those with early glaucomatous optic nerve damage.

The most sensitive technique for identifying early field loss is probably one degree by one degree, manual, threshold perimetry with instruments like the Tubinger Oculus.[31] But the ponderous exactitude of this approach ensures it can only investigate a small fraction of the visual field and is too time-consuming for routine diagnostic and screening purposes. In addition, the small, subtle paracentral scotomas, considered the earliest of glaucomatous changes and for which this technique is particularly well suited, were identified in careful prospective studies of patients with ocular hypertension or early glaucoma.[32,34,38,39] Unfortunately, these studies lacked suitable controls for assessing specificity. It seems likely that many such subtle defects identified in routine testing would prove to be false-positives.

A good deal of thought and energy has gone into the development of manual screening techniques that reduce examination time and increase specificity.[40,41] The original investigations were limited to cooperative, experienced subjects with good acuity and "reliable" results, and included subjects with (readily detectable) advanced field loss.[41] Even so, compared with more detailed perimetry, sensitivity varied from 90% to 96%, and specificity from 62% to 89%.

Automated perimetry has not resolved these problems. Because of the uncertainty that surrounds every threshold value, only depressions of at least 5 or 6 dB begin to suggest true pathology.[4,6,8,10,12,19,20,30,33,42,43]

These "depressions" may be defined in relation to the values of surrounding points[9,10,22,23,30,43] or to an external standard, often generated from experienced, relatively alert subjects lacking the associated ocular pathology commonly encountered in routine practice.[9,14,21,44,45] Evaluation of threshold fields obtained with Octopus equipment, commonly follows the approach of Fankhauser and colleagues.[12,16,46] They collected exhaustive series of repeat threshold measurements on relatively small numbers of normal, highly trained subjects, thereby defining in detail a normal range of *intra*-subject variability. Statistical extrapolations were then used to formulate criteria for abnormality. As to be expected, numerous problems were experienced in applying these criteria.[12,14,15] They failed to take into account significant *inter*-subject variation, especially among typically untrained, often elderly patients, with concommitant ocular disease.

The very large number of depressed points encountered in our controls, suggests that attempting to base diagnostic criteria on isolated or even contiguous depressions of 5 or 6 dB is likely to prove frustrating.[44] Dramatic results claimed for this approach[9,10] have not withstood careful scrutiny.[7,8]

With experience, Heijl, Fankhauser and others have restricted their criteria for field loss in an attempt to improve specificity[22,23] and begun to modify and refine theoretical approaches in light of empirical findings.[14] Until an entirely different and more effective measure of ganglion cell function emerges, we will continue to seek an optimal balance between the number of true positive and false positives, sensitivity and specificity. Fankhauser put it nicely: "the diagnosis of an observed scotoma is always a statistical decision and not an absolute matter."[12]

Both of our analytic strategies yielded high sensitivity rates. The few false negatives all had early, equivocal field loss on manual perimetry. The specificity however, was disappointing, especially in comparison with some reports of suprathreshold techniques.[1,3-5,8,47] Close examination of these studies, and subsequent directions taken by many of their investigators, suggest potential shortcomings in the areas of study design, suitability of controls, definition of criteria, and the rigor with which field defects were sought by the reference manual technique.

Greve and Verduin[3] initially reported low false-positive rates for the Friedmann Visual Field Analyzer, but the following year, pointed out the limitations of its multiple spot technique and suggested it was inadequate for anything but mass screening purposes.[48] He has since concentrated on more sophisticated field examinations.[20] Notwithstanding Greve's experience and that of Drance and co-workers,[41] a recent report claims extraordinary sensitivity and specificity for this instrument.[4] However, the ref-

erence criteria are not detailed; the manual technique was apparently unstandardized, and there is considerable confusion over whether the manual "Goldmann" examination was actually the reference standard (as is suggested for the 20 "glaucoma" eyes). If it was, the "100% sensitivity" claimed for the Friedmann Analyzer on "glaucoma" eyes is inconsistent with its missing 14 of 30 "unknown," but "probable" glaucomatous eyes that met these same "Goldmann" reference criteria.

The evolution of Heijl's technique[8] and results are illuminating. using fast, 3-minute suprathreshold automated screening, he reported a sensitivity of 100% and specificity of 84% (in relation to manual "Armaly" screening). Heijl and co-workers[22] subsequently moved to more exhaustive threshold examinations. For the data-set used to develop their diagnostic threshold criteria, sensitivity was 94% and specificity 92%, even after omitting all subjects with poor fixation. Slight modification of their criteria resulted in a drop in specificity to 78%, which suggests his analytic strategy was unstable. Retesting the same subjects, using the original criteria, reduced specificity from 92% to 82%, confirming this suspicion. When subsequently comparing three different instruments, Heijl and Drance[23] tightened their criteria for the Competer. Sensitivity for the Competer, Octopus, and Perimetron ranged between 86% and 93%, and specificity between 81% and 87%, even after excluding subjects who were untrainable, who had poor fixation, or whose fields contained lens-edge artifact.

Johnson and Keltner's[1,2] techniques and criteria for the Fieldmaster have undergone continuous evolution.

The more successful of our two strategies assumed early field disturbance was asymmetric about the horizontal meridian.[32,34,35,38,49] Both Hart and Becker[34] and Mikelberg and Drance[36] have shown that initial field defects are usually limited to one or the other altitudinal hemifield. Over a 10-year follow-up period, the opposite hemifield remains uninvolved in most instances.

Comparing threshold values of corresponding areas (groups of test points) above and below the horizontal meridian, we correctly identified a large proportion of abnormal and control eyes. The pattern and criteria developed from the first group of subjects was not entirely stable. Results from the second group suggested the need for minor alterations. This approach therefore remains unvalidated, until proven stable in another group of subjects. Drance had earlier explored differences in threshold values above and below the nasal horizontal meridian.[33] Heijl employed this comparison in his initial diagnostic criteria for the Competer.[22] In a subsequent study with Drance however, Heijl abandoned this analytic

approach for the Competer, but applied it to his interpretation of Octopus data.[23]

Our other strategy assumed that early glaucomatous field disturbance is already accompanied by diffuse neuronal damage, as suggested by recent clinical and histopathologic examinations of the NFL[28,50] and by a variety of subtle psychophysical tests[24,25,27]; and that this would be accompanied by increased variability in threshold values. There is greater fluctuation and variability of threshold in glaucomatous than in hypertensive eyes, and least of all in normal eyes.[17,29,30] Indeed, data reported by Holmen and Krakau[26] suggest variability is greater in the "normal" area of a glaucomatous eye than it is in an entirely normal eye, and is greatest of all in the area of a field defect. Our results suggest this is a promising parameter for distinguishing normal from abnormal fields. Although the specificity was not as high as our other approach, our criterion was more stable. The coefficient of variation level established from the first set of subjects, produced comparable rates of sensitivity and specificity in the second; the balance could not be improved by altering the original criterion. Fankhauser and others[12] have used a number of measures of variability, primarily for establishing "expected" variation of individual points or for generating interpollated grey-scale values.

Bebie et al[15] cautions against global estimators, as these might miss small localized defects. This does not appear to have been a serious problem with the coefficient of variation, though conceivably, it forced us to use a criterion with modest specificity. Of course a more "sensitive" manual reference technique might have yielded less favorable results.

Both of our approaches to threshold analysis satisfy our initial design-requirements: they make extensive use of the data available; are independent of external standards (once general criteria are established); and analysis and decision-making can be carried out by the perimeter's on-board computer. Though not yet ideal, sensitivity was extremely high, confirming suggestions that careful assessment of the central field will uncover most instances of established field loss.[22] Specificity, however, was disappointing. It is conceivable that continued refinement of both strategies, some combination of the two, or elimination of, as yet unrecognized sources of artifact, will further improve results. It must be remembered however, that a substantial proportion of the real world, especially the elderly, are simply not suited to automated testing, at least not as it is presently performed. The flexibility and responsiveness of a well-trained and experienced technician can better deal with these problems, than can the present crop of computers. In the words of Parrish and co-workers[42]: "while the objectivity of an automated test eliminates the

influence of poor judgement (bias), it also eliminates the benefit of sound judgement."

REFERENCES

1. Johnson CA, Keltner JL, Balestrery PG: Suprathreshold static perimetry in glaucoma and other optic nerve disease. *Ophthalmology* 1979; 86:1278-1286.
2. Johnson CA, Keltner JL: Comparative evaluation of the Autofield-I, CFA-110, and Fieldmaster 101-PR automated perimeters. *Ophthalmology* 1980; 87:777-784.
3. Greve EL, Verduin WM: Mass visual field investigation in 1834 persons with supposedly normal eyes. *Albrecht Von Graefes Arch Klin Exp Ophthalmol* 1972; 183:286-293.
4. Batco KA, Anctil JL, Anderson DR: Detecting glaucomatous damage with the Friedmann Analyzer compared with the Goldmann perimeter and evaluation of stereoscopic photographs of the optic disc. *Am J Ophthalmol* 1983; 95:435-447.
5. Mills RP: A comparison of Goldmann, Fieldmaster 200, and Dicon AP2000 perimeters used in a screening mode. *Ophthalmology* 1984; 91:347-354.
6. Henson DB, Dix SM, Oborne AC: Evaluation of the Friedman Visual Field Analyzer Mark II. Part 1. Results from a normal population. *Br J Ophthalmol* 1984; 68:458-462.
7. Drance SM: Discussion of presentations by Drs Johnson et al, McCrary et al, and Li et al. *Ophthalmology* 1979; 86:1313-1316.
8. Heijl A: Automatic perimetry in glaucoma visual field screening. *Albrecht Von Graefes Arch Klin Exp Ophthalmol* 1976; 200:21-37.
9. Schmeid U: Automatic (Octopus) and manual (Goldmann) perimetry in glaucoma. *Albrecht Von Graefes Arch Klin Exp Ophthalmol* 1980; 213:239-244.
10. Li SG, Spaeth GL, Scimeca HA, et al: Clinical experiences with the use of an automated perimeter (Octopus) in the diagnosis and management of patients with glaucoma and neurologic disease. *Ophthalmology* 1979; 86:1302-1312.
11. Keltner JL: Comments on automated perimetry papers. *Ophthalmology* 1979; 86:1317-1319.
12. Fankhauser F, Bebie H: Threshold fluctuations, interpolations, and spacial resolution in perimetry. *Docum Ophthalmol Ser* 1979; 19:295-309.
13. Van Veenendaal WG, Langerhorst CT, Van Den Berg TJTP, et al: New programs for the scoperimeter, in EL Greve, A Heijl (eds): *Fifth International Visual Field Symposium*. The Hague, Dr W. Junk, 1983, pp 323-329.
14. Gloor B, Schmied U, Fassler A: Changes of glaucomatous field defects: Degree of accuracy of measurements with the automatic perimeter Octopus. *Int Ophthalmol* 1980; 3:5-10.
15. Bebie H, Fankhauser F, Spahr J: Statis perimetry: Accuracy and fluctuations. *Acta Ophthalmol* 1976; 54:339-348.
16. Koerner F, Fankhuaser F, Bebie H, et al: Threshold noise and variability of field defects in determinations by manual and automatic perimetry. *Doc Ophthalmol Proc Ser* 1977; 14:53-59.
17. Flammer J, Drance SM, Zulauf M: Differential light threshold: Short and long-term fluctuation in patients with glaucoma, normal controls and patients with suspected glaucoma. *Arch Ophthalmol* 1984; 102:704-706.
18. Drance SM, Berry V, Hughes A: Studies in the reproducibility of visual field areas in normal and glaucomatous subjects. *Am J Ophthalmol* 1966; 1:14-23.
19. Hart WM, Kolker AE: Computer-generated display for three-dimensional static perimetry: Correlation of optic disc changes with glaucomatous defects, in EL Greve, A Heijl (eds): *Fifth International Visual Field Symposium*. The Hague, Dr W. Junk, 1983, pp 43-49.

20. Greve EL, Bakker D: Some possibilities of the peritest automatic and semi-automatic perimeter, in EL Greve, A Heijl (eds): *Fifth International Visual Field Symposium*. The Hague, Dr W. Junk, 1983, pp 313-321.
21. Bebie H, Fankhauser F: Statistical program for the analysis of perimetric data. *Doc Ophthalmol Proc Ser* 1981; 26:9-10.
22. Heijl A, Drance SM, Douglas GR: Automatic perimetry (Competer): Ability to detect early glaucomatous field defects. *Arch Ophthalmol* 1980; 98:1560-1563.
23. Heijl A, Drance SM: A clinical comparison of three computerized automatic perimeters in the detection of glaucoma defects. *Arch Ophthalmol* 1981; 99:832-836.
24. Lakowski R, Drance SM, Goldthwaite D: Chromatic extrafoveal dark adapatation function in normal and glaucomatous eyes. *Mod Probl Ophthalmol* 1976; 17:304-310.
25. Atkin A, Bodis-Wollner I, Wolkstein M, et al: Abnormalities of central contrast sensitivity in glaucoma. *Am J Ophthalmol* 1979; 88:205-211.
26. Holmin C, Krakau CET: Variability of glaucomatous visual field defects in computerized perimetry. *Albrecht Von Graefes Arch Klin Exp Ophthalmol* 1979; 210:235-250.
27. Anctil JL, Anderson DR: Early foveal involvement and generalized depression of the visual field in glaucoma. *Arch Ophthalmol* 1984; 102:363-370.
28. Sommer A, Quigley HA, Robin AL, et al: Evaluation of nerve fiber layer assessment. *Arch Ophthalmol* 1984; 102:1766-1771.
29. Werner EB, Saheb N, Thomas D: Variability of static visual threshold responses in patients with elevated IOPs. *Arch Ophthalmol* 1982; 100:1627-1631.
30. Werner EB, Drance SM: Early visual field disturbances in glaucoma. *Arch Ophthalmol* 1977; 95:1173-1175.
31. Aulhorn E, Harms H: Early visual field defects in glaucoma. Glaucoma Symposium, Tutzing Castle, 1966. Basel, Karger, 1967, pp 151-186.
32. Aulhorn E, Karmeyer H: Frequency distribution in early glaucomatous visual field defects, in EL Greve (ed): *Second International Visual Field Symposium*, Tubingen, 1976. The Hague, Dr W. Junk, 1977, pp 75-83.
33. Drance SM, Fairclough M, Thomas B, et al: The early visual field defect in glaucoma and the significance of nasal steps, in EL Greve (ed): *Second International Visual Field Symposium*, Tubingen, 1976. The Hague, Dr W. Junk, 1977, pp 119-126.
34. Hart WM, Becker B: The onset and evaluation of glaucomatous visual field defects. *Ophthalmology* 1982; 89:268-279.
35. Heijl A, Lundqvist L: The location of earliest glaucomatous visual field defects documented by automatic perimetry, in EL Greve, A Heijl (eds): *Fifth International Visual Field Symposium*. The Hague, Dr W. Junk, 1983, pp 153-158.
36. Mikelberg FS, Drance SM: The mode of progression of visual field defects in glaucoma. *Am J Ophthalmol* 1984; 98:443-445.
37. *Field Analyzer Owners Manual*. Model 610. San Leandro, California, Humphrey Instruments Company, 1983.
38. Drance SM, Wheeler C, Pattullo M: The use of static perimetry in the early detection of glaucoma. *Can J Ophthalmol* 1967; 2:249-258.
39. Morin JD, Changes in the visual fields in glaucoma: Static and kinetic perimetry in 2,000 patients. *Trans Am Ophthalmol Soc* 1979; 77:622-642.
40. Armaly MF: Selective perimetry for glaucomatous defects in ocular hypertension. *Arch Ophthalmol* 1972; 87:518-524.
41. Rock WJ, Drance SM, Morgan CW: Visual field screening in glaucoma. *Arch Ophthalmol* 1973; 89:287-290.
42. Parrish RK, Schiffman J, Anderson DR: Static and kinetic visual field testing: Reproducibility in normal volunteers. *Arch Ophthalmol* 1984; 102:1497-1502.
43. Portney GL, Krohn MA: The limitations of kinetic perimetry in early scotoma detection. *Ophthalmology* 1978; 85:287-293.
44. Wilensky JT, Joondeph BC: Variation in visual field measurements with an automated perimeter. *Am J Ophthalmol* 1984; 97:328-331.

45. Brechner RJ, Whalen WR: Creation of the transformed Q statistic probability distribution to aid in the detection of abnormal computerized visual fields. *Ophthalmic Surg* 1984; 15:833-836.
46. Fankhauser F, Spahr J, Bebie H: Three years of experience with the Octopus automatic perimeter. *Doc Ophthalmol Proc Ser* 1977; 14:7-15.
47. Rabin S, Kolesar P, Podos SM, et al: A visual field screening protocol for glaucoma. *Am J Ophthalmol* 1981; 92:530-535.
48. Greve EL, Wijnans M: The statistical evaluation of measurements in static compimetry and its consequences for multiple stimulus compimetry. *Ophthalmol Res* 1972/73; 4:355-366.
49. Gramer E: Topography of early visual field defects in computerized perimetry. *Klin Monatsbl Augenheilkd* 1983; 180:515-523.
50. Quigley HA, Addicks EM, Green WR: Optic nerve damage in human glaucoma: III. Quantitative correlation of nerve fiber loss and visual field defect in glaucoma, ischemic neuropathy, pailledema, and toxic neuropathy. *Arch Ophthalmol* 1982; 100:135-146.

DISCUSSION

Dr Robert N. Shaffer. Doctor Sommer and co-workers are to be congratulated for devising a method for earlier detection of glaucoma damage using data generated by the new automated threshold perimeters. Throughout the country, investigators are demonstrating optic nerve damage far earlier than in the past. Ophthalmoscopy by red-free light will show early defects in the nerve fiber layer. Damage to nerve function can be demonstrated at increasingly early periods by color perimetry, flicker-fusion peripheral visual acuity, etc. Who of us would have dreamed that there could be a 40% to 50% diffuse loss of ganglion cells before a field loss could be detected by Goldmann perimetry? Doctor Quigley in Baltimore and Doctor Drance in Vancouver have estimated that normal eyes have an average drop-out of 5000 neurons per year beginning in the early 20s. As increasingly sensitive methods are devised, even this normal decrease in retinal sensitivity may become detectable. Will that patient then be considered to have early glaucoma?

Doctor Sommer's technique of comparing corresponding groups of test points above and below the horizontal meridian, achieved an admirable sensitivity of 95%. The specificity was a somewhat disappointing 85%. Can we base therapeutic decisions on nerve damage revealed by such subtle methods? I am reminded of our surprise in the 50s, when a long-term study of untreated ocular hypertensives showed no field loss by Goldmann perimetry in 9 of 10 cases over a 10-year period.

Some of those nine might well show damage by these more sophisticated methods. What degree of nerve damage must there be before a diagnosis of glaucoma can be confirmed? Knowing the side effects, the expense and the nuisance of drugs, when are we justified in beginning therapy? Doctor Sommer's work is important, but the course of these patients must be monitored for years before we can be sure whether or not therapy should be instituted at an earlier period than at present.

DR BRIAN R. YOUNGE. I would certainly like to congratulate Doctor Sommer and colleagues for thinking about the interpretation of threshold perimetry data and putting this on paper where we can actually work on it. We, in fact, have been doing the same kind of thing without the same formal mechanisms, and I think the title of my talk "Computer-assisted Perimetry" reflects this. I do not like "automated perimetry" as a term because I really think that we need to use clinical judgement in the interpretation of these visual fields. Many patients make errors in their responses and one of the first things we learned about computer-assisted perimetry is that the first test on a patient is less accurate; in fact it is less accurate whether it be the first time on a tangent screen or Goldmann perimeter. Another point that I would like to make is that in disease patterns or disease entities one needs to think about the nerve fiber distribution above and below the horizontal raphae; and comparing the numbers that test these loci may be useful. It is nice to be able to submit these numbers to mathematical analysis, but you need to examine the grey scales and the numbers yourself for proper interpretation. By knowing the patient's fundus and optic nerve fibers, as well as the probable disease entity, one can make an intelligent interpretation.

DR MARVIN SEARS. Mr President and Mr Secretary, I would like to commend Doctor Sommer for this wonderful presentation. Not only in this presentation but in many others Doctor Sommer has brought the wonderful world of statistics to ophthalmology and we are very grateful to him for doing that. In doing that, however, several questions come to mind. One is, I wonder about the retest strategy that you use, particularly in the global variation technique, and, whether or not there was any effect of the interval between test and retest on your subsequent result. Second question: in listening to your beautiful presentation, what was the justification for making a ratio between one particular locus in the visual field and adjacent loci that have been stimulated? Is the resultant ratio a biologic property of the eye or is it something like P_0 over $c(P_0/c)$ where the numerator and denominator are so closely related, have the same dependents, that an arithmetic correlation develops without actually addressing the underlying biologic problem. I wonder if you could speak to this.

DR ALFRED SOMMER. I would like to thank the three discussants for their kind comments and address them in turn. Doctor Younge made the important point that the first test of an eye is less accurate than subsequent tests, which is what I referred to as the learning phenomenon. In fact, it has been repeatedly shown that repeating a threshold test improves sensitivity on the average by 2 decibels. The cogent suggestion to relate our analyses to the distribution of the arcuate nerve fiber is, of course, one of the reasons we developed the cross-horizontal meridian analysis. While we also examined many other parameters, it was gratifying to find that this one in particular, which made the most biologic sense, worked best.

Doctor Sears raised two questions, one of which probably reflects some incoherence in my presentation. Our global variation strategy was not a comparison of multiple tests on the same eye. Each eye was tested only once. The global variation analysis measured the degree of variation in the field of a single test of the central 30 degrees. This eliminates the need for retesting individuals and for dealing with other factors that might influence temporal changes in perimetric results. Of course a patient with ocular hypertension or early glaucoma will be tested as many times as necessary for the clinician to feel comfortable with the results and his or her appreciation of the biology of the disease process. But we are hoping that this test would be useful for screening individuals, either in the population at large or in the average ophthalmologic practice. The vast majority of such subjects will not have glaucoma, and we didn't want to have to subject such people to repeated testing. As to the second question, why use a ratio, I have no biologically cogent answer. We applied a number of different statistical techniques and found empirically that the ratio provided the cleanest separation between abnormals and controls. It may well be that future tests, on larger numbers of cases and controls, will suggest a superior approach.

I would like to thank Professor Shaffer in particular for his careful review of the material and for his kind remarks. He raised the question that I thought more people would jump at: what do we do with all this new sophisticated material on patients who lack definite field loss or increased cupping by present testing methods, especially when we now know that many lose optic nerve tissue early in the glaucoma process. Unfortunately, we are still between the same "rock and a hard place" we have been for years. When a patient presents with high pressure we follow them very closely or initiate therapy. When a patient presents with a 0.9 cup or absence of neuroretinal rim and an elevated pressure, or with definite, reproducible visual field defects, we initiate therapy. These are not the problem cases. Th real quandary is the patient who comes in for the first time with a 0.30 to 0.6 cup and a pressure of 23 or a vague, often transient shallow scotoma in the paracentral area. Should this patient be on therapy? Quite honestly we don't have the answer. Doctor Shaffer has already alluded to the fact we will only learn the value of these new, sensitive prognostic tests by following such patients over time and learning which develop definite, persistent damage.