SUPRATHRESHOLD PERIMETRY: ESTABLISHING THE TEST INTENSITY

DAVID B. HENSON[1], PAUL H. ARTES[1], SHAILA J. CHAUDRY[1] and BALWANTRAY C. CHAUHAN[2]

[1]Department of Ophthalmology, University of Manchester, Manchester, UK; [2]Department of Ophthalmology, Dalhousie University, Halifax, Canada

Abstract

Purpose: To establish the precision of two techniques for setting the test intensity in suprathreshold perimetry.
Method: Full-threshold data (Humphrey Visual Field Analyzer (HFA) Program 30-2) from a normal database (*n*=146) and a POAG database (*n*=69) were analyzed to establish the precision of both an age-related and the HFA threshold-related technique for deriving the test intensity in suprathreshold perimetry. Computer simulation was also used to establish the precision of the HFA technique.
Results: For the normal database, there was little difference in precision between the age-related and the HFA threshold-related techniques (SD 1.318 versus 1.445). With the POAG database, the age-related technique was found to be more precise than the HFA threshold-related technique (SD 1.618 versus 1.924). Simulated results for the HFA technique compared well to the analysis of clinical data.
Conclusions: The current HFA threshold-related technique is no more precise than an age-related technique at setting the test intensity in suprathreshold perimetry. Both techniques are less precise than a semi-automated multiple stimulus technique (SD 0.86). Further research is needed to improve the precision of setting the test intensity in automated single stimulus threshold-related suprathreshold perimetry.

Introduction

Suprathreshold perimetry involves the presentation of stimuli at intensities calculated to be above the patients' threshold. If the stimuli are seen, we assume that no significant defect exists. Suprathreshold perimetry has been widely used to screen for visual field loss when speed has been of paramount importance[1-3] and the prevalence of visual field defects is low. In its original form[4] the test intensity varied only with eccentricity (to compensate for the sensitivity gradient between the center and the periphery of the visual field) with no account being taken of age or individual differences in sensitivity. Friedmann[5] proposed that the technique could be improved if the test intensity took into account the relationship between sensitivity and age[6-9]. This

Address for correspondence: D.B. Henson, PhD, Department of Ophthalmology, University of Manchester, Royal Eye Hospital, Oxford Road, Manchester M13 9WH, UK

Perimetry Update 1998/1999, pp. 243–252
Proceedings of the XIIIth International Perimetric Society Meeting,
Gardone Riviera (BS), Italy, September 6–9, 1998
edited by M. Wall and J.M. Wild

back

modified technique became known as age-related suprathreshold perimetry. Greve[10] and Gutteridge[11] pointed out that there were often considerable differences in the height of the hill of vision within a particular age group, and that the performance of suprathreshold perimetry could be further improved if it took these into account. This technique, threshold-related suprathreshold perimetry, requires a routine that establishes the appropriate test intensities at the onset of the examination.

In non-automated instruments the threshold-related approach relies upon either a staircase or bracketing technique to derive appropriate test intensities. The test sequences or termination criteria were rarely defined. This situation changed with the development of fully automated perimetry. The nature of these tests required specific test algorithms with clearly defined test sequences and termination criteria. The method adopted in the Humphrey Visual Field Analyzer (HFA) was to use the standard full-threshold strategy at four visual field locations (12.7° from fixation along the 45, 135, 225 and 315° meridians) and to base the suprathreshold test intensity on the second most sensitive of the four locations. An upper limit is placed on the intensity to guard against it being set too high in cases where sensitivity is markedly reduced at the thresholded locations.

Setting of the test intensity has important implications with respect to the sensitivity and specificity of the suprathreshold test. If the intensity is set too high, the test is likely to be insensitive, if set too low, it is likely to produce false-positive results. The precision of the HFA technique, or an age-related technique has not, however, been fully evaluated within the literature.

Full-threshold data from the HFA 24-2 Program can be used to establish a gold standard measure of sensitivity. In the data reported in this paper we will be using the seventh most sensitive location as the gold standard. Within the test pattern of the 24-2 program are the four test locations used to derive the test intensity in HFA threshold-related screening program. A single full-threshold HFA 24-2 result can, therefore, be used to calculate the precision of the HFA technique for establishing the suprathreshold test intensity.

Full-threshold HFA 24-2 data can also be used to derive the relationship between sensitivity and age. This in turn can be used to calculate the precision of an age-related technique for establishing the suprathreshold test intensity.

This paper reports on the precision of both the HFA threshold-related and an age-related technique of establishing the test intensity in suprathreshold perimetry in populations of normals and patients with primary open-angle glaucoma (POAG). The results obtained are then compared to those from a computer simulation which models both the HFA test algorithm and the responses obtained from normals and patients with glaucomatous visual field loss. This part of the paper demonstrates that the simulation can accurately predict the results obtained from the analysis of clinical data.

Methods

Analysis of clinical data

Patient data

The analysis is based upon two databases of full-threshold (HFA Program 30-2) results from the Department of Ophthalmology at Dalhousie University. The first database contains the results from 146 normal controls, while the second contains the results from 69 patients with POAG. All analyses are based upon a single test result from one randomly selected eye of each subject. All subjects had prior experience with full-threshold perimetry on the HFA.

Appropriate test intensity gold standard

This study is concerned with the precision, *i.e.*, the spread of the differences between the estimated sensitivity and a gold standard measure of sensitivity. We chose the gold standard to be the seventh most sensitive point in the 24-2 set of test locations. Choosing the seventh most sensitive value, rather than the most sensitive, guards against the influence of false-positive response errors. The seventh most sensitive value is also relatively immune to the presence of all but the most advanced visual field defects.

Age-related technique

The relationship between age and sensitivity was derived from a subset of the normal database. Because false-positive and false-negative response errors can artificially raise or lower the sensitivity estimate[12], the subset excluded patients who responded more than once to either the false-positive or false-negative catch trials (Table 1). A linear regression was performed on the variables age and sensitivity at the seventh most sensitive location. The distribution of the differences between the predicted values and observed values (residuals) was used as a measure of the techniques precision.

Threshold-related values, current HFA technique

Threshold-related values were derived from both the normal and POAG databases. The results from the four thresholded locations used for determination of suprathreshold test intensity were extracted from the full-threshold data, and the second most sensi-

Table 1. Details of databases used in this study

Database	*Size*	*Age (range, mean)*
Normals (excluding those with more than one false-positive or false-negative)	119	30-85, mean 51
Normals (all included)	146	30-85, mean 52
POAG	69	30-90, mean 63

tive derived. The distribution of the differences between this measure and the gold standard was used as a measure of the techniques precision.

Perimetric test simulation

The simulation was performed on a PC with programs written in the Delphi programming language. The programs contained a perimeter unit that simulated the normal operation of a perimeter, and a response unit that simulated responses from both normal and glaucomatous patients. The perimeter unit would randomly select one out of the four locations and calculate the appropriate stimulus intensity for the next presentation. The stimulus parameters (location and intensity) are then passed to the response unit. Depending on sensitivity, threshold variability and the probability of response errors, the response unit would then return a value of ‘seen’ or ‘not seen’. On the basis of this response, the perimeter unit would update its stored information before selecting another stimulus location.

Threshold variability at each test location was set according to the following equation:

$$\log_e (\mathrm{SD}) = -0.08 * \text{sensitivity} + 3.22$$

where sensitivity was given in dBs.

This equation is the best fitting (least squares) function to the gradient of the frequency-of-seeing (FOS) versus sensitivity data collected from a population of glaucomatous/ocular hypertensive (n=44) and normal eyes (n=22). The FOS data were collected with an adaptive procedure which presented a minimum of 20 stimuli at six or more intensities straddling the threshold (see Henson, Chaudry and Artes[13] in this volume for further details). The relationship between the gradient of the FOS curve and sensitivity was similar to that already reported by Weber and Rau[14] and Chauhan *et al.*[15]. At the beginning of each simulation (1000 simulations were run for each condition), the false-positive and false-negative response rates were set to match those of a randomly selected patient from the normal database.

The standard two-reversal 4-2 algorithm[16] was used at the four locations. The last seen response was taken as the threshold estimate and the starting level was set to 30dB. Each simulation was terminated when sensitivity estimates had been obtained at each of the four locations. The second most sensitive was taken as a predictor of the appropriate test intensity.

At the onset of each simulation, the sensitivity at the four locations was set. Its value varied, from one simulation to the next, in a manner similar to that within the normal database (SD=1.44). One thousand simulations were run for conditions in which zero, one, two and three of the four locations were defective (18dB defect).

Results

Age-related measures

Figure 1 shows the relationship between the seventh most sensitive location and age in the population of 146 normal eyes. The regression line (fitted to the subset of 119

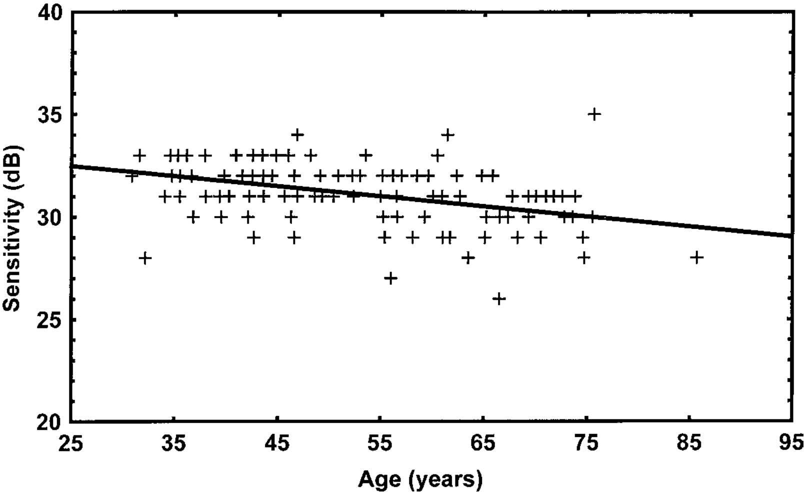

Fig. 1. Sensitivity at the seventh most sensitive location versus age in a population of 146 eyes from 146 patients. The regression line is fitted to a subset of 119 eyes in which the subjects produced no more than one false-positive or false-negative catch trial response.

eyes) has a gradient of –0.059dB/year and an intercept of 34.08dB. Figure 2a gives the distribution of errors (gold standard minus age-related estimate) for the normal database, while Figure 2c gives the respective results for the POAG database.

Threshold-related values, current HFA technique

Clinical data

Figures 2b and 2d give the distribution of errors (gold standard minus the second most sensitive thresholded location) for the normal and POAG databases.

Computer simulation

Figure 3 gives the distribution of errors (true threshold minus the second most sensitive thresholded location) for conditions in which zero, one, two and three of the four thresholded locations are defective (18dB defect).

Discussion

The regression line fitted to the variables age and sensitivity at the seventh most sensitive location has a gradient of –0.059dB/year. This value is similar to that reported by both Haas *et al.*[7] (–0.058dB/year) and Zulauf *et al.*[17] (–0.064dB/year), but below that reported by both Jaffe *et al.*[9] (–0.074 and –0.088dB/year) and Johnson and Choy[18] (–0.08dB/year).

The distributions of errors for the normal database are similar for the age-related (Fig. 2a) and HFA threshold-related (Fig. 2b) techniques. The distributions of errors for the age-related technique show a negative shift of approximately 2dB in the POAG database (Fig. 2c). This shift is due to a reduction in sensitivity at the seventh most

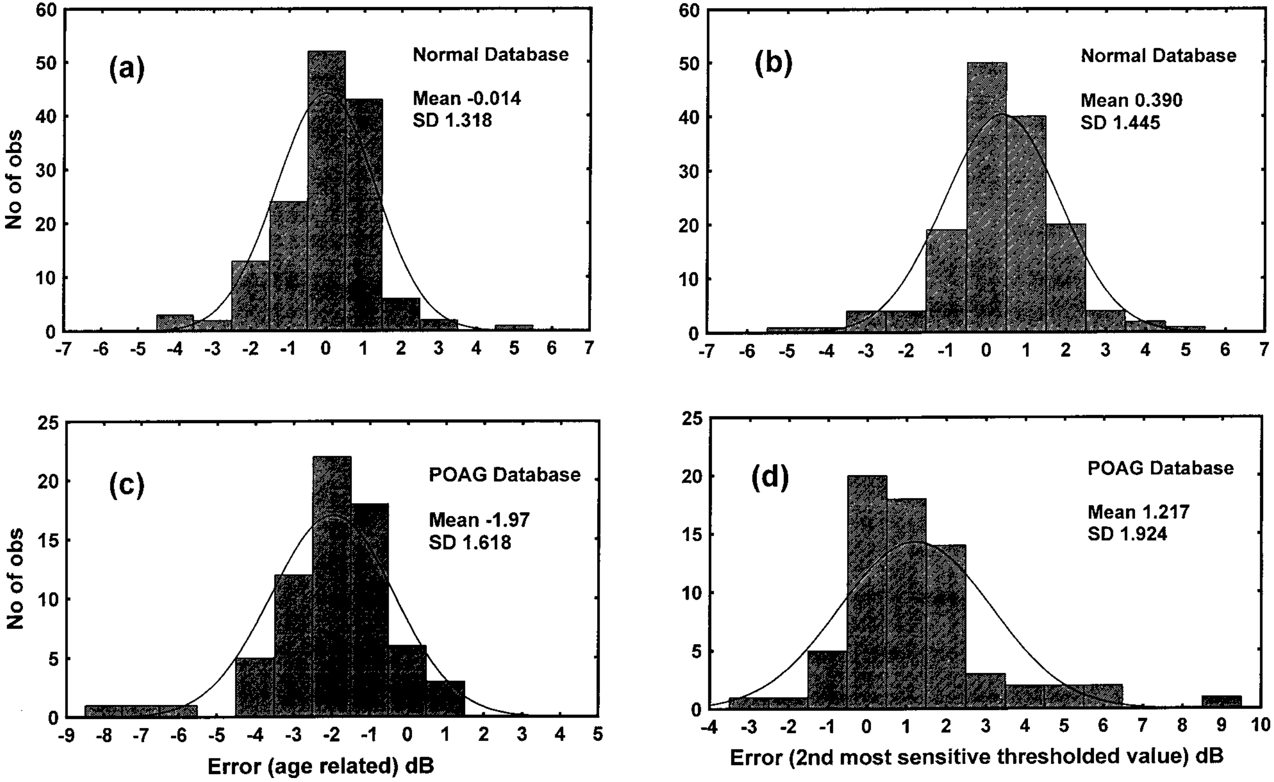

Fig. 2. Errors in the estimation of the appropriate test intensities for *a.* an age-related technique and normal population; *b.* an HFA threshold-related technique and normal population; *c.* an age-related technique and POAG population; and *d.* an HFA threshold-related technique and POAG population. Also included on each frequency histogram is the best-fitting gaussian distribution and the distributions mean and SD.

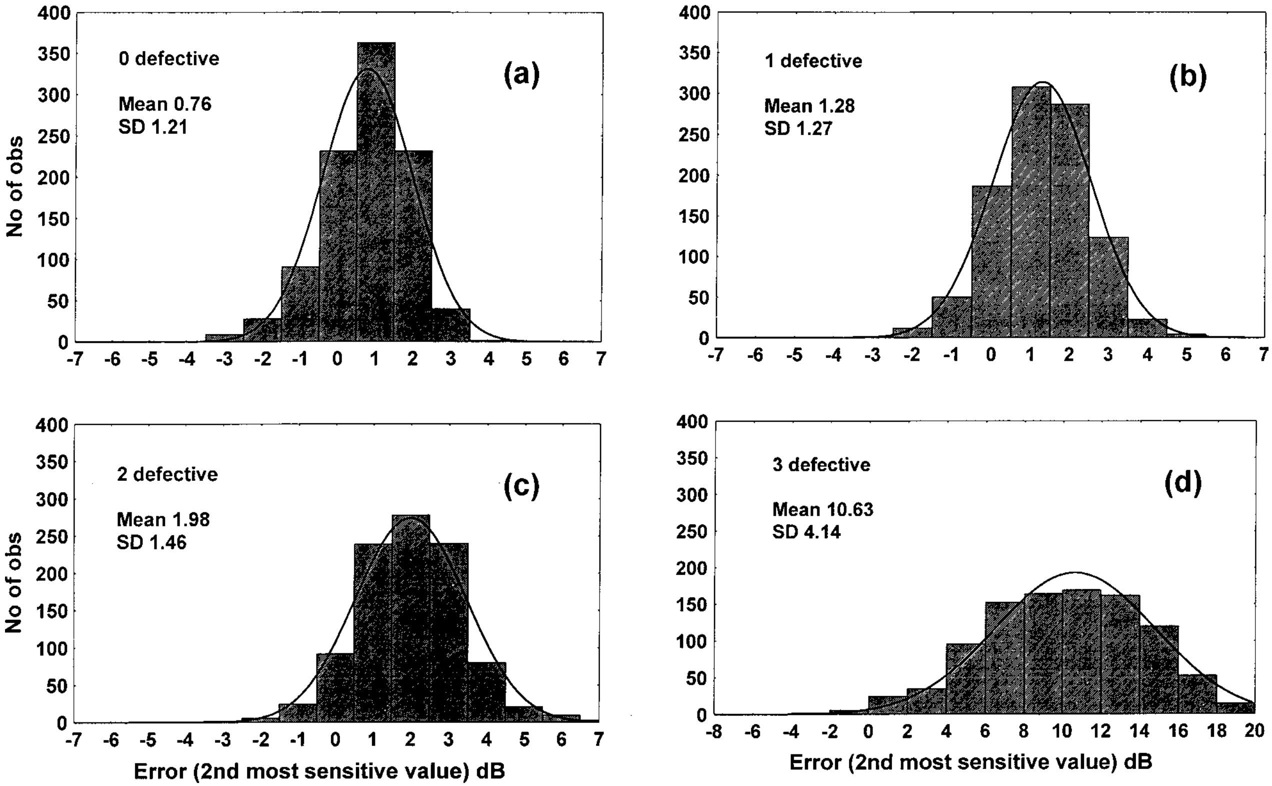

Fig. 3. Errors in the estimation of the appropriate test intensities, established by simulation of the HFA threshold-related technique, for conditions in which *a.* none of the four thresholded locations are defective; *b.* one of the four thresholded locations has a defect of 18dB; *c.* two of the four thresholded locations have a defect of 18dB; and *d.* three of the four thresholded locations have a defect of 18dB. Also included on each frequency histogram is the best-fitting gaussian distribution and the distributions mean and SD.

sensitive location in the POAG eyes, indicative of a diffuse depression in sensitivity[19-21] or a large area of loss.

The distributions of errors in the POAG database are also broader (Figs. 2c and 2d) than those from the normal database with the age-related distribution (Fig. 2c) having a negative skew (skewness -1.16) and the threshold-related distribution (Fig. 2d) a positive skew (skewness 1.42). The negative skew on the age-related distribution results from some of the POAG patients having a reduced gold standard measure of sensitivity. If an age setting had been used to screen the visual field of these patients, then it is highly likely that they would have missed a large number of the stimuli and failed the test (true positives). The positive skew on the threshold-related distribution results from the second most sensitive of the four seed locations having a lower sensitivity than the gold standard. In these situations, the suprathreshold test intensity would be set too high and the sensitivity of the suprathreshold test reduced (false-negatives). The HFA partially guards against this problem by putting an upper limit on the test intensity.

In summary, the results given in Figure 2 show no improvement in precision of the HFA threshold-related technique over that of an age-related one. If anything, the threshold-related measure is slightly less accurate, with a tendency to underestimate the sensitivity which would result in a higher percentage of false-negative test results.

The simulation of the HFA routine gave results that, for the condition of zero defective locations (Fig. 3a), were similar to the normal data (Fig. 2b). When the number of defective locations increased to one, two and three of the four locations, the variance of the errors increased and the mean error became more positive.

The POAG database contained patients with a large range of defects. The majority had early to intermediate loss that was likely to affect few of the thresholded locations. There were, however, some patients whose visual field loss was more severe. Looking at the data in Figure 3, we can see how the distribution of errors in the POAG database (Fig. 2d) can be explained by combining part of the distribution from 3c and 3d with that from 3a and 3b. The finding that the simulation reproduces the results from the normal population, and can also explain the distribution found in the POAG population, validates the use of this simulation to explore alternate strategies for establishing the appropriate test intensities in threshold-related suprathreshold perimetry.

For simulations to accurately predict clinical data, it is important that they include false-positive and false-negative response errors. While there have been a number of studies evaluating the error rates in both normal and POAG populations[22,23], there is no published report on the relationship between false-positives and false-negatives. The simulation used in this study overcame this problem by setting the false-positive and false-negative error rate of each simulation to equal that of a randomly selected patient from the normal database. This approach takes into account *1.* the variability in false-positive and false-negative errors found within clinical populations, and *2.* the existence of any relationship between false-positive and false-negative errors; and *3.* gives results mirroring those found within clinical populations.

The simplicity of an age setting with its widely understood limitations has much to commend it: *1.* it is no less precise than the current HFA threshold-related technique; *2.* it is independent of the extent of loss; and *3.* it is faster. However, its overall precision is not very good, with 18% of estimates being in error by two or more dB (Fig. 2a). Henson and Anderson[24] reported on the accuracy of a range of different techniques for setting the suprathreshold test intensity in semi-automated multiple stimulus suprathreshold perimetry. They found that a single staircase technique, in which the patients reported whether or not they could see one or more of the stimuli in patterns of three or four, had an SD of 0.86dB. This is much smaller than that reported for the age-related and the HFA threshold-related techniques, and highlights the potential advantages of a threshold-related technique not currently being achieved with the HFA algorithm.

A disadvantage of the age-related technique is that, in a clinical situation, the values would be based upon results collected at a remote site where the perimetric protocol and test environment may be different, as may be the age and quality of the instrument (see Gutteridge[11] for an example of differences between instruments). The combined effect of these factors could lead to a bias in the age setting which would further reduce either the sensitivity or specificity of the suprathreshold test.

The simulation used in this study gave results close to those found in the retrospective analysis of visual field data. This validates the use of simulation to look at alternate techniques that cannot be evaluated via an analysis of existing visual field data.

Conclusions

Threshold-related suprathreshold perimetry is generally accepted as being better than age-related suprathreshold perimetry as it takes into account overall shifts in sensitivity when selecting the suprathreshold test intensity. The current HFA threshold-related technique has, however, been found to be no more precise than an age-related technique which is likely to be in error by × ≤2dB in 20% of cases. These techniques are considerably less precise than a semi-automated one using multiple stimulus presentations. There is a need to investigate alternate single stimulus techniques for use in automated suprathreshold perimetry, which can increase the precision of establishing appropriate test intensities and the subsequent sensitivity and specificity of suprathreshold perimetry.

Acknowledgments

Supported by the International Glaucoma Association, Guide Dogs for the Blind and the Manchester Royal Eye Hospital endowment funds (DBH, PHA, and SJC) and Medical Research Council of Canada MT-11357 (BCC).

References

1. Klein BEK, Klein R, Sponsel WE, Franke T, Cantor LB, Martone J, Menage MJ: Prevalence of glaucoma: the Beaver-Dam eye study. Ophthalmology, 99:499-1504, 1992
2. Katz J, Tielsch JM, Quigley HA, Javitt J, Witt K, Sommer A: Automated suprathreshold screening for glaucoma: the Baltimore Eye Survey. Invest Ophthalmol Vis Sci 34:3271-3277, 1993
3. Sponsel WE, Ritch R, Stamper R, Higginbotham E, Anderson DR, Wilson MR, Zimmerman TJ: Prevent Blindness America visual field screening study. Am J Ophthalmol 120:699-708, 1995
4. Harrington DO, Flocks M: The multiple-pattern method of visual field examination: a five year evaluation of its effectiveness as a visual screening technique. Arch Ophthalmol 61:755-765, 1959
5. Friedmann AI: Serial analysis of change in visual field defects employing a new instrument to determine the activity of diseases involving the visual pathways. Ophthalmologica 152:1-12, 1966

6. Brenton RS, Phelps CD: The normal visual field on the Humphrey Visual Field Analyzer. Ophthalmologica 193:56-74, 1986
7. Haas A, Flammer J, Schneider U: Influence of age on the visual fields of normal subjects. Am J Ophthalmol 101:199-203, 1986
8. Heijl A, Lindgren E, Olsson J: Normal variability of static perimetric threshold values across the central visual field. Arch Ophthalmol 105:1544-1549, 1987
9. Jaffe GJ, Alvarado JA, Juster RP: Age-related changes of the normal visual field. Arch Ophthalmol 104:1021-1025, 1986
10. Greve EL: The visual field analyzer and threshold. Br J Ophthalmol 55:704-708, 1971
11. Gutteridge IF: The working threshold approach to Friedmann Visual Field Analyser screening. Ophthalmol Physiol Opt 3:41-46, 1982
12. Johnson CA, Chauhan BC, Shapiro LR: Properties of staircase procedures for estimating thresholds in automated perimetry. Invest Ophthalmol Vis Sci 33:2966-2974, 1992
13. Henson DB, Chaudry SJ, Artes PH: The relationship between sensitivity and variability in normal and glaucomatous visual fields. *This Volume*, pp. 95-101
14. Weber J, Rau S: The properties of perimetric thresholds in normal and glaucomatous eyes. German J Ophthalmol 1:79-85, 1992
15. Chauhan BC, Tompkins JD, LeBlanc RP, McCormick TA: Characteristics of frequency-of-seeing curves in normal subjects, patients suspected with glaucoma, and patients with glaucoma. Invest Ophthalmol Vis Sci 34:3534-3540, 1993
16. Bebie H, Fankhauser F, Spahr J: Static perimetry: strategies. Acta Ophthalmol (Kbh) 54:325-338, 1976
17. Zulauf M, Leblanc RP, Flammer J: The normal visual-fields measured with Octopus-program G1.2: global visual-field indices. Graefe's Arch Clin Exp Ophthalmol 232:516-522, 1994
18. Johnson CA, Choy D: On the definition of age-related norms for visual function testing. Appl Optics 8:1449-1454, 1987
19. Drance SM, Douglas GR, Airaksinen PJ, Schulzer M, Hitchings RA: Diffuse visual field loss in chronic open-angle glaucoma and low-tension glaucoma. Am J Ophthalmol 104:577-580, 1987
20. Drance SM: Diffuse visual field loss in open-angle glaucoma. Ophthalmology 98:1533-1538, 1991
21. Chauhan BC, LeBlanc RP, Shaw AM, Chan AB, McCormick TA: Repeatable diffuse visual field loss in open-angle glaucoma. Ophthalmology 104:532-538, 1997
22. Katz J, Sommer A: Reliability indexes of automated perimetric tests. Arch Ophthalmol 106:1252-1254, 1988
23. Corallo G, Gandolfo E, Morescalchi F, Semino E, Magnasco A, Novaro C, Micco MD: Reliability indices in automated perimetry. In: Mills RP (ed) Perimetric Update 1992/1993, pp 509-513. Amsterdam: Kugler Publ 1994
24. Henson DB, Anderson R: Threshold-related suprathreshold field testing: which is the best technique of establishing the threshold? In: Mills RP, Heijl A (eds) Perimetry Update 1990/1991, pp 367-372. Amsterdam/Milano: Kugler & Ghedini Publ 1991