Agreement to detect glaucomatous visual field progression by using three different methods: a multicentre study

M Iester,[1] E Capris,[1] F De Feo,[2] M Polvicino,[2] P Brusini,[3] P Capris,[4] G Corallo,[1] M Figus,[5] P Fogagnolo,[6] P Frezzotti,[7] G Manni,[8] A Perdicchi[9]

[1]Laboratorio clinico anatomo-funzionale per la diagnosi e il trattamento del glaucoma e della malattie neurooftalmologiche, Clinica Oculistica, Department of Neurological Sciences, Ophthalmology, Genetic, University of Genoa, Italy
[2]Divisione di Oculistica, Ospedale San Paolo, Savona, Italy
[3]Divisione di Oculistica, S. Maria della Misericordia Hospital, Udine, Italy
[4]Ophthalmic Division of the G. Gaslini Institute, Genoa, Italy
[5]Clinica Oculistica, University of Pisa, Italy
[6]G. B. Bietti Eye Foundation for the Study and Research in Ophthalmology - IRCCS, Rome, Italy
[7]Dipartimento di chirurgia, sezione Oftalmologia, University of Siena, Italy
[8]Clinica Oculistica, University of Tor Vergata, Rome, Italy
[9]Clinica Oculistica, University Sapienza 2, Rome, Italy

Correspondence to
Michele Iester, MD, PhD, University Eye Clinic, Viale Benedetto XV, 5, 16132 Genoa, Italy; iester@unige.it

This study has been presented in part at the Association for Research and Vision in Ophthalmology (ARVO) meeting 2–6 May 2010, in Ft Lauderdale, Florida, USA.

Accepted 28 October 2010
Published Online First 3 December 2010

ABSTRACT

Aim To examine the level of agreement among nine clinicians in assessing progressive deterioration in visual field (VF) overview using three different methods of analysis.

Methods Each visual field was assessed by Humphrey Field Analyzer (HFA), program 24-2 SITA Standard. Nine expert clinicians assessed the progression status of each series by using HFA 'overview printouts' (HFA OP), the Guided Progression Analysis (GPA) and the Guided Progression Analysis (GPA2). VF series were presented in random order, but each patient's VF remained in chronological order within a given field series. Each clinician adopted his personal methods based on his knowledge to evaluate VF progression. The level of agreement between the clinicians was evaluated by using weighted κ statistics.

Results A total of 303 tests, comprising 38 visual field series of 7.9±3.4 tests (mean±SD), were assessed by the nine glaucoma specialists. When the intra-observer agreement was evaluated between HFA OP and GPA, the mean κ statistic was 0.58±0.13, between HFA OP and GPA2, κ was 0.55±0.06 and between GPA and GPA2 it was 0.56±0.17. When the inter-observer agreement was analysed κ statistic was 0.65 for HFA OP, 0.54 for GPA and 0.70 for GPA2.

Conclusions Using any procedure for evaluating the progression of a series of VF, agreement between expert clinicians is moderate. Clinicians had higher agreement when GPA2 was used, followed by HFA OP and GPA printouts, but these differences were not significant.

INTRODUCTION

Glaucoma is an asymmetric, progressive disease whose treatment can slow down the changes but usually cannot stop it. Intraocular pressure (IOP) is the most important risk factor to treat.[1] Clinicians must detect glaucomatous changes by observing the optic nerve head (ONH) and the visual fields. It is not easy to identify these visual field (VF) changes because of the short and long term fluctuation of the sensitivity. Many different algorithms have been introduced to distinguish between fluctuation and progression, but none has shown to be the best, even if some of these have proven to be useful for progression detection in some clinical trials.[2–10] The simple examination of all the graphical plots and the behaviour of VF indices is fundamental for the evaluation of VF progression or stability, which requires clinical experience and time consumption.[11–13]

Linear regression of the VF indices or of the sensitivity of the tested points has been used in many statistical programs such as Glaucoma Change Probability, Change Analysis or Progressor. Recently the 'Guided Progression Analysis' (GPA2), which is a statistical program to evaluate VF progression, has been introduced in clinical practice. The purpose of this study was to examine the level of agreement among nine clinicians in assessing VF progression by using three different methods of analysis: the standard Humphrey 'overview printout' (HFA OP), the 'Guided Progression Analysis' (GPA) and the GPA2 to better understand which method could be more useful in clinical practice.

PATIENTS AND METHODS

This was a retrospective chart review study of at least 5 years of follow-up (F/U). It included primary open angle glaucoma patients from the population attending the Glaucoma Clinic of the University of Genoa. The study, approved by the institutional ethical committee of the department, was in agreement with the tenets of the Declaration of Helsinki.

Patients were not excluded on the basis of gender, age or race. Visual fields were assessed by Humphrey Field Analyzer 750 II, (HFA, Carl Zeiss Meditec, Dublin, California, USA), using the 24-2 SITA standard (Swedish Interactive Thresholding Algorithm) test.

Patients were classified as having primary open angle glaucoma when they had a typical abnormal ONH and/or a typical glaucomatous VF, open angle at gonioscopy, IOP>21 mm Hg before treatment and no clinically apparent secondary cause for their glaucoma.[14] All the included patients were expert in performing VFs and they had already done 3 VFs, which were not considered in this study to avoid any learning effects and were not included in the printouts.

The abnormal ONH classification[15] was based on the presence of an optic rim notch or of diffuse/generalised loss of optic rim tissue, vertical cup/disc diameter ratio asymmetry unexplained by side differences in optic disc size, or disc haemorrhage. A glaucomatous VF defect[15] was defined as: three adjacent points depressed by 5 dB, with one of the points depressed by at least 10 dB; two adjacent points depressed by 10 dB, or a 10 dB difference across the nasal horizontal meridian in two adjacent points. None of the points could be edge points unless immediately above or below the nasal horizontal meridian. In addition, visual field testing

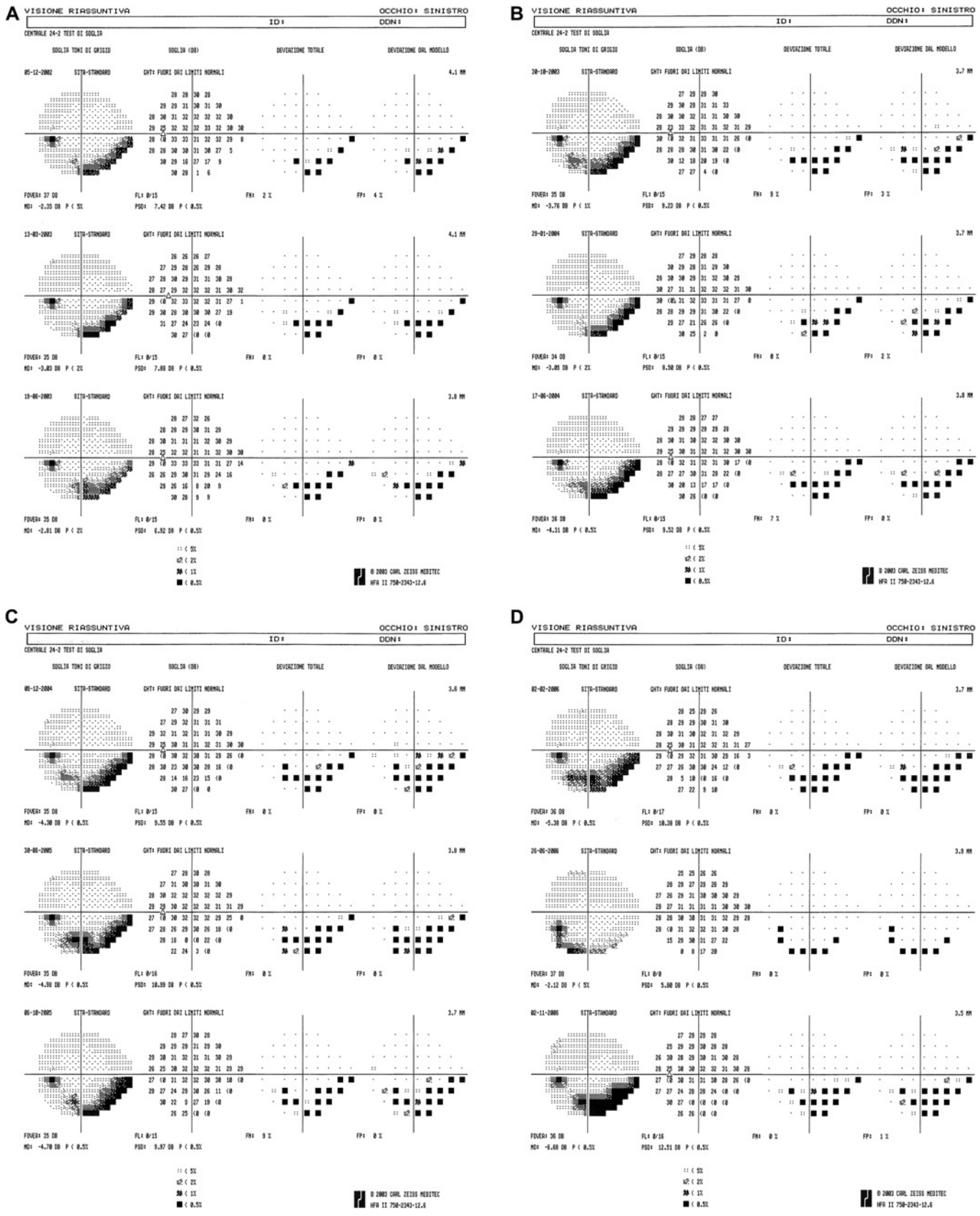

Figure 1 Overview printout of one patient included in the study.

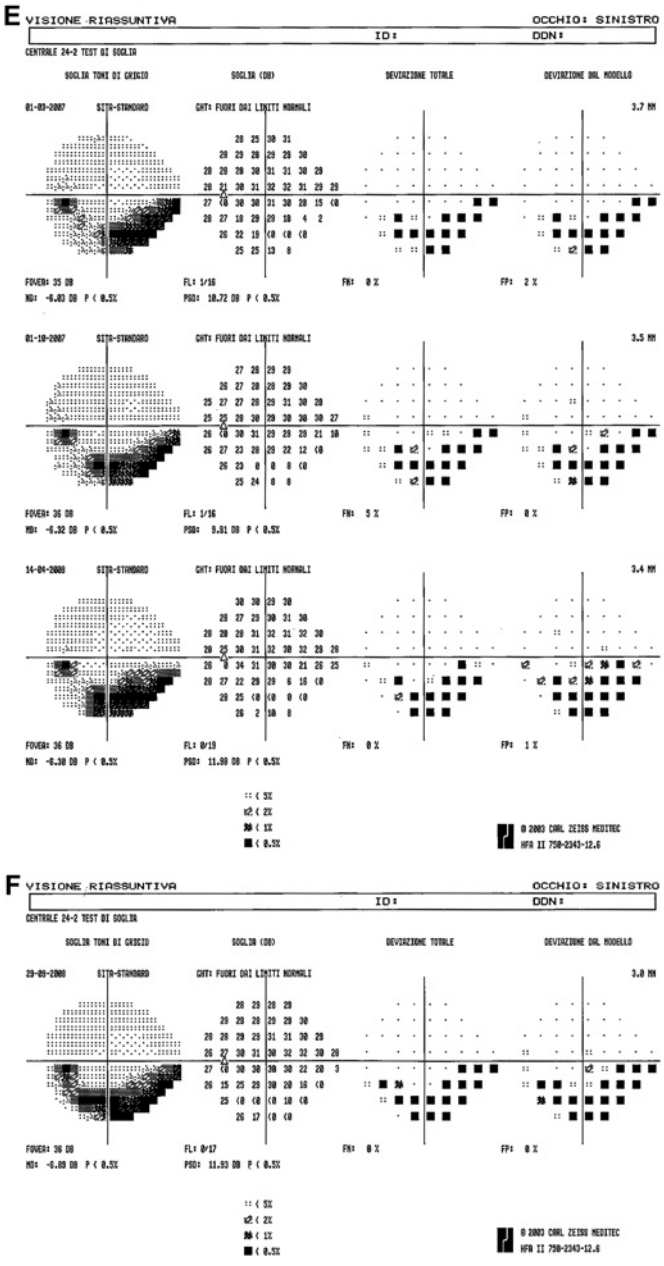

Figure 1 (Continued).

was considered reliable only when false-negative responses and fixation losses were less than 20%; unreliable VFs were not included in the analyses. Mean deviation (MD) and pattern SD were considered in the study to describe the included patients.

The database satisfied the following filters in order to obtain a group of VF series such as those commonly encountered in clinical practice when the degree of VF 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's age over 40 years; patients were expert in perimetry; each VF series consisted of 24-2 SITA Standard tests; continuous F/U data of at least 5 years and a minimum of 5 24-2 SITA standard VFs during a minimum of 5 years, without considering the very first three VFs; Full Threshold strategy was also accepted only for baseline tests; each VF test was required to be reliable according to the above mentioned criteria; presence of a typical glaucomatous VF was required.

Exclusion criteria were: concomitant ocular disease (eg, cataract); previous ocular surgery; systemic disease or medication known to affect the VF; refractive error exceeding 8 D spherical equivalent or 3 D of astigmatism, and visual acuity <20/50 at baseline or during the F/U.

Nine clinicians (PC, GC, PF, PF, MF, GM, PB, AP and MI) assessed the progression status of each VF series using standard HFA OP (which shows the grey scale maps, the threshold absolute values maps, and the total and pattern deviation probability plots), GPA printouts and GPA2 printouts. All of them were glaucoma specialists. They were all experienced in the interpretation of series of standard Humphrey visual field printouts in order to determine progression status, and were familiar with GPA and GPA2.

Clinicians were asked to judge the presence of stability or progression of each field series of the Humphrey printouts by using the three printouts, considering the first two baseline VF examinations and the behaviour of the glaucomatous defect over time (figures 1–3). Every time, the VF series were presented in random order, but obviously, each patient's VF remained in chronological order within a given field series. The velocity to classify VF changes was considered for each user and compared among the three different methods.

Guided progression analysis (GPA)

The GPA is statistical software available for the HFA for the evaluation of progression according to statistical criteria of the VF deterioration in glaucoma.[3 16] In order to improve some limits of the previous software (Glaucoma Change Probability Analysis), all the statistical evaluations about progression are carried out utilising the Pattern Deviation Plot values rather than the Total Deviation Plot values.[3 16]

The GPA software compares a patient's baseline visual fields to each subsequent VF in a series. The baseline values are obtained by the average of the two first exams. In each F/U field, every test point is evaluated relative to the baseline. The evaluation of progression is carried out comparing threshold modifications to a database of stable glaucoma patients who were tested over a very short period of time, taking into account fluctuation related to eccentricity and advancing disease. A single examination test with GPA Probability Plots printout, and the F/U overview printout are available.

Guided progression analysis 2 (GPA2)

The GPA2 is a new version of the Guided Progression Analysis that differs from the first in some characteristics: GPA2 uses the Visual Field Index (VFI) that allows a quantification of the VF according to the comparison of the defect depth with the normal age-adjusted visual field taking into account the functional damage related to eccentricity to correlate with ganglion cell density.[17] This parameter is less affected by cataract and other media changes and shows field status as a percentage. On the new GPA2 Summary report, VFI is used to quantify rate of progression, where it is plotted relative to patient age to calculate the rate of functional loss.

In the middle of the report the VFI Plot graphs the VFI values of all exams included in GPA analysis as a function of the patient's age. The VFI Plot also provides a linear regression analysis of the VFI over time. A minimum of five exams over 3 years must be included in GPA for the linear regression results to be presented.

Besides the VFI Plot, a histogram (the VFI Bar) indicates the patient's current VFI value. In addition, when the results of the regression analysis are displayed, the VFI Bar will also graphically indicate the 3–5 year projection of the linear regression

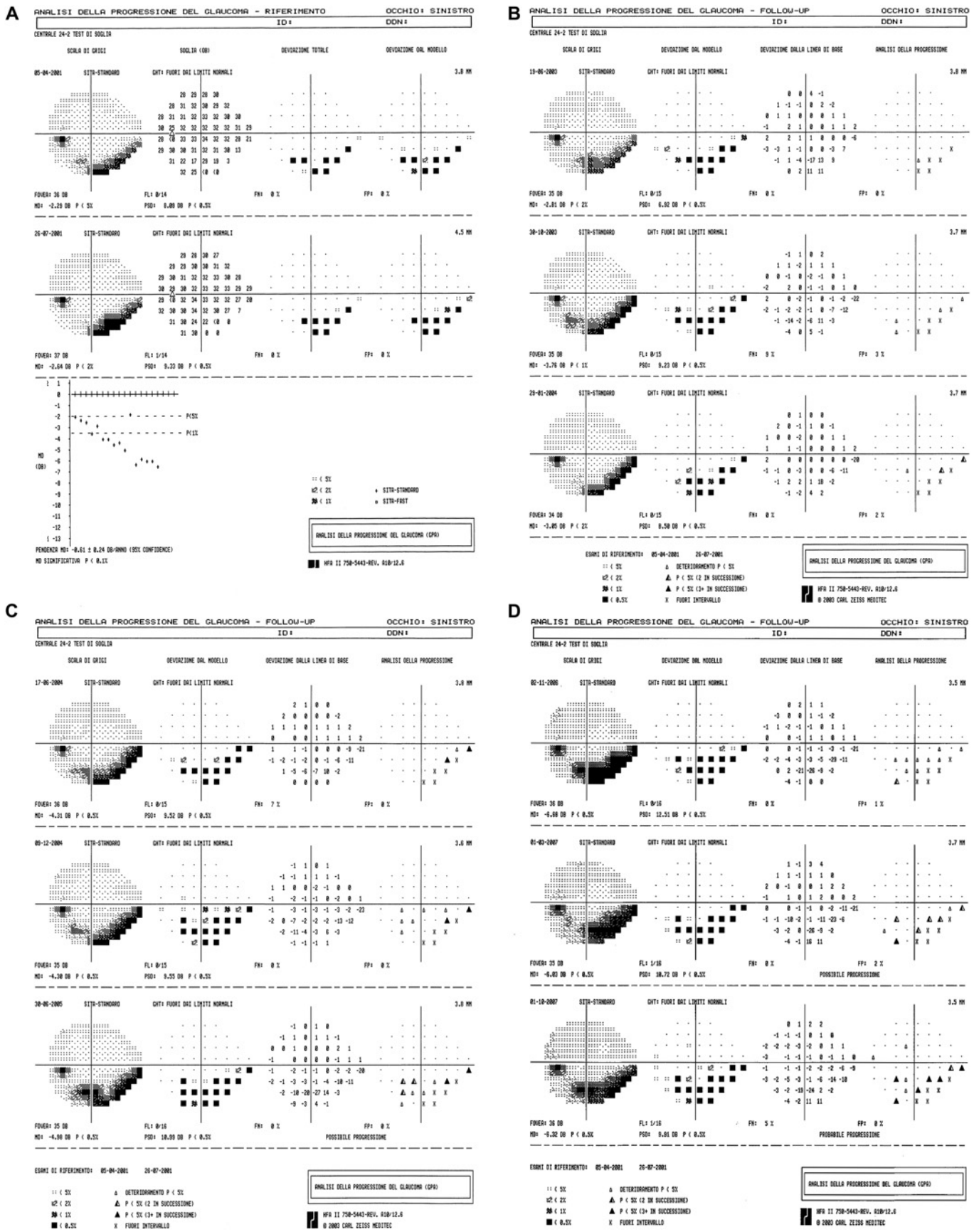

Figure 2 Guided Progression Analysis of one patient included in the study.

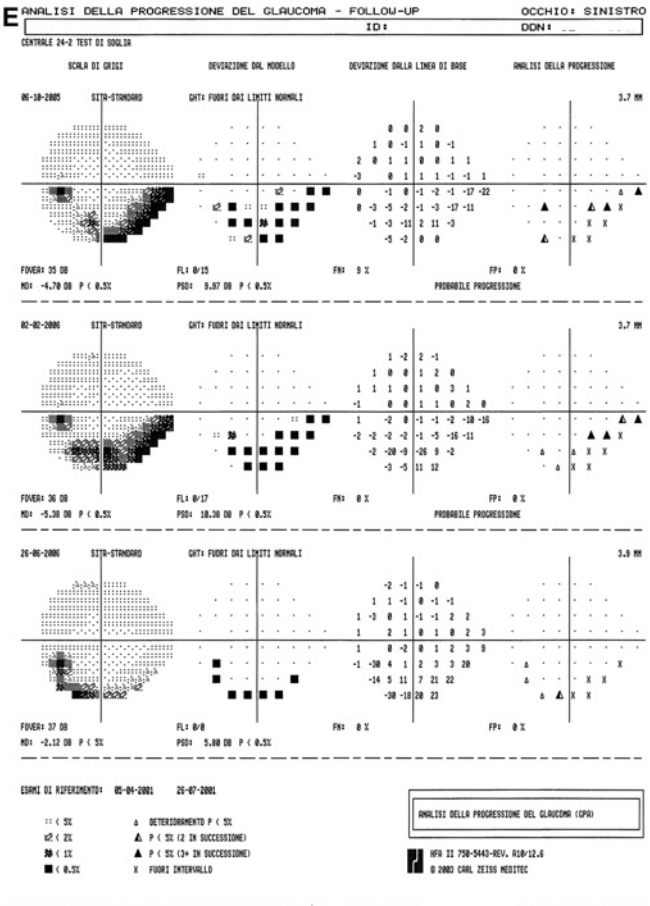

Figure 2 (Continued).

line, shown as a broken line. The length of projection is equal to the number of years of GPA data that is available, up to a maximum projection time of 5 years.

STATISTICAL ANALYSIS

The inter and intra-agreement among the nine clinicians (A, B, C, D, E, F, G, H, I) and the three different methods were analysed by using κ statistics[18]. Student t test was used to compare VF indices, while ANOVA test was used to compare the time used to assess all the printouts.

RESULTS

A total of 303 tests, comprising 38 VF series of 7.9±3.4 tests (mean±SD), were assessed by the nine glaucoma specialists. The mean F/U time was 6.16±1 years. The mean age was 73.8±13.43 years, all the patients were Caucasian and 53% were female. At baseline MD was −7.34±7.18 dB and at the end of the F/U was −9.25±8.65, and this change was statistically significant (p=0.02). At baseline PSD was 5.67±4.09 dB and at the end of the F/U it was 6.92±4.67, and this change was also statistically significant (p<0.001). The global agreement among observers and methods was 0.56.

When the time used to assess all the printouts was considered for each observer, statistically significant (p<0.001) differences were found among the three methods; in particular for the HFA OP the mean time was 31.25±5.1 min, for GPA it was 19.56±7.8 min and 10.19±2.7 min for GPA2.

When the intra-observer agreement was evaluated between HFA OP and GPA, the mean κ statistic was 0.58±0.13 (range 0.41−0.79), between HFA OP and GPA2, κ was 0.55±0.06 (range: 0.48−0.63) and between GPA and GPA2 it was 0.56±0.17 (range: 0.22−0.79). The details are listed in table 1.

When the inter-observer agreement was analysed, if HFA "overview printouts" were used, κ statistic was 0.65, when GPA was used, κ statistic was 0.54 and when GPA2 was used, κ statistic was 0.70. Agreement was calculated for all the possible pairs of the clinicians (table 2). The mean inter-observer agreement between HFA OP and GPA was 0.58, while between HFA OP and GPA2 it was 0.55, and between GPA and GPA2 it was 0.56.

DISCUSSION

The evaluation of progression of the defects in glaucomatous patients both in structure and/or in function is fundamental to decide the treatment. VFs are usually carried out by comparing only the global parameters, such as MD, PSD or loss variance,[5] but, even if these VF interpretative aids are important for the evaluation of progression in glaucoma clinics, the global assessment of the VF could be difficult and vague by using just a few numbers. The possibility to use all the data of the printout with different topographical maps that each instrument provides[3–8] is more fitting for a correct VF evaluation; however, the clinical judgement is often difficult owing to many variable factors: the short and long-term threshold fluctuation, that is more consistent in glaucoma even in patients with stable condition, the lack of defined criteria of progression, and the influence of cataract on the general sensitivity threshold.[9 19–23]

Different types of statistical software have been proposed by several instrument manufacturers, based on clinical studies, and the inter-observer disagreement, even among expert clinicians, is due to the different importance attributed to threshold modifications, the topography of deterioration, the time interval between tests and the influence of media transparency.[6 12 17 23]

In this study, three different methods to evaluate progression have been analysed and an interesting result was that when using GPA2 the time to review all the printouts was significantly shorter compared with the other two methods. These data could be due to the influence on the clinicians when they saw the graphical representation of the VFI, which could simplify the interpretation of the printout (ie, stable or progression), but make them lose some important details on the location and the depth of the defect. On the other hand, the OP took so much

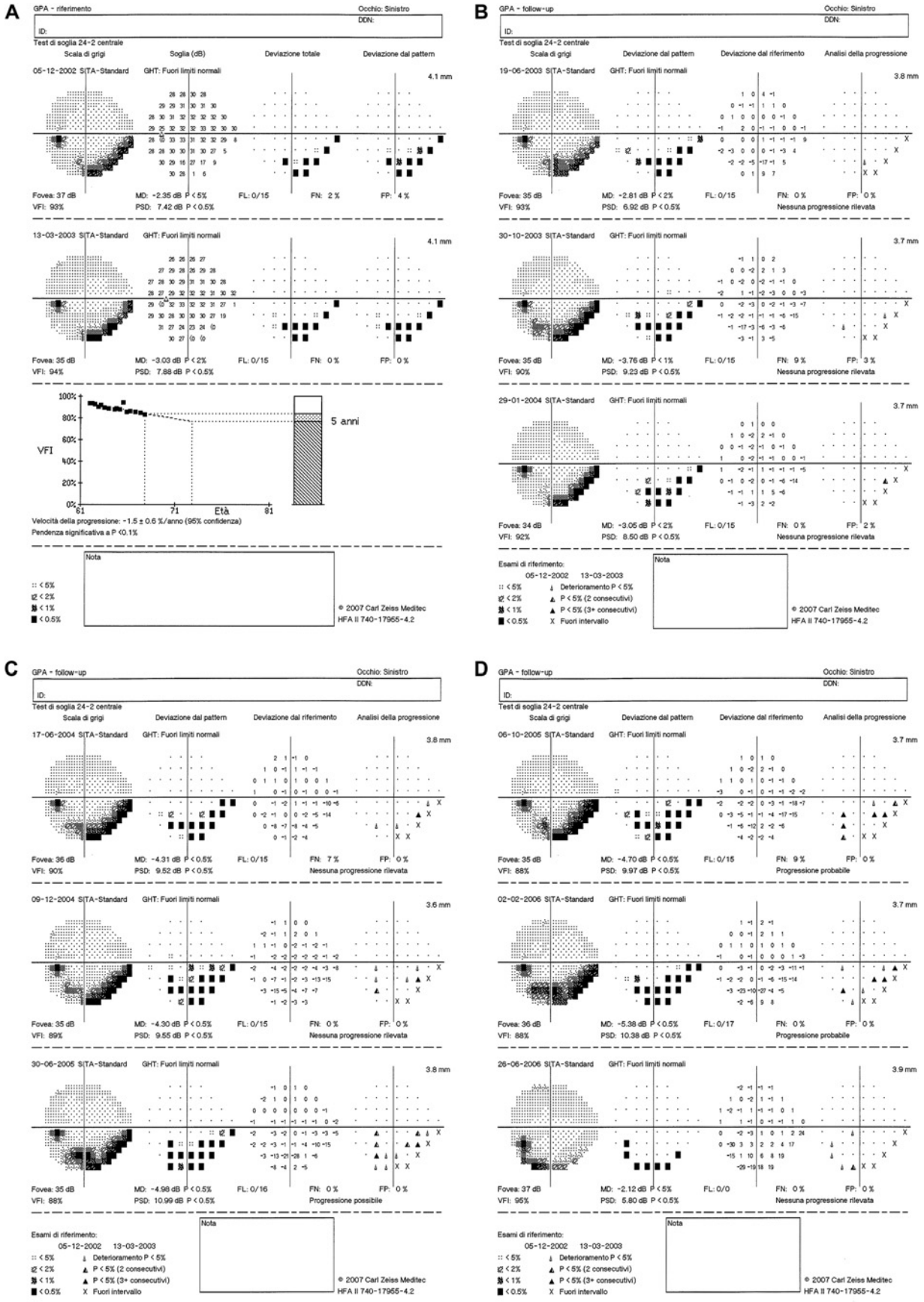

Figure 3 Guided Progression Analysis 2 of one patient included in the study.

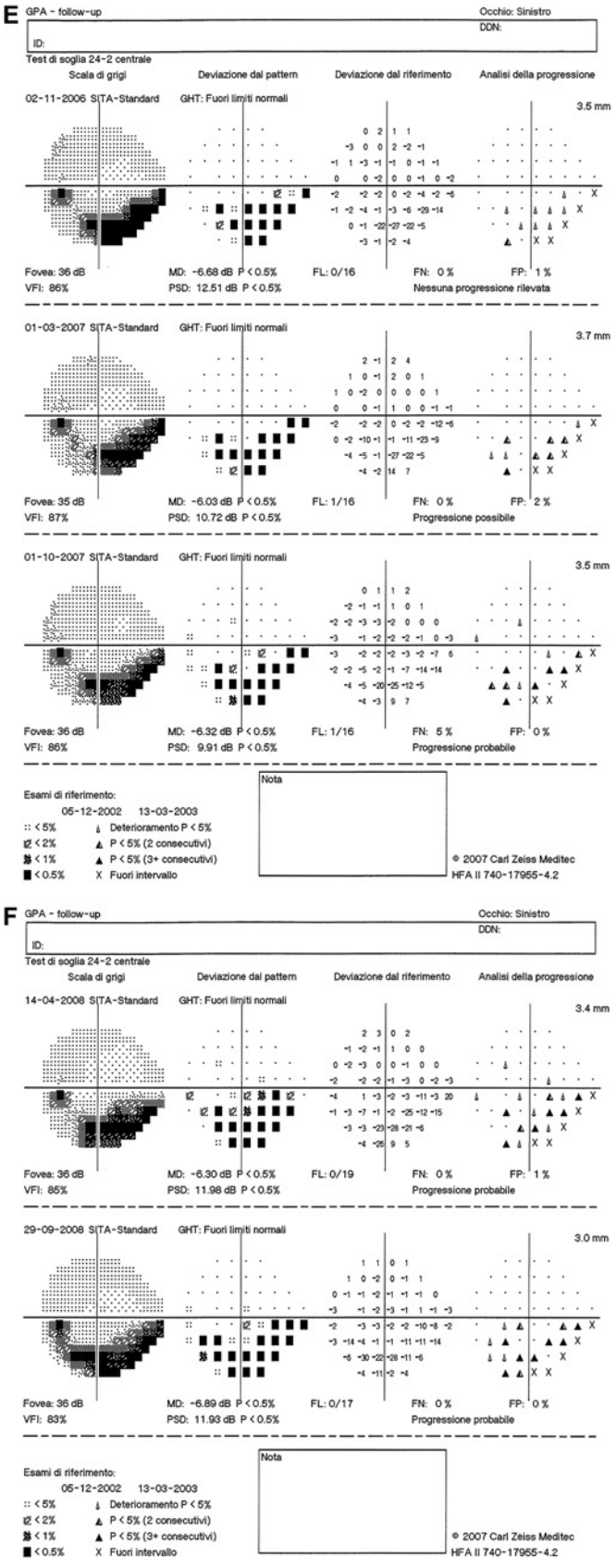

Figure 3 (Continued).

time because clinicians had to look for changes without any help from the software.

In our study, the intra-observer agreement was very similar among the nine observers (Table 1). Among the nine clinicians

Table 1 Intra-observer agreement for the nine clinicians

	A	B	C	D	E	F	G	H	I
OP versus GPA	0.62	0.58	0.74	0.41	0.43	0.53	0.47	0.79	0.64
OP versus GPA2	0.63	0.58	0.63	0.52	0.57	0.48	0.58	0.52	0.48
GPA versus GPA2	0.5	0.45	0.79	0.66	0.22	0.5	0.58	0.63	0.7

OP, overview printout; GPA, Guided Progression Analysis; GPA2, Guided Progression Analysis 2.

no agreement was previously set up about the criteria to adopt for the clinical evaluation of progression. All were glaucoma specialists. The good agreement among the nine clinicians in the evaluation according to the above mentioned criteria utilising the HFA OP was shown by the κ statistics whose values ranged from 0.46 to 0.70 after an analysis of 300 tests with a mean of 7.9 VFs per patient. When the HFA OP was used, the clinician's evaluation was based on the available plots (grey scale, threshold absolute values, total and pattern deviation plots) and the VF indices (MD, PSD, reliability indices) and they evaluated the severity of the sensitivity loss, the general or local depression, the topography of the defect and its worsening in deepness or extension.[9–11]

Viswanathan *et al* found that the number of VF tests was not very influenced by the agreement among clinicians. In clinical practice the evaluation of progression is usually carried out considering not only the first and the last VF examinations but, similarly to the criteria adopted by GPA, also the behaviour of defects in time independently from the number of examinations.[12] In a different study, Chauhan *et al* analysed the

Table 2 Inter-observer agreement for each method

	B	C	D	E	F	G	H	I	
A	0.58	0.79	0.42	0.74	0.55	0.74	0.79	0.63	OP
	0.22	0.38	0.45	0.3	0.34	0.47	0.42	0.55	GPA
	0.81	0.68	0.83	0.55	0.88	0.78	0.55	0.88	GPA2
	A	C	D	E	F	G	H	I	
B	0.58	0.68	0.63	0.74	0.74	0.74	0.79	0.63	OP
	0.22	0.52	0.72	0.64	0.73	0.58	0.74	0.72	GPA
	0.81	0.62	0.73	0.47	0.79	0.68	0.58	0.68	GPA2
	A	B	D	E	F	G	H	I	
C	0.79	0.68	0.42	0.74	0.61	0.53	0.58	0.42	OP
	0.38	0.52	0.31	0.48	0.25	0.53	0.37	0.46	GPA
	0.68	0.62	0.77	0.61	0.82	0.72	0.61	0.69	GPA2
	A	B	C	E	F	G	H	I	
D	0.42	0.63	0.42	0.47	0.58	0.69	0.53	0.48	OP
	0.45	0.72	0.31	0.49	0.71	0.64	0.68	0.77	GPA
	0.83	0.73	0.77	0.62	0.94	0.95	0.73	0.83	GPA2
	A	B	C	D	F	G	H	I	
E	0.74	0.74	0.74	0.47	0.69	0.69	0.84	0.58	OP
	0.3	0.64	0.48	0.49	0.59	0.63	0.68	0.49	GPA
	0.55	0.47	0.61	0.62	0.56	0.68	0.35	0.44	GPA2
	A	B	C	D	E	G	H	I	
F	0.55	0.74	0.61	0.58	0.69	0.79	0.84	0.68	OP
	0.34	0.73	0.25	0.71	0.59	0.43	0.58	0.55	GPA
	0.88	0.79	0.82	0.94	0.56	0.89	0.67	0.88	GPA2
	A	B	C	D	E	F	H	I	
G	0.74	0.74	0.53	0.69	0.69	0.79	0.74	0.79	OP
	0.47	0.58	0.53	0.64	0.63	0.43	0.74	0.64	GPA
	0.78	0.68	0.72	0.95	0.68	0.89	0.68	0.77	GPA2
	A	B	C	D	E	F	G	I	
H	0.79	0.79	0.58	0.53	0.84	0.84	0.74	0.63	OP
	0.42	0.74	0.37	0.68	0.68	0.58	0.74	0.68	GPA
	0.55	0.58	0.61	0.73	0.35	0.67	0.68	0.55	GPA2

OP, overview printout; GPA, Guided Progression Analysis; GPA2, Guided Progression Analysis 2.

glaucoma progression of 32 VF series using a computer animated graphic technique corrected for test–retest variability: they found a κ value of 0.572.[7]

More attention is currently given to the confirmation of the behaviour of defects in subsequent examinations. In a recent study, Chauhan *et al* showed that to detect a 4 dB VF change in MD, six visual fields in 2 years were needed.[24]

When the HFA OP was used, the better agreement between clinicians was probably due to the different criteria of VF progression adopted by the observers, who considered deterioration not only when more than three points reached the GPA progression significance in at least three examinations, but also other factors, such as their location in critical areas for glaucoma in the visual fields. Furthermore, the GPA printout did not show the Total Deviation Plot which allowed the visualisation of the total sensitivity loss, sometimes representing a sign of glaucomatous visual field deterioration. However, Bengtsson and Heijl showed that all the glaucomatous localised defects were associated with a different level of diffuse damage and in the Pattern Deviation Probability map the probability symbols increased up to a MD of −20 dB, but when the MD was more than −20 dB the Total Deviation map was more useful than the pattern deviation map.[17] On the other hand, the GPA2 printout did not show the Pattern Deviation Plot; however, it was found to be the fastest method because the observers used mainly the VFI regression graph to classify the visual fields, while the GPA alert was not considered much by clinicians.

Both GPA programs were useful for the identification of points to watch for which could significantly progress, but it could not replace the clinical evaluation of the morphological modifications of the defects. Actually there is no internationally validated method to detect progression. The different available software are a step toward the improvement of this methodology[6]; however, the clinical judgement is still fundamental to interpret the results of these new methods.

Our analysis strongly depends on the quality of the evaluators; in particular in this study, glaucoma specialists were involved. A lower κ-statistic value would be expected if general ophthalmologists were considered. However, it is likely that teaching programs or the development of common strategies and procedures to evaluate progression over a VF series would further improve the agreement. These issues were outside the aim of our study, but they deserve consideration because they could improve the management of glaucoma patients.

Competing Interest None to declare.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Leske MC,** Heijl A, Hussein M, *et al*. Factors for glaucoma progression and the effect of treatment: the early manifest glaucoma trial. *Arch Ophthalmol*. 2003;**121**:48–56.
2. **Advanced Glaucoma Intervention Study Investigators.** The Advanced Glaucoma Intervention Study (AGIS), 7: the relationship between control of intraocular pressure and visual field deterioration. *Am J Ophthalmol* 2000;**130**:429–40.
3. **Heijl A,** Leske MC, Bengtsson B, *et al*; Early Manifest Glaucoma Trial Group. Reduction of intraocular pressure and glaucoma progression: results from the Early Manifest Glaucoma Trial. *Arch Ophthalmol* 2002;**120**:1268–79.
4. **Smith SD,** Katz J, Quigley HA. Analysis of progressive change in automated visual fields in glaucoma. *Invest Ophthalmol Vis Sci* 1996;**37**:1419–28.
5. **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.
6. **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.
7. **Chauhan BC,** Drance SM, LeBlanc RP, *et al*. Technique for determining glaucomatous visual field progression by using animation graph. *Am J Ophthalmol* 1994;**118**:485–91.
8. **Fitzke FW,** Hitchings RA, Poinoososawmy D, *et al*. Analysis of visual field progression in glaucoma. *Br J Ophthalmol* 1996;**80**:40–8.
9. **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.
10. **Boeglin RJ,** Caprioli J, Zulauf M. Long-term fluctuation of visual field in glaucoma. *Am J Ophthalmol* 1992;**113**:396–400.
11. **Katz J.** Scoring systems for measuring progression of visual field loss in clinical trials of glaucoma treatment. *Ophthalmology* 1999;**106**;391–5.
12. **Viswanathan AC,** Crabb DP, McNaught AI, *et al*. Inter-observer agreement on visual field progression in glaucoma: a comparison methods. *Br J Ophthalmol* 2003;**87**:726–30.
13. **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.
14. **European Glaucoma Society.** *Terminology and Guidelines for Glaucoma 2008*. 3rd ed. Savona: Dogma, 2008;(Ch 2):86–7.
15. **Iester M,** Swindale NV, Mikelberg FS. Sector-based analysis of optic nerve head shape parameters and visual field indices in healthy and glaucomatous eyes. *J Glaucoma* 1997;**6**:371–6.
16. **Heijl A,** Lindgren G, Olsson J, *et al*. Visual field interpretation with empiric probability maps. *Arch Ophthalmol* 1989;**107**:204–8.
17. **Bengtsson B,** Heijl A. A visual field index for calculation of glaucoma rate of progression. *Am J Ophthalmol* 2008;**45**:343–53.
18. **Landis JR,** Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics* 1977;**33**:363–74.
19. **Grødum K,** Heijl A, Bengtsson B. A comparison of glaucoma patients identified through mass screening and in routine clinical practice. *Acta Ophthalmol Scand* 2002;**80**:627–31.
20. **Arnalich-Montiel F,** Casas-Llera P, Muñoz-Negrete FJ, *et al*. Performance of glaucoma progression analysis software in a glaucoma population. *Graefes Arch Clin Exp Ophthalmol* 2009;**247**:391–7.
21. **Diaz-Aleman VT,** Anton A, de la Rosa MG, *et al*. Detection of visual-field deterioration by Glaucoma Progression Analysis and Threshold Noiseless Trend programs. *Br J Ophthalmol* 2009;**93**:322–8.
22. **Casas-Llera P,** Rebolleda G, Munoz-Negrete FJ, *et al*. Visual field index rate and event-based glaucoma progression analysis: comparison in a glaucoma population. *Br J Ophthalmol* 2009;**93**:1576–9.
23. **Perdicchi A,** Iester M, Scuderi G, *et al*. Visual field damage and progression in glaucomatous myopic eyes. *Eur J Ophthalmol* 2007;**17**:534–7.
24. **Chauhan BC,** Garway-Heath DF, Goñi FJ, *et al*. Practical recommendations for measuring rates of visual field change in glaucoma. *Br J Ophthalmol* 2008;**92**:569–73.

Agreement to detect glaucomatous visual field progression by using three different methods: a multicentre study

M Iester, E Capris, F De Feo, et al.

Br J Ophthalmol 2011 95: 1276-1283 originally published online December 3, 2010
doi: 10.1136/bjo.2010.189456