CLINICAL SCIENCE

Interobserver agreement on visual field progression in glaucoma: a comparison of methods

A C Viswanathan, D P Crabb, A I McNaught, M C Westcott, D Kamal, D F Garway-Heath, F W Fitzke, R A Hitchings

....

See end of article for
authors' affiliations
....

Correspondence to:
Fred W Fitzke, PhD,
Institute of Ophthalmology,
11-43 Bath Street, London
EC1V 9EL, UK;
f.fitzke@ucl.ac.uk

Accepted for publication
11 September 2002
....

Aim: To examine the level of agreement between clinicians in assessing progressive deterioration in visual field series using two different methods of analysis.
Methods: Each visual field series satisfied the following criteria: more than 19 reliable fields, patient age over 40 years, macular threshold at least 30 dB. The first three fields in each series were excluded to minimise learning effects: the following 16 were studied. Five expert clinicians assessed the progression status of each series using both standard Humphrey printouts and pointwise linear regression (PROGRESSOR). The level of agreement between the clinicians was evaluated using a weighted kappa statistic.
Results: A total of 432 tests comprising 27 visual field series of 16 tests each were assessed by the clinicians. The level of agreement on progression status between the clinicians was always higher when they used PROGRESSOR (median kappa = 0.59) than when they used Humphrey printouts (median kappa = 0.32). This was statistically significant (p = 0.006, Wilcoxon matched pairs signed rank sum test).
Conclusions: Agreement between expert clinicians about visual field progression status is poor when standard Humphrey printouts are used, even when the field series studied are long and consist solely of reliable fields. Under these ideal conditions, clinicians agree more closely about patients' visual field progression status when using PROGRESSOR than when inspecting series of Humphrey printouts.

"At present, there is no generally accepted technique for detecting change in the visual field over time using automated perimetry . . . most clinicians who use automated perimetry are probably employing simple visual inspection of the perimetric data to diagnose progressive visual field loss in patients with glaucoma."[1]

Although these statements were made over a decade ago they are probably still true today, even though the knowledge of whether a patient's glaucoma is progressive or stable remains central to the management of the condition. Several scoring systems have been devised in order to identify visual field progression for the purposes of research[2–5] but none has found widespread acceptance in general clinical practice. However, unaided clinical judgment is inconsistent: even expert observers show considerable disagreement about whether a given visual field series signifies progression or stability.[1] One possible reason for this is that the standard output of most automated perimeters provides inadequate information relating to progression or stability. Thus, when the clinician is attempting to decide whether or not a given series of outputs constitutes progressive disease, the task involves manually comparing the decibel sensitivity values (or processed versions thereof) or graphical plots for all fields in the series. This task is further complicated by the contributions of "within test" variability (short term fluctuation)[6] and "between test" variability (long term fluctuation),[7] both of which are known to be increased above normal in glaucoma.[8] For these reasons, and because the grids of numbers produced by automated perimeters are easily amenable to numerical analysis, a great variety of software and statistical approaches have been taken to aid in the determination of visual field progression in glaucoma. One set of methods rely on estimates of change in summary measures of the field such as regression analysis of the mean defect value,[9] mean deviation,[10] other global measures,[10] measurement of whole field and quadrantic

sensitivity losses,[11] and trend and regression analysis of various estimates of the sensitivity of the whole field or parts of it.[12–14] However, the analysis of summary measures, whether based on the whole field or on clusters of points within it, has been found to be "remarkably poor"[15] and "of little value"[16] in detecting glaucomatous change. Summary measures largely or completely ignore the detailed spatial information contained within computerised field tests and are insensitive to early localised change.[17] Furthermore, different regions of the visual field may deteriorate at different rates.[14 18 19]

PROGRESSOR, which has been described fully previously in the *BJO*,[20] avoids these problems by performing a point by point linear regression analysis of sensitivity on time for the whole visual field series. This technique has been used for several years to investigate glaucomatous visual field change[21 22] and has recently been re-examined.[10] The pointwise linear model has been demonstrated to provide a valid framework for detecting and forecasting glaucomatous loss.[23] Using equivalent progression criteria, PROGRESSOR has been found to compare favourably with other pointwise analyses such as the STATPAC 2 glaucoma change probability analysis[24] for the Humphrey field analyser (Humphrey Instruments Inc, Dublin, CA, USA).[25] Pointwise linear regression has been found to agree more closely with expert clinical judgment about progression status than glaucoma change probability analysis.[26]

Since there is no general consensus on what pointwise value of regression slope and p value constitutes progression, or whether there should be a requirement for contiguous points to show this behaviour and whether it should be maintained in subsequent fields, at present PROGRESSOR remains a subjective analysis. For this reason, and because there is no universally accepted gold standard for visual field progression, this study investigates the usefulness of PROGRESSOR by determining the level of agreement between expert observers using both PROGRESSOR and standard clinical techniques (manually comparing serial printouts from an automated perimeter).

Table 1 Classification of field series using Humphrey printouts

	Clinician B			
Clinician A	Definitely stable	Probably stable	Probably progressing	Definitely progressing
Definitely stable	**5**	4	1	0
Probably stable	4	**1**	2	1
Probably progressing	0	3	**1**	3
Definitely progressing	0	1	0	**1**

*Weighted kappa = 0.28 (SE 0.08).
Classification of the 27 visual field series by clinician A and clinician B when using Humphrey printouts. These two clinicians agreed exactly on the assessment of eight series (along the diagonal of the table). Note that one field series was assessed as being "probably stable" by clinician A but "definitely progressing" by clinician B. This level of disparity in assessment occurred with three series in total. These "serious disagreements," which are accounted for in the calculation of the weighted kappa value, were a feature of the results from other pairs of clinicians when Humphrey printouts were the method of assessment.

Table 2 Classification of field series using PROGRESSOR

	Clinician B			
Clinician A	Definitely stable	Probably stable	Probably progressing	Definitely progressing
Definitely stable	**7**	0	0	0
Probably stable	0	**0**	1	0
Probably progressing	0	4	**2**	4
Definitely progressing	0	1	2	**6**

*Weighted kappa = 0.63 (SE 0.13).
Classification of the 27 visual field series by clinician A and clinician B when using PROGRESSOR. These clinicians now agreed exactly on the assessment of 15 series (along the diagonal of the table). "Serious disagreements" have been reduced and the table suggests closer conformity between the clinicians when they assessed the series with PROGRESSOR

METHODS
Selection of visual field data
The visual field series presented to the five clinicians in the study for evaluation were drawn from the clinical visual field database of the glaucoma service at Moorfields Eye Hospital, a tertiary referral centre also serving the local community. At the time of the study this database contained Humphrey visual field tests performed between 16 January 1985 and 6 August 1997. All data were the result of standard clinical testing on Humphrey model 630 perimeters. The following filters were applied to the database in order to obtain a group of visual field series such as those commonly encountered in clinical practice when the degree of visual field deterioration must be estimated, but without specifying the presence, absence, or nature of progression in the visual field series to be studied based on any a priori assumptions of what constitutes progression:

- Patient age greater than 40 years

- Each field series consisted solely of tests using the Humphrey 24-2 and 30-2 test grid patterns with standard 4–2 dB staircase thresholding strategy

- A white stimulus of Goldmann size III was used throughout

- Each field test was required to meet reliability criteria as set by the Humphrey perimeter: fewer than 20% fixation losses, fewer than 33% false positive and fewer than 33% false negative responses

- The macular threshold of each field test was required to be at least 30 decibels. This criterion was introduced so as to rule out visual field tests with defects arising from significant media opacities or macular disease

- Each field series included in the study contained at least 19 fields satisfying the foregoing criteria. The first three fields in each series were ignored to obviate learning effects[27]: the following 16 fields were presented to the clinicians for analysis

- If both eyes of a particular patient fulfilled the inclusion criteria, one eye was selected at random.

Analysis of visual field data
Each of the visual field series was assessed by all five clinicians (AMcN; MW, DK, DG-H, RH). All five clinicians were glaucoma specialists. They were all experienced in the interpretation of series of standard Humphrey visual field printouts (raw sensitivity values, the grey scale plot, total deviation values and plot, pattern deviation values and plot, global indices, and glaucoma hemifield test) in order to determine progression status, both from clinical practice and from research into visual field deterioration. At the time of the study two of the clinicians (AMcN and RH) were familiar with PROGRESSOR: the others had a working knowledge from the published literature. Initially, the clinicians were asked to examine the visual field series presented as standard Humphrey printouts. Although each patient's visual fields were, of course, presented in chronological order within a field series, the field series themselves were presented in a random order determined by a software random number generator. Clinicians were asked to use their judgment to assess each field series and to assign it to one of four categories: definitely stable, probably stable, probably progressing, or definitely progressing. Then, the clinicians were asked to use their judgment to rate the field series using PROGRESSOR. For this analysis the field series were presented in a different random order from that used for the Humphrey printouts. (Naturally, each patient's visual fields remained in chronological order within a given field series.)

The clinicians were allowed to use any of the options within PROGRESSOR which they considered were necessary for an accurate analysis. These included the cumulative graphical output, a choice of progression criteria for slope and p value, grey scale plots, animation analysis, and Gaussian filtering. The latter has been shown to reduce long term fluctuation[28] without delay in the detection of visual field deterioration.[29] The clinicians were asked to categorise the field series using PROGRESSOR into the same four categories previously described for use with the Humphrey printouts. All the clinicians evaluating the visual field series were aware of the purpose of the study.
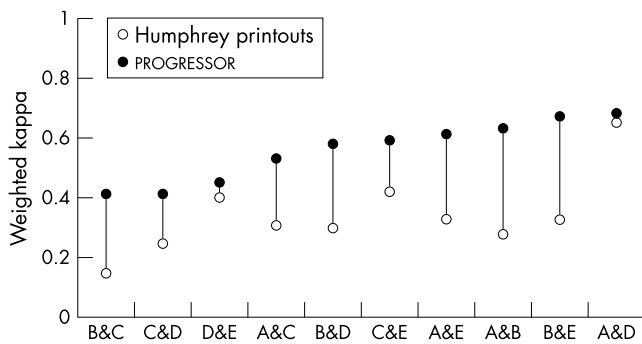
Figure 1 Weighted kappa values of agreement for all 10 pairs of clinicians. Each pair is represented by an open symbol representing the weighted kappa value when using Humphrey printouts and a closed symbol when using PROGRESSOR. The pairs are ranked along the x-axis by magnitude of the weighted kappa value. In all cases the weighted kappa value of agreement was higher when clinicians used PROGRESSOR compared to when using Humphrey printouts. The range of standard error for all the weighted kappa values was 0.08 to 0.14.

In order to measure of the level of intraobserver agreement two of the clinicians were asked to re-examine all the field series using both Humphrey printouts and PROGRESSOR 3 months after their original analysis. These clinicians were given no warning that intraobserver variation was to be measured as part of the study. Visual field series were evaluated independently by each clinician. All field series were anonymised. For the intraobserver agreement study clinicians were masked to their previous evaluations.

Statistical analysis

The level of agreement among five individuals (A, B, C, D, E) may be expressed as the levels of agreement of the 10 possible pairs (A&B, A&C, A&D, A&E, B&C, B&D, B&E, C&D, C&E, D&E). The level of interobserver agreement was measured using the weighted kappa statistic[30] for all 10 possible pairs of clinicians. Weighted kappa is an appropriate chance adjusted measure of agreement between two observers when there are more than two ordered categories of classification. The statistic, which ranges from 0 (agreement no better than chance) to 1 (perfect agreement), gives partial credit for partial agreement in accordance with a linear weighting scheme. Weighted kappa is obtained by giving weights to the frequencies in each cell of the table according to their distance from the diagonal that indicates agreement. A simple linear weighting scheme was adopted. For the cell in row i and column j, with observed frequency f_{ij}, the weight was calculated as

$$w_{ij} = 1 - \frac{|i - j|}{3}$$

Additionally, the level of the observer concordance was assessed: the proportion of the observers agreeing about the

category of a particular field series was measured. Similar statistics were used to measure the level of intraobserver agreement.

All statistical analysis was performed using the software package S-PLUS 3.2 for Windows (StatSci Europe, MathSoft Inc, Oxford, UK).

RESULTS

Twenty seven visual field series were assessed by the clinicians. Since each series consisted of 16 visual field tests each clinician assessed a total of 432 visual field tests. The median age of the patients used in this study at the first visual field in the series analysed was 61 years (range 44–72 years). The median Humphrey mean deviation (MD) of the first visual field in each series analysed was −7.7 dB (range −0.1 to −14.8 dB). The median length of follow up for the visual field series was 5.7 years (range 3.3–7.7 years).

Tables of agreement were produced for all 10 possible pairs of the clinicians. For example, Table 1 shows the classification of the 27 field series by one pair of the clinicians (A and B) when using the Humphrey printouts. The weighted kappa value for the agreement exhibited is 0.28 (SE 0.08). Table 2 shows the results for the same pair of clinicians when using PROGRESSOR analysis. The weighted kappa value for the agreement between the two clinicians has increased substantially to 0.63 (SE 0.13). Approximate 95% confidence intervals may be calculated from $2 \times SE$ (kappa) according to Fleiss.[30]

The weighted kappa values for all the pairs of clinicians using the different methods of analysis are shown graphically in Figure 1. All pairs of clinicians demonstrated greater weighted kappa values for agreement when PROGRESSOR was used compared to Humphrey printouts. The median weighted kappa value for all pairs of clinicians using PROGRESSOR was 0.59 compared to a median of 0.32 when using Humphrey printouts. Qualitative interpretation of kappa levels of agreement are imprecise but guidelines have been published.[31] Generally a value less than 0.40 indicates only "slight" agreement and a value above 0.60 indicates "substantial" agreement. The approximate 95% confidence intervals for the weighted kappa values of PROGRESSOR and Humphrey printouts for individual pairs of clinicians overlap suggesting that statistical significance has not been achieved for these individual cases. However, the finding that all pairs of clinicians consistently demonstrated higher levels of agreement when using PROGRESSOR compared to Humphrey printouts is statistically significant at the 5% level (p = 0.006, Wilcoxon matched pairs signed rank sum test.)

Tables 3 and 4 show the concordance between the clinicians in classifying the visual field series using the different methods. One hundred per cent concordance was defined as all five clinicians either agreeing a series was progressing (either probably or definitely) or all five clinicians agreeing a series was stable (either probably or definitely). This rate of concordance was achieved in only 10 of the 27 series (37%) when the clinicians used Humphrey printouts. In comparison, the 100% rate of concordance was doubled when the clinicians used PROGRESSOR: all five clinicians had identical opinions on 20 of the 27 series (74%).

Tables 3 and 4 also show that 13 of the 27 series (48%) were classified as progressing by the majority (at least three or 60%) of the clinicians using Humphrey printouts. In contrast, 19 of the 27 series (70%) were classified as progressing by the majority of the clinicians when using PROGRESSOR.

Intraobserver agreement was assessed in two of the clinicians as described in the Methods section. The weighted kappa statistic for intraobserver agreement for clinician A was 0.43 (SE 0.11) using Humphrey printouts and 0.71 (SE 0.13) when using PROGRESSOR. The weighted kappa statistic for intraobserver agreement for clinician B was 0.60 (SE 0.12) using Humphrey printouts and 0.83 (SE 0.09) when using PROGRESSOR.

Table 3 Concordance between clinicians when using Humphrey printouts

Concordance among clinicians	Series classified as progressing	Series classified as stable	Total
100%	5	5	10 (37%)
80%	2	6	8 (30%)
60%	6	3	9 (33%)
Total	13	14	27 (100%)

Categories were collapsed into progressing (probably and definitely) and stable (probably and definitely). 100% concordance represents exact agreement among all five clinicians, 80% concordance represents exact agreement among four clinicians, and 60% concordance represents exact agreement among three clinicians.

Table 4 Concordance between clinicians when using PROGRESSOR

Concordance among clinicians	Series classified as progressing	Series classified as stable	Total
100%	13	7	20 (74%)
80%	3	0	3 (11%)
60%	3	1	4 (15%)
Total	19	8	27 (100%)

Categories were collapsed into progressing (probably and definitely) and stable (probably and definitely). 100% concordance represents exact agreement among all five clinicians, 80% concordance represents exact agreement among four clinicians, and 60% concordance represents exact agreement among three clinicians.

DISCUSSION

It is not particularly surprising that only slight interobserver agreement was found when standard Humphrey printouts were used as a basis for a decision about glaucomatous visual field progression (median weighted kappa value 0.32). In a previous study, Werner and colleagues measured the level of agreement between six experienced observers in rating the progression status of automated visual field series from 30 glaucoma patients.[1] Although the weighted kappa value was not reported in that paper, it has been calculated subsequently as 0.402.[32] This is comparable to our findings. It is interesting that Werner and colleagues found a slightly higher level of agreement than ours, even though in their study the mean number of visual fields in each series was 6.3 whereas in the present study it was fixed at 16. This suggests that the clinicians in the present study did not benefit from being asked to analyse relatively more visual field data per subject. In fact, longer periods of visual field follow up may actually make the task of deciding about progression status more difficult and complex, when this task is based on standard automated visual field printouts. Werner and colleagues found that all the observers agreed about progression status in 11 of their 30 subjects. The present study found that there was complete agreement about 10 of the 27 subjects. Although Werner and colleagues used six observers whereas the present study used five these results appear comparable. Once again, the level of agreement between observers in rating visual field progression based on standard automated perimeter output does not seem to have been influenced by the relatively longer follow up in the present study.

A consistently higher level of interobserver and intraobserver agreement was found when PROGRESSOR was used to rate progression status rather than standard perimeter output. This suggests that clinicians are better able to make meaningful, systematic decisions about visual field progression status when using PROGRESSOR rather than standard automated perimeter output. Chauhan and colleagues conducted a similar study to the present one in which five observers rated the progression status of 32 visual field series using a computer animated graphics technique which corrected for test-retest variability in order to aid the recognition of progression.[32] They found a weighted kappa value of 0.572 which is very similar to the figure of 0.59 obtained in the present study. Chauhan and colleagues found that at least four out of five observers (80% or greater concordance) agreed on progression status in 27 out of 32 visual field series (84%). This is remarkably similar to our findings of 80% or better concordance on 23 out of 27 field series (85%). The level of complete agreement between all five observers (100% concordance) is, however, greater in the present study: 74% compared to Chauhan and colleagues' figure of 56%. Both these studies suggest that the level of agreement about progression status rises when clinicians analyse displays of visual field series which highlight information pertaining to progression which is not obvious in the standard perimeter outputs.

The number of visual field tests in each series in our study was fixed at 16. This filter was applied to the database of visual field series in order both to optimise the chances of agreement between observers and to produce a similar number of visual field series to those used by Werner and colleagues (30 series) and Chauhan and colleagues (32 series) without resorting to random or other potentially biased methods of subselection. As a consequence of this procedure the length of visual field series used in the present study was longer than those used by Werner and colleagues (mean series length 6.3) or Chauhan and colleagues (median series length 7.0). As discussed previously, the difference in series length does not appear to have greatly influenced the level of agreement when observers are considering standard Humphrey printouts. Levels of agreement also seem similar between PROGRESSOR and the technique used by Chauhan and colleagues despite the difference in series length, but it is difficult to speculate what effect shorter series length might have on the level of agreement using PROGRESSOR. As a further consequence of the filtering procedure, field series with relatively frequent (approximately 4 monthly) tests were obtained. This frequency of testing is undertaken when progression is suspected[33] and for the routine follow up of normal tension glaucoma patients in our department: these groups of patients are likely to be relatively over-represented in the study sample.

The requirement for the macular threshold of each visual field test in a series to be at least 30 dB almost certainly affected the results of this study. Had this criterion not been imposed, the clinicians may have shown less agreement on visual field series progression status, or they may have shown higher levels of agreement over visual field deterioration which was not in fact truly glaucomatous in nature. Media opacities in particular are an important source of error in studies which specify visual field deterioration in glaucoma as an outcome: various strategies have been employed to mitigate against the effects of media opacities, including those reported in the Collaborative Normal Tension Glaucoma Study[34] and the Early Manifest Glaucoma Trial.[3] In the present study, however, visual field progression was not an outcome measure and the inclusion of visual field series with significant contributions from media and macular pathology would have presented our clinicians with an inappropriate task: in the clinical scenario, the task of judging progression in these cases is informed by clinical examination of the patient. The 30 dB macular threshold criterion may have resulted in visual field tests showing advanced glaucomatous damage being excluded from the study. However, many of the test locations in such cases fall outside the dynamic range of the perimeter and thus yield little useful information about progression status. Strictly, though, the results of this study may only be generalised to similar clinical scenarios (that is, learning effects accounted for, all tests reliable, long series with good macular sensitivity threshold).

The clinicians in the present study were given no explicit guidelines upon which to base the diagnosis of stability or progression for each visual field series. They were merely asked to use their clinical judgment based on the perimetric output. Thus, it is possible that a different group of observers might have used different "personal progression criteria" which would have led to different results. In the absence of a gold standard for the diagnosis of glaucomatous visual field progression, however, the analytical task presented to the clinicians in the present study resembles the conditions encountered in clinical practice more closely than if predetermined progression criteria had been imposed. Although the level of agreement between the observers was similar for PROGRESSOR and for the glaucoma change analysis described by Chauhan and colleagues, speculation about the personal progression criteria used by the observers in the two studies is difficult, since the analytical task of diagnosing progression or stability using PROGRESSOR is quite different from that using the technique described by Chauhan and colleagues.

The use of PROGRESSOR led to the clinicians classifying a higher proportion of the field series as progressing than when the standard printouts were used. In the absence of an external gold standard for visual field progression, this may indicate either an increased sensitivity in the detection of progression or alternatively a reduction in specificity. An approach using software simulation of visual field data[35] has suggested that pointwise linear regression has a sensitivity and specificity of over 90% to detect significant rates of deterioration of 2.5 dB per year or worse, when 10 fields per series are available for study. However, if the number of fields available is limited to five, the sensitivity is reduced to around 25% (though the specificity is maintained owing to the requirement for a statistically significant deterioration). This lack of sensitivity when few fields are available for analysis is common to all forms of progression assessment. There is some evidence that the reliability of PROGRESSOR in detecting visual field deterioration in glaucoma compares favourably with other automated analyses such as STATPAC 2.[25] Furthermore, more visual field series were classified as stable with 100% concordance using PROGRESSOR than using standard Humphrey printouts: in only one of the eight field series classified as stable using PROGRESSOR was there less than 100% concordance. It is also possible that the higher levels of agreement found with PROGRESSOR are partly explained by the higher proportion of series classified as progressing using PROGRESSOR, since this in itself will tend to generate more agreement. It is difficult to estimate the size of this effect since agreement and decisions about progression status cannot be studied separately.

It is likely that new developments and refinements in the area of computer assisted diagnosis of visual field progression in glaucoma will yield higher levels of agreement between clinicians in the future. However, the interpretation of any algorithm and its application to patient management within the clinical context will remain a subjective matter of clinical expertise.

ACKNOWLEDGEMENTS

. . . .

Authors' affiliations

A C Viswanathan, F W Fitzke, Institute of Ophthalmology, London, UK
D P Crabb, Nottingham Trent University, Nottingham, UK
A I McNaught, Gloucestershire Eye Unit, Cheltenham, UK
M C Westcott, D Kamal, D F Garway-Heath, R A Hitchings, Moorfields Eye Hospital, London, UK

REFERENCES

1 **Werner EB**, Bishop KI, Koelle J, et al. A comparison of experienced clinical observers and statistical tests in detection of progressive visual field loss in glaucoma using automated perimetry. *Arch Ophthalmol* 1988;**106**:619–23.
2 **Musch DC**, Lichter PR, Guire KE, et al. The Collaborative Initial Glaucoma Treatment Study: study design, methods, and baseline characteristics of enrolled patients. *Ophthalmology* 1999;**106**:653–62.
3 **Leske MC**, Heijl A, Hyman L, et al. Early Manifest Glaucoma Trial: design and baseline data. *Ophthalmology* 1999;**106**:2144–53.
4 **AGIS**. The advanced glaucoma intervention study (AGIS): 7. The relationship between control of intraocular pressure and visual field deterioration.The AGIS Investigators. *Am J Ophthalmol* 2000;**130**:429–40.
5 **Collaborative Normal-Tension Glaucoma Study Group**. Comparison of glaucomatous progression between untreated patients with normal-tension glaucoma and patients with therapeutically reduced intraocular pressures [see comments]. *Am J Ophthalmol* 1998;**126**:487–97.
6 **Flammer J**. The concept of visual field indices. *Graefes Arch Clin Exp Ophthalmol* 1986;**224**:389–92.
7 **Boeglin RJ**, Caprioli J, Zulauf M. Long-term fluctuation of the visual field in glaucoma. *Am J Ophthalmol* 1992;**113**:396–400.
8 **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**:704–6.
9 **Weber J**, Koll W, Krieglstein GK. Intraocular pressure and visual field decay in chronic glaucoma. *Ger J Ophthalmol* 1993;**2**:165–9.
10 **Smith SD**, Katz J, Quigley HA. Analysis of progressive change in automated visual fields in glaucoma. *Invest Ophthalmol Vis Sci* 1996;**37**:1419–28.
11 **Wegner A**, Ugi I, Hofman A. A long-term visual field evaluation of glaucoma patients treated topically with timolol or carteolol. In: Mills-RP, ed. *Perimetry update 1992/1993*. Amsterdam: Kugler and Ghedini, 1993:143–5.
12 **Holmin C**, Krakau CE. Regression analysis of the central visual field in chronic glaucoma cases. A follow-up study using automatic perimetry. *Acta Ophthalmol Copenh* 1982;**60**:267–74.
13 **Wu D**, Schwartz B, Nagin P. Trend analyses of automated visual fields. *Doc Ophthalmol Proc Ser* 1987:175–89.
14 **O'Brien C**, Schwartz B, Takamoto T, et al. Intraocular pressure and the rate of visual field loss in chronic open-angle glaucoma. *Am J Ophthalmol* 1991;**111**:491–500.
15 **Chauhan BC**, Drance SM, Douglas GR. The use of visual field indices in detecting changes in the visual field in glaucoma. *Invest Ophthalmol Vis Sci* 1990;**31**:512–20.
16 **Birch MK**, Wishart PK, O'Donnell N. Determining progressive field loss. In: Mills RP, Wall M, eds. *Perimetry update 1994/1995*. Amsterdam: Kugler and Ghedini, 1995:31–6.
17 **Wild JM**, Hussey MK, Flanagan JG, et al. Pointwise topographical and longitudinal modeling of the visual field in glaucoma. *Invest Ophthalmol Vis Sci* 1993;**34**:1907–16.
18 **Hoskins HD**, Jensvold N, Zaretsky M, et al. Rate of progression of discrete areas of the visual field. In: Heijl A, ed. *Perimetry update 1988/1989*. Amsterdam: Kugler and Ghedini, 1989:173–6.
19 **O'Brien C**, Schwartz B. The visual field in chronic open angle glaucoma: the rate of change in different regions of the field. *Eye* 1990;**4(Pt 4)**:557–62.
20 **Fitzke FW**, Hitchings RA, Poinoosawmy D, et al. Analysis of visual field progression in glaucoma. *Br J Ophthalmol* 1996;**80**:40–8.
21 **Noureddin BN**, Poinoosawmy D, Fitzke FW, et al. Regression analysis of visual field progression in low tension glaucoma. *Br J Ophthalmol* 1991;**75**:493–5.
22 **Poinoosawmy D**, Wu J, Fitzke FW, et al. Discrimination between progression and non-progression visual field loss in low tension glaucoma using MDT. In: Mills-RP, ed. *Perimetry update 1992/1993*. Amsterdam: Kugler and Ghedini, 1993:109–14.
23 **McNaught AI**, Crabb DP, Fitzke FW, et al. Modelling series of visual fields to detect progression in normal tension glaucoma. *Graefes Arch Clin Exp Ophthalmol* 1995;**233**:750–755.
24 **Heijl A**, Lindgren G, Lindgren A, et al. Extended empirical statistical package for evaluation of single and multiple fields in glaucoma: Statpac 2. In: Mills RP, Heijl A, eds. *Perimetry update 1990/1991*. Amsterdam: Kugler and Ghedini, 1991:303–15.
25 **Viswanathan AC**, Fitzke FW, Hitchings RA. Early detection of visual field progression in glaucoma: a comparison of PROGRESSOR and STATPAC 2. *Br J Ophthalmol* 1997;**81**:1037–42.
26 **Nouri-Mahdavi K**, Brigatti L, Weitzman M, et al. Comparison of methods to detect visual field progression in glaucoma [published erratum appears in *Ophthalmology* 1998;**105**:7]. *Ophthalmology* 1997;**104**:1228–36.
27 **Werner EB**, Adelson A, Krupin T. Effect of patient experience on the results of automated perimetry in clinically stable glaucoma patients. *Ophthalmology* 1988;**95**:764–7.
28 **Fitzke FW**, Crabb DP, McNaught AI, et al. Image processing of computerised visual field data.: *Br J Ophthalmol* 1995;**79**:207–12.
29 **Viswanathan AC**, Hitchings RA, Fitzke FW. Spatial filtering of glaucomatous visual fields using PROGRESSOR for Windows. In: Wall M, Heijl A, eds. *Perimetry update*. Amsterdam/New York: Kugler Publications, 1996-7:311–19.
30 **Fleiss JL**. *Statistical methods for rates and proportions*. New York: John Wiley, 1981:212–25.
31 **Landis JR**, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159–74.
32 **Chauhan BC**, Drance SM, LeBlanc RP, et al. Technique for determining glaucomatous visual field progression by using animation graphics. *Am J Ophthalmol* 1994;**118**:485–91.
33 **Viswanathan AC**, Hitchings RA, Fitzke FW. How often do patients need visual field tests? *Graefes Arch Clin Exp Ophthalmol* 1997;**235**:563–8.
34 **Collaborative Normal-Tension Glaucoma Study Group**. The effectiveness of intraocular pressure reduction in the treatment of normal-tension glaucoma [see comments]. *Am J Ophthalmol* 1998;**126**:498–505.
35 **Spry PGD**, Johnson CM, Bates AB, et al. Discriminatory power of pointwise linear regression for detection of glaucomatous visual field defect progression. In: Wall M, Mills RP, eds. *Perimetry update*. The Hague, Netherlands: Kugler, 2000-1:363–9.