Agreement Among Spectral-Domain Optical Coherence Tomography, Standard Automated Perimetry, and Stereophotography in the Detection of Glaucoma Progression

Sebastián A. Banegas,[1] Alfonso Antón,[2] Antonio Morilla-Grasa,[2] Marco Bogado,[2] Eleonora M. Ayala,[2] and Javier Moreno-Montañes[3]

[1]Institut Oftalmòlògic Quirón Barcelona, Spain
[2]Institut Català de Retina, Universidad Internacional de Cataluña, Parc Salut Mar, Barcelona, Spain
[3]Clínica Universidad de Navarra, Pamplona, Spain

Correspondence: Sebastián A. Banegas, Institut Oftalmològic Quirón Barcelona, Av/Diagonal 632, Barcelona 08017, Spain; sebanegas@hotmail.com.

Submitted: June 10, 2014
Accepted: December 16, 2014

Citation: Banegas SA, Antón A, Morilla-Grasa A, Bogado M, Ayala EM, Monero-Montañes J. Agreement among spectral-domain optical coherence tomography, standard automated perimetry, and stereophotography in the detection of glaucomatous progression. *Invest Ophthalmol Vis Sci.* 2015;56:1253–1260. DOI:10.1167/iovs.14-14994

Purpose. The purpose of this study was to evaluate the agreement among spectral-domain optical coherence tomography, standard automated perimetry (SAP), and optic disc stereo photography in the detection of glaucomatous progression.

Methods. This was an observational cohort study enrolling 246 eyes (148 patients) followed for an average of 31.8 ± 9 months. Images were obtained every 6 to 12 months with optical coherence tomography (OCT), visual field test, and optic disc stereo photography. Progression was determined with OCT using guided progression analysis (GPA) software, in perimetry with Humphrey field analyzer GPA, and by masked assessment of stereo photograph series. Agreement among methods was reported using the κ coefficient, prevalence-adjusted bias-adjusted κ (PABAK), Gwet's first-order agreement coefficient (AC1), overall percentage agreement (OPA), percentage of positive agreement (Ppos), and percentage of negative agreement (Pneg).

Results. Progression by stereo photos, SAP, and OCT was found in 17 eyes (6.9%), 37 eyes (15%), and 63 eyes (25.6%), respectively. Most cases with detectable changes were only identified by one examination method, resulting in low Ppos (<33%). On the contrary, 147 eyes (59.7%) were identified as nonprogressing cases by all three methods, showing high OPA (72.8–89.8) and high Pneg (83.8–94.5). PABAK and AC1 between methods reached 0.67 to 0.88. Measurements of agreement showed a trend toward better agreement between photos and visual field (VF) than between photos and OCT. Spectral-domain OCT parameters reflected a tendency toward better agreement with stereo photos than with VF.

Conclusions. Methods obtained acceptable agreement outcomes in terms of PABAK, AC1, and OPA. However, most cases with detectable changes were identified only by one examination method, resulting in low Ppos.

Keywords: agreement, glaucoma, optical coherence tomography, progression

Early detection and control of disease progression or worsening is one of the most important and difficult challenges in glaucoma management. Conventional methods used to identify progression, such as stereophotography or standard automated perimetry (SAP), differ in their abilities to identify changes and results show poor agreement.[1–6] One of the limitations of functional tests like SAP when performing the task of progression detection is high test-retest variability.[7,8] Imaging devices are less dependent on patient performance, and in particular, the latest spectral-domain optical coherence tomography (SD-OCT) versions offer highly reproducible measurements.[9] Many studies found excellent intravisit and intervisit reproducibility of peripapillary retinal nerve fiber layer (RNFL) measurements based on the interclass correlation coefficient, coefficient of variation, and test-retest standard deviation in normal subjects and glaucoma patients.[10,11] This high reproducibility offers a good theoretical basis to enable the detection of structural changes over time. Progression algorithms have recently been applied to imaging technologies such as scanning laser polarimetry (GDX), confocal scanning laser ophthalmoscopy (Heidelberg retinal tomography [HRT]) and OCT. These algorithms could be useful tools to identify progression not requiring active patient participation.[12–20] However, they require thorough clinical evaluation to determine their capabilities and limitations.

The purpose of this study was to evaluate agreement between SD-OCT guided progression analysis (GPA) with Humphrey field analyzer GPA software and with expert assessment of optic disc stereo photos for detection of glaucomatous progression. Will OCT and a series of stereophotographs, both structural diagnostic tests, show a trend toward better agreement than with a functional test such as SAP? Do OCT event parameters have better agreement in identifying progression with visual field (VF) event analysis than

with VF trend algorithm? How do OCT event and trend parameters' rates of progression detection differ? In this study, we sought to answer theses queries. To the best of our knowledge, this is the first longitudinal study to report a thorough appraisal of all Cirrus SD-OCT (software version 5.0; Carl Zeiss Meditec, Dublin, CA, USA) GPA parameters agreement with well-established modalities such as stereo photos and SAP.

Materials and Methods

This was an observational cohort study. A total of 246 eyes of 148 patients were participants on a larger prospective study conceived to assess diagnostic performance of different structural and functional modalities in glaucoma and were chosen on the basis of minimum number of examinations with VF, SD-OCT, and stereo photos (discussed below).

Patients with glaucoma or suspected of having glaucoma and healthy patients were consecutively enrolled as they presented at our clinic, Institùt Català de Retina at Barcelona, Spain, and were evaluated by established protocol series of structural and functional diagnostic procedures. Reasons for the visit for glaucoma and suspects were scheduled first or follow-up visit. Healthy subjects were identified among patients who scheduled visits to rule out glaucoma (family history), loss of vision, or were suspected of having cataract.

Medical history and full ophthalmic examinations were performed at baseline and during follow-up visits. Findings were electronically recorded on a designed glaucoma visit form. Ocular examinations included visual acuity (Early Treatment Diabetic Retinopathy Study chart), manifest refraction, Goldmann applanation tonometry, slit-lamp biomicroscopy, gonioscopy, and dilated fundus evaluation using a 78-diopter (D) hand-held lens. Visual fields with 24-2 Swedish interactive threshold algorithm (SITA) standard strategy (Carl Zeiss Meditec, Jena, Germany) were performed every 6 months and optic disc stereo photos were acquired every year with a model TRC-NW7SF fundus camera (Topcon, Tokyo, Japan). Images of the optic disc and peripapillary region were obtained annually with Cirrus SD-OCT (Carl Zeiss Meditec) using the optic disc Cube 200 × 200 scan protocol.

In the present study, inclusion criteria were 18 years of age or older, a best-corrected visual acuity of 20/30 or better, spherical refractive error of +5 to −5 D, less than 3 D of cylinder, and open-angle defined as ≥3 in the Shaffer classification. Subjects were excluded if they had a history of surgery or laser retina procedures or any ocular or systemic disease that could cause optic disc or retina abnormalities or loss of VF. Patients with severely depressed VFs in which Humphrey field analyzer (HFA) GPA could not perform both event and trend analyses were excluded. There were no exclusion criteria regarding type of glaucoma. Patients with mild cataract were included.

Examinations

After subsequent visits, the sample was divided into 4 groups: glaucoma, suspected glaucoma, ocular hypertension, and normal, based on two subsequent fields: optic disc appearance and ocular pressure. Glaucoma was considered if eyes presented consistent glaucomatous visual field defects with spatially corresponding optic nerve damage on more than two consecutive occasions. Glaucomatous visual field defect was defined as having three or more significant nonedge-contiguous points outside the 95% normal limits in the pattern deviation plot ($P < 5\%$), with at least 1 at the level of $P < 1\%$ in at least two consecutive examinations, or a glaucoma hemifield test result (GHT) outside normal limits. Optic nerve damage was defined by the presence of a localized notch, rim thinning, a cup-to-disc ratio of >0.8 or a difference between these ratios in each eye of >0.2, disc hemorrhage, or RNFL defects. Eyes classified as suspected of having glaucoma had normal visual fields but glaucomatous optic disc appearance. Patients with ocular hypertension (OHT) had high IOP (>21 mm Hg), normal fields, and normal discs in three different consecutive examinations. Normal subjects had no history of high IOP (>21 mm Hg) in three different consecutive examinations, reliable normal visual field with no defects outside 95% normal limits in the pattern deviation plot, a GHT result within normal limits at two consecutive examinations, and normal fundus examination results.

During the follow-up period, each patient was treated at the discretion of the attending ophthalmologist. The study was conducted according to the principles of the Declaration of Helsinki. All subjects gave informed consent, and the study was approved by the Ethics Committee of the Hospital Universitari Sagrat Cor.

Main outcome measurements were identification of changes in stereo photos, VF, and SD-OCT.

1. Stereo photos were assessed by two independent expert graders. These experts were masked to both patient data and temporal sequence of the images. Baseline and follow-up photographs were compared using a stereoscopic viewer (Screen-Vu PS Mfg., Portland, OR, USA). The experts were asked to look for rim thinning, increase in cup-to-disc ratio or disc asymmetry, disc hemorrhage, or onset or enlargement of retinal nerve fiber defects. Graders were required to identify chronological sequences correctly in order to classify a progressing case. If consensus was not reached, a third independent grader decided whether progression was present;
2. Visual fields. Only VF series with at least five reliable tests (fixation losses of <30%, false-positive rates of <20%, and false-negative rates of <30%) were included in the analysis. Guided progression analysis event and trend analyses were used to assess the presence of VF change. Guided progression analysis event analysis compares the changes between baseline and follow-up examinations. Significant progression is considered if the change detected is greater than the deterioration expected to occur less than 5% of the time at the location in a stable glaucoma patient ($P < 0.05$). Guided progression analysis alert messages of possible progression (significant deterioration in three locations in two consecutive tests) or likely progression (significant deterioration in three locations in three consecutive tests) in the last VF visit, were classified as VF progression. Guided progression analysis trend analysis was applied to calculate the rate of progression of VF index (VFI) with a regression equation represented at the VFI plot.[21] Significant negative slopes with P values of <5% were considered VF progression; and
3. Change in SD-OCT images was analyzed with SD-OCT GPA version 5.0 software. Only scans with signal strength of >5 that were well centered and without artifacts were included. Spectral-domain OCT GPA also used event and trend analyses to compare RNFL thickness in follow-up images to that of baseline images (event analysis) and determined whether a statistically significant change occurred since baseline. The software also calculated the rate of change (trend analysis). The first two good quality scans were established as baseline examinations. A minimum of three examinations were

needed to generate a GPA printout. If a statistically significant change was detected once in follow-up scans, a yellow alert message of possible progression was indicated. When significant thinning was noted in two consecutive examinations, a red likely progression message was indicated. Both of the alert messages, if observed in any of the progression parameters in the last OCT acquirement visit, were considered progression in SD-OCT. The following parameters were analyzed: event analysis: RNFL thickness map progression (thickness map), and RNFL thickness profile progression (thickness profile). Trend analysis: average RNFL thickness progression (average thickness), overall thickness, superior thickness and inferior thickness. Figures 1A and 1B show an example of good agreement between changes detected by VF and OCT.

Statistical Analysis

Statistical analyses were performed using SPSS version 20.0 software (SPSS, Inc., Chicago, IL, USA) and WINPEPI (PEPI-for-Windows version 11.15; Microsoft, Redmond, WA, USA). In all analyses, a P value of <0.05 was considered statistically significant. A Venn diagram showed cases of progression identified by the different tests and baseline group classification of these cases (Fig. 2). Agreement among the different methods was reported using the κ coefficient, adjusted prevalence-adjusted bias-adjusted κ (PABAK), Gwet's first-order agreement coefficient (AC1), overall percentage agreement (OPA), Cicchetti-Feinstein indices percentage of positive agreement (Ppos), and percentage of negative agreement (Pneg). Table 1 presents the formulas used in this study.

Kappa coefficient is a chance-corrected statistic widely used for measuring the level of agreement between raters. Kappa ranges from -1 (perfect disagreement) to $+1$ (perfect agreement), when κ has the value zero, it indicates no agreement better than chance. It is highly dependent on the prevalence of the condition and decreases greatly where there is an unbalanced predominance of either "yes" or "no" answers.[22,23] For this reason, evaluation of agreement using other indices is recommended. PABAK is used to adjust κ in unbalanced situations. PABAK also ranges from -1 to $+1$ and assumes an average of the prevalence of each category of the two raters and absence of bias and is more consistent with OPA than κ. Gwet's AC1 is a novel and robust chance-corrected statistic based upon the assumption that only a portion of the observed ratings will potentially lead to agreement by chance. This index is also less influenced by differences in the propensity to give positive ratings and differences in prevalence of the response category, avoiding paradoxical results.

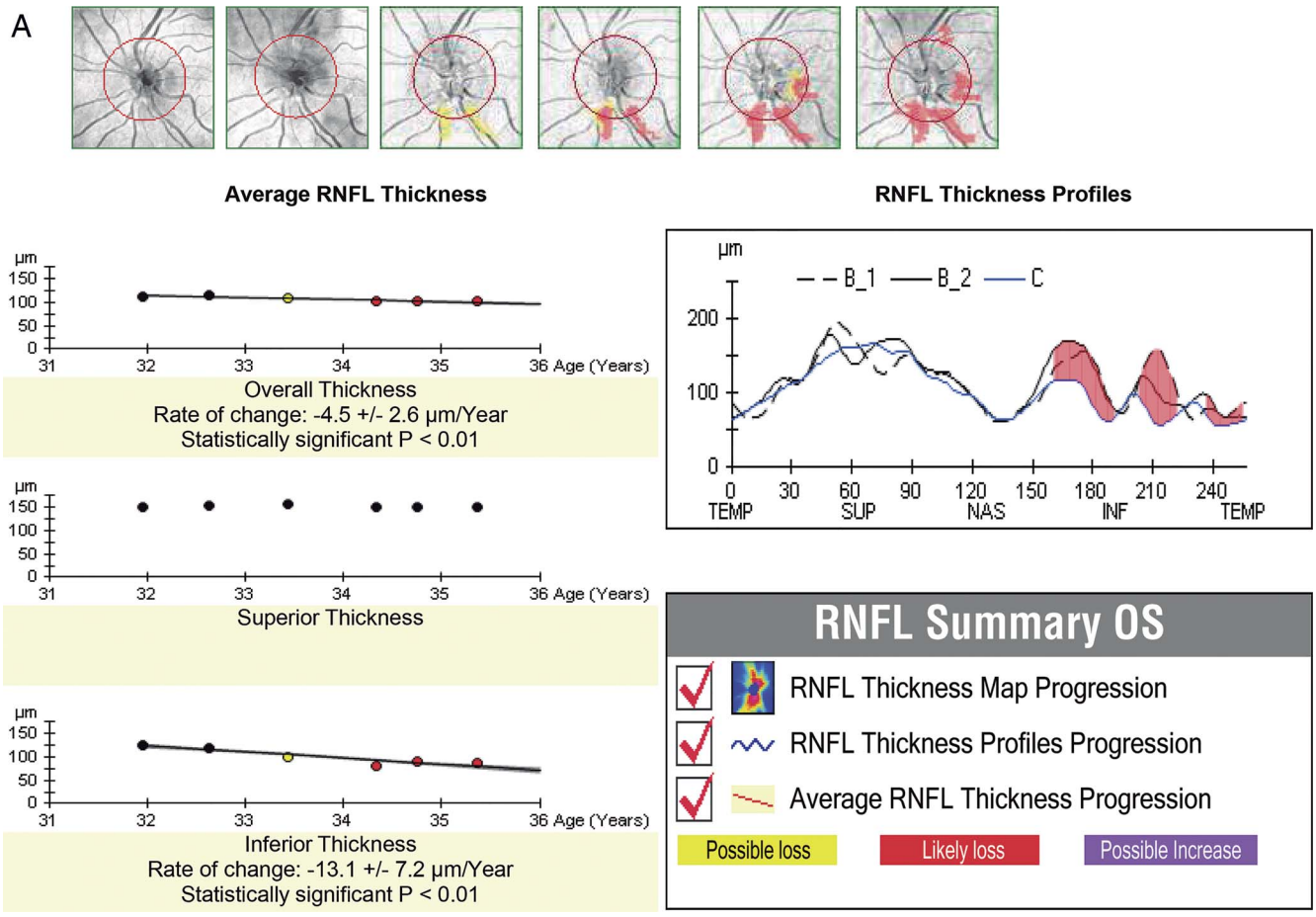

FIGURE 1. (A, B) Example of agreement detecting glaucomatous progression between VF and OCT. Visual field's GPA alert detected significant deterioration in three consecutive tests at several points of the superior hemifield, indicating likely progression. Trend analysis of the patient's VFI rate of progression, performed by regression line at the visual field index plot, indicated negative significant slope, P, of $<1\%$, considered glaucomatous visual field progression. Significant RNFL loss in the inferior sector was detected by OCT's event parameters, thickness map progression, and thickness profiles progression; significant rates of change by trend parameters, overall thickness, and inferior thickness progression.

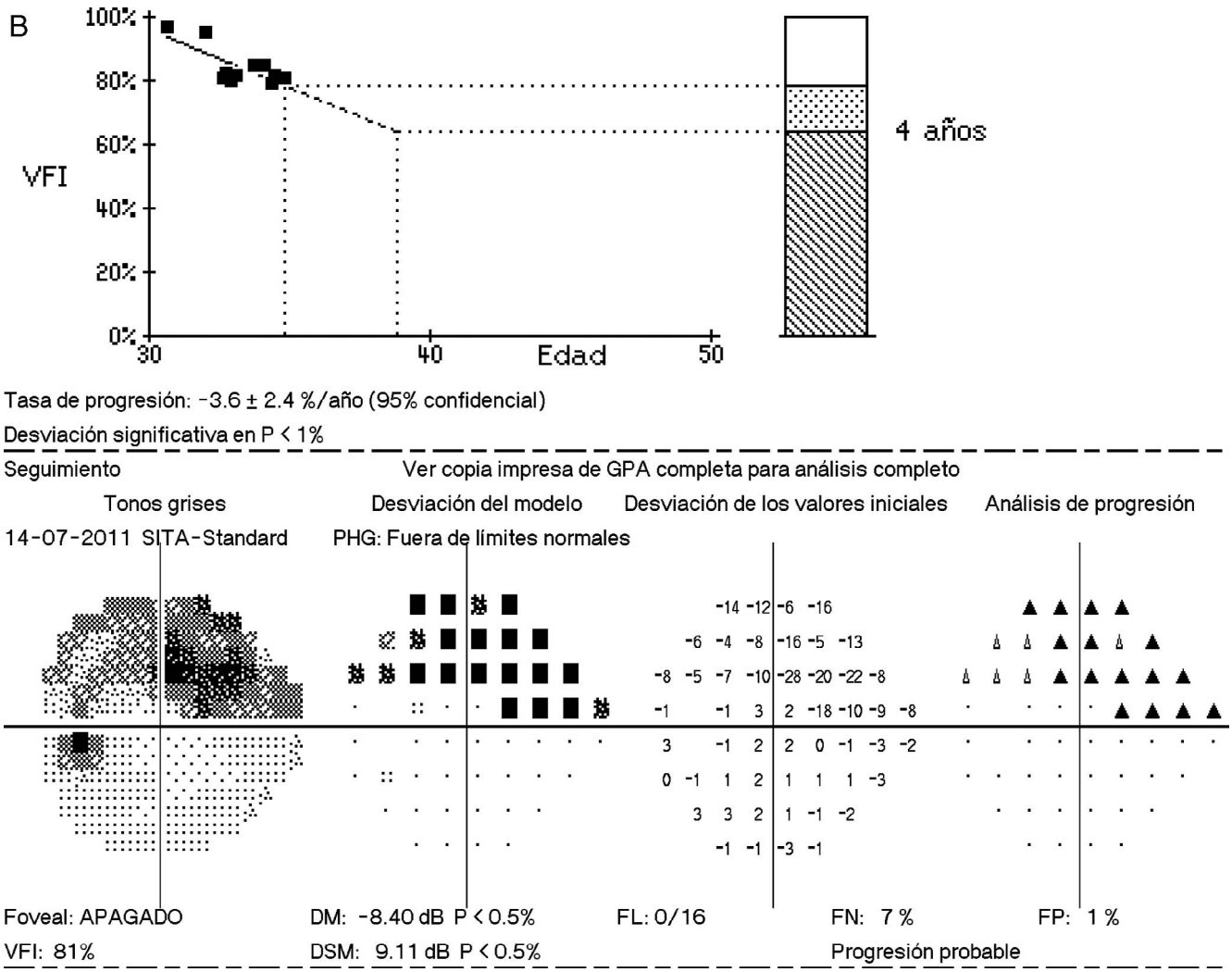

Figure 1. Continued.

Gwet's first-order agreement coefficient also ranges in value from −1 to +1.[22] Overall percentage agreement ranging from 0% to 100% is not corrected for chance agreement and tends to be high when there is an unbalanced prevalence of the condition (very high or low). In addition, OPA does not differentiate between agreement on the positives and agreement on the negatives. Cicchetti-Feinstein indices Ppos and Pneg (ranging from 0%–100%) are measurements proposed to overcome the limitations of OPA. They are concordant ratings, positive or negative as a percentage of all positive or negative ratings. They represent the probability that if a subject has been given a certain rating by a typical observer, a second typical observer would assign the same rating.[22]

Results

A total of 276 patients were chosen on the basis of a minimum number of VF, SD-OCT, and stereo photography examinations

Table 1. Formulas Used in This Study

Name	Formula	Rater B	Rater A (GS) Positive	Rater A (GS) Negative	Total
κ	$[(a+d)/n] - (rt + su/n^2)/\ [1 - (rt + su/n^2)]$	Positive	a	b (FP)	r
PABAK	$2[(a+d)/n] - 1$	Negative	c (FN)	d	s
Gwet's AC1	$[(a+d/n) - (t + r/2n)]/\ 1 - (t + r/2n)$	Total	t	u	n
OPA	$(a+d)/n$				
Ppos	$2a/(2a+b+c)$				
Pneg	$2d(2d+b+c)$				

κ, kappa; GS, Gold Standard; a, true positive; b (FP), false positive; c (FN), false negative; d, true negative; t, total number of defectives by gold standard; u, total number of not defectives by gold standard; r, total number of positive test results by Rater B; s, total number of negative test results by Rater B; n, total number of test results.

Table 2. Subjects Characteristics*

Characteristic	Normal	OHT	Glaucoma Suspected	Glaucoma	Sigma *P*
No. of eyes (no. of subjects)	17 (14)	69 (41)	63 (41)	97 (60)	0.000
Age, y (±SD)	54.3 (8.2)	62.5 (11.5)	66 (10.6)	69.37 (12.8)	0.001
Follow-up, mo (±SD)	31.84 (9.6)	30.54 (7.7)	32.56 (8.9)	32.12 (9.1)	0.381
Paquimetry, μ (±SD)	540.36 (38.5)	553.23 (40.9)	540.1 (35.7)	533.45 (30.9)	0.007
Median no. of OCT examinations (1st Q/3rd Q)	3 (3/4)	3 (3/4)	3 (3/4)	3 (3/4)	0.631
Median no. of visual field examinations (1st Q/3rd Q)	5 (5/6)	6 (5/7)	6 (5/8)	7 (6/9)	0.000
Median no. of stereo photos (1st Q/3rd Q)	2 (2/3)	2 (2/2)	2 (2/3)	2 (2/3)	0.287
Baseline average RNFL thickness, μ (±SD)	90.88 (8.1)	88.17 (9.1)	83.56 (11.8)	75.26 (13.4)	0.000
Baseline MD, dB (±SD)†	0.5 (0.9)	−0.37 (1.7)	−1.66 (3.5)	−4.27 (5.5)	0.000

MD, mean deviation; Q, quartile; Sigma *P*, Kruskal-Wallis test for nonnormally distributed data.
* $N = 148$ (59 male; 89 female).

(five, three, and two, respectively). However, 128 participants were excluded because of poor quality test results (110 patients were excluded because of poor quality OCT results, 15 because of poor VF results, and 3 because of poor stereophotography results). In the current study, 246 eyes of 148 subjects (59 males, 89 females) were recruited and followed for an average of 31.8 ± 9 months. Patient mean age was 65.1 ± 12.25 (range, 34–86 years of age). A total of 97 eyes (39.4%) were classified as glaucomatous, 63 eyes (25.6%) were suspected of being glaucomatous, 69 eyes (28%) had OHT, and 17 eyes (6.9%) were classified as normal. The demographic characteristics of the study sample are summarized in Table 2.

Progression by stereo photos was found in 17 eyes (6.9%). Agreement between graders was obtained in 83.9% of cases (207 cases), with 18.2% Ppos agreement for judging progression (5 cases), and 91.1% Pneg for judging nonprogression (202 cases). Interobserver agreement (κ) was poor at 0.12, but PABAK was 0.68, and AC1 was 0.8. In 39 cases, the third grader was required to grade.

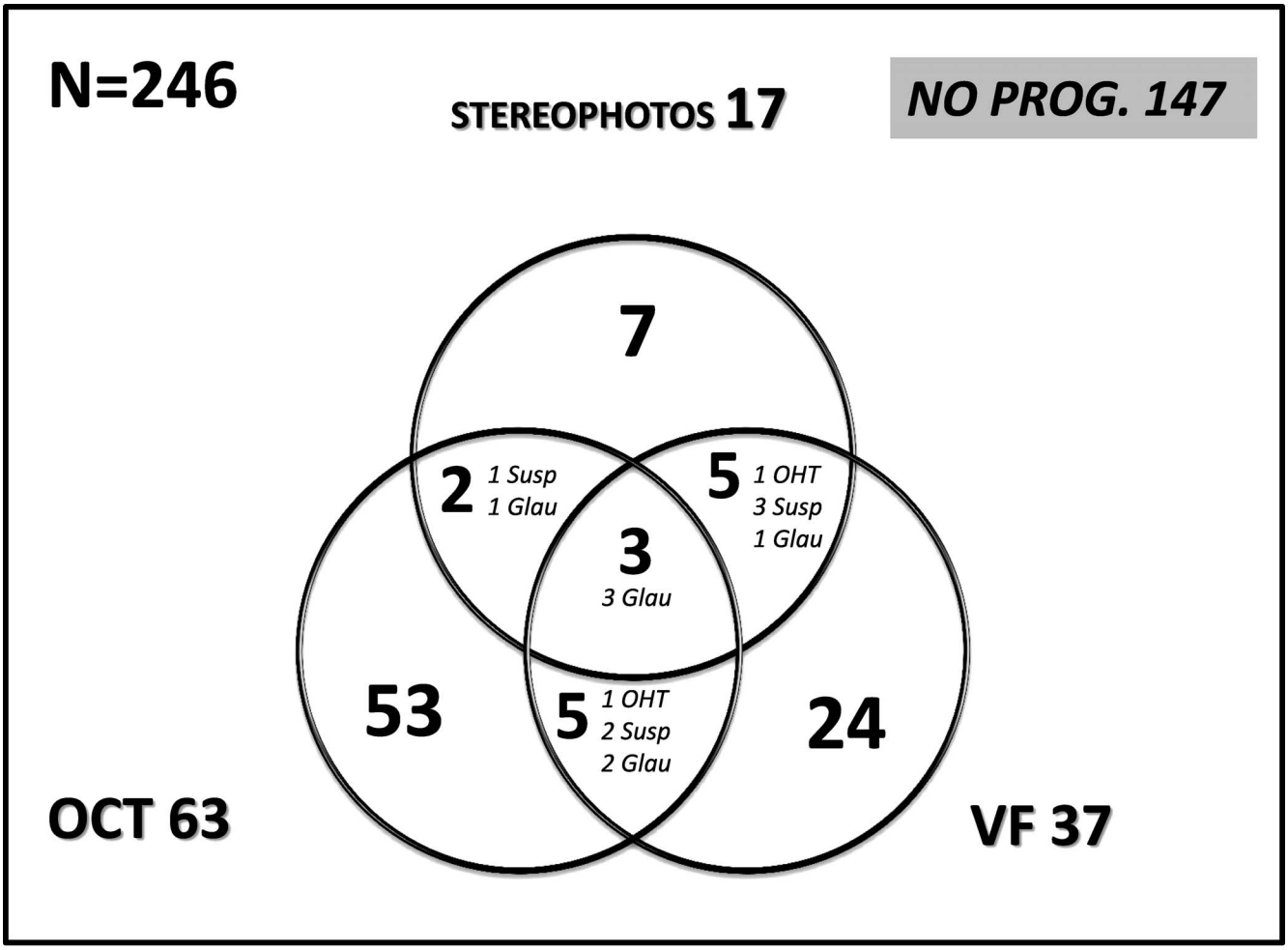

Figure 2. Venn diagram shows cases with progression identified by different tests.

Table 3. Frequency Table Showing Different Test and Algorithms Rates of Progression Detection and Specificity Values

Methods/Algorithms	Progressing Cases	Glaucoma	ONH Suspected	OHT	Normal	Specificity, %
No. of stereo photos (no. of subjects)	17 (16)	8	7	2	0	100
VF progression (no. of subjects)	37 (26)	21	8	5	3	82
VF GPA event	26	18	5	2	1	94
VF GPA trend	20	10	4	4	2	88
SD-OCT progression (no. of subjects)	63 (42)	24	16	21	2	88
Thickness map	52	19	15	17	1	94
Thickness profile	31	12	9	9	1	94
Average thickness	30	14	8	7	1	94
Overall thickness	17	11	2	4	0	100
Superior thickness	11	5	3	3	0	100
Inferior thickness	14	7	5	1	1	94

Specificity of each method detecting progression was calculated by subtracting the number of "progressing cases" of the Normal group from 17, dividing by 17, and multiplying by 100.

Progression by SD-OCT was identified in 63 eyes (25.6%). Thickness map and thickness profile, both event parameters, detected significant change in 52 eyes (21.1%) and in 31 eyes (12.6%), respectively. Average thickness, a parameter that groups trend analysis parameters, detected progression in 31 eyes (12.6%). Overall thickness, superior thickness, and inferior thickness identified progression in 17 (6.9%), 11 (4.4%), and 14 eyes (5.6%), respectively.

Progression by VF was observed in 37 eyes (15%). Visual field event analysis detected change in 26 eyes and VF trend analysis in 20 cases. Table 3 is a frequency table showing different tests and algorithm rates of progression detection and specificity values.

The three techniques obtained good outcomes in terms of specificity (82%–100%). Stereo photos and the OCT trend analysis parameters overall thickness and superior thickness reported the highest specificity results of 100% (Table 3).

Complete agreement detecting progression by the three modalities was found in three eyes (1.2%), whereas 147 eyes (59.7%) were identified as nonprogressing cases by all three methods (Fig. 2). Most cases with detectable changes were identified by only one examination method, resulting in low Ppos; 27.9% and 32.4% for VF event and VF trend analysis, respectively, with stereo photos, and all under 25% for SD-OCT parameters and stereo photos. Nevertheless, concordance in cases with "no progression detected" was high, showing high OPA (72.8%–89.8%) and high Pneg (83.8%–94.5%). The highest OPA, Ppos, and Pneg results were found between stereo photos and VF trend analysis, followed by stereo photos and overall thickness (Table 4). Despite the high OPA values (all greater than 70, most greater than 80), κ coefficients among SD-OCT, VF, and stereo photos were poor (less than 0.3). The highest κ values were obtained with photos-VF trend analysis and photos-VF event analysis, 0.27 and 0.21, respectively (Table 4). Measures of agreement showed a trend toward better agreement between photos and VF trend analysis and between stereo photos and overall thickness than by any other pair, obtaining acceptable PABAK and AC1 values (>0.78 and 0.88, respectively). Spectral-domain OCT parameters reflected a tendency toward better agreement with stereo photos than with VF. Additionally, OCT parameter obtained better agreement coefficient outcomes with VF trend analysis than with VF event analysis. Finally, photos and VF trend analyses tended to agree more than photos-VF event analyses in all indices (Table 4).

Discussion

Detection of worsening of glaucomatous damage is not a simple task. There are three main difficulties: first, differentiating real change from instrument variability; second, limited knowledge about the amount of change that is clinically relevant; and third, no gold standard test with which to compare the rest of the tests. In fact, consensus using conventional methods is difficult to achieve, even among glaucoma experts.[6,24] The present study evaluated the agreement between two structural tests (optic disc stereo photos and SD-OCT) and the one functional test standard automated perimetry.

The progression detection rate using stereo photos (6.9%) was similar to those observed by Alencar et al.[5] (3.7%) and Heijl et al.[25] (4%) in the treatment group of the EMGT study. Our results showed that SD-OCT GPA has a higher propensity to classify the presence of progression (25%) than VF GPA (15%) or photographs with the given definitions of change.

In this study, conventional modalities detected fewer progressing cases. Are these OCT positive ratings "false positives"? Or is it because over the span of 3 to 4 years few eyes will meet the requirements for being classified by SAP as possible progression or as abnormal change identified in series of photographs. Are these OCT positive ratings clinically useful? Will standard methods detect these progressing cases in future years? There is a clear necessity to expand the study and include more patients over a longer follow-up period in order to elucidate these questions.

It seems evident, however, that these diagnostic modalities provide data about very different aspects of the glaucomatous disease process. Several studies suggest that test performances for detecting progression could vary throughout the different stages of the disease. Subtle progression of glaucoma in the early stages is probably undetectable in disc photos or VF but could be identified in RNFL, and in contrast, changes in advanced stages are better assessed by VF.[1,17,26–28] All studies indicate the need to combine functional and structural tests when assessing changes in glaucoma[3,29,30] and to thoroughly evaluate any progression algorithm. Our results agree with these statements. In our study, most OCT progressing cases, 37 of 63 cases (58.7%), were initially classified at baseline as OHT (33.3%) or as glaucoma suspect cases (25.4%). In contrast, most VF progressing cases, 21 of 37 cases (56.8%), were initially described as glaucoma cases. SAP detected abnormal changes in only eight glaucoma suspects (21.6%) and five OHT cases (13.5%). When analyzing the Venn diagram, we observed most agreement on progressing cases on glaucomatous eyes[8] and optic nerve head (ONH) suspects,[7] whereas diagnostic test results concurred in classifying progression in only two OHT cases.

In addition, results indicate substantial differences in agreement for detection of change among different SD-OCT

Investigative Ophthalmology & Visual Science

TABLE 4. Glaucoma Progression Measurement Data

	κ			PABAK			Gwet's AC1			OPA			Ppos			Pneg	
Measurement	Photos	VF Event	VF Trend	Photos	VF Event	VF Trend	Photos	VF Event	VF Trend	Photos	VF Event	VF Trend	Photos	VF Event	VF Trend	Photos	VF Event	VF Trend
Thickness map	−0.02	0.01	−0.01	0.48	0.46	0.47	0.66	0.63	0.65	74	72.8	73.6	8.6	15.2	11.0	84.8	83.8	84.5
Thickness profile	−0.01	0.03	−0.02	0.64	0.60	0.62	0.78	0.75	0.77	82.1	80.1	80.9	8.3	14.0	7.8	90.1	88.7	89.3
Average thickness	0.08	0.03	−0.02	0.67	0.60	0.62	0.80	0.75	0.77	83.7	80.1	80.9	16.7	14.0	7.8	91.0	88.7	89.3
Overall thickness	0.18	0.06	0.04	0.79	0.70	0.73	0.88	0.82	0.84	89.4	85.0	86.6	23.5	14.0	10.8	94.3	91.8	92.7
Superior thickness	−0.06	−0.07	−0.06	0.77	0.70	0.75	0.87	0.83	0.86	88.6	85.0	87.4	0.0	0.0	0.0	94.0	91.9	93.3
Inferior thickness	0	0.03	−0.01	0.76	0.71	0.74	0.87	0.83	0.85	88.2	85.4	87.0	6.5	10.0	5.9	93.7	92.0	93.0
VF event analysis	0.21		0.33	0.75		0.77	0.85		0.86	87.4		88.6	27.9		39.1	93.1		93.7
VF trend analysis	0.27	0.33		0.80	0.77		0.88	0.86		89.8	88.6		32.4	39.1		94.5	93.7	

parameters with standard methods. Higher levels of agreement among tests were obtained when analyzing with OCT trend parameters than with OCT event parameters. This could indicate more similarities in detection capabilities and temporal appearance of change between OCT trend analysis and conventional methods.

There is poor agreement among the tests detecting progression described in many previous studies (Moreno-Montañes J, *IOVS* 2010: ARVO E-Abstract 4015/A294).[3–5,13–15,17,26,27] Xin et al.[4] compared functional and structural changes in glaucoma assessed with Humphrey VF, frequency doubling perimetry, multifocal visual evoked potentials, Stratus OCT (Carl Zeiss Meditec) and stereo photos. Although progression percentages were similar with all methods (from 16.4%–23.6%) and similar to the ones obtained in the present study (15%–25%), no eye showed progression in all tests at the same time. Eleven eyes showed progression by OCT, but only two of these cases presented VF change, and only four showed evidence of changes in stereo photos. These results were consistent with those in the study by Leung et al.,[1] who found agreement in only 3 of 40 progressing eyes, both by Stratus GPA trend analysis and by VF trend analysis. Similarly, Strouthidis et al.[3] found poor agreement between progression assessed by HRT and VF in 198 subjects with ocular hypertension after 7 years' follow-up. Alencar et al.[5] reported agreement among stereo photos, VF, and GDX in only 2 of 34 cases. A κ index value 0.48 and Gwet's AC1 value of 0.92 were found between GDX and conventional tests. These were the highest published agreement figures.

The lack of agreement among tests to identify progression in glaucoma may be due to differences in the capability of instruments and algorithms to detect changes and/or to natural temporal differences in the appearance of structural and functional changes. Whether this is a true characteristic of the disease or a consequence of the capabilities of each method to detect progression requires further study.

Our study has limitations. First, the follow-up time could be considered relatively short. However, the SD-OCT GPA was implemented in 2008, so 5 years was the longest available follow-up period. Second, the prevalence of progression with all three methods was relatively low, limiting performance and agreement assessment, but it was very similar to that of other studies and reflects the reality of daily glaucoma practice. Third, there are certain statistically computed trends by SAP and SD-OCT in our study that may not be clinically significant; for example, a slope of <1.5 VFI %/y or <1 dB/y for VF[31,32] or slopes less than 0.94 to 1.18 μm loss per year in average RNFL thickness.[33,34] In the current study, there were only three cases with negative slopes less than 0.5 dB/y (data not presented); if these cases had been considered not progressed, the specificity of the VF GPA trend analysis would have improved to 94% and the agreement between trend analysis and the other methods would also improve slightly. Fourth, a gold standard for comparison with the rest of the modalities is lacking. Most clinical trials in glaucoma have used VF and/or optic disc photos, despite the frequent disagreement in progression detection found between them in previous publications. Despite these limitations we believe the results increase our understanding of progression detection with available tests and algorithms.

To summarize, the three methods obtained acceptable agreements outcomes in terms of PABAK, AC1, and OPA. However, most cases with detectable changes were identified by only one examination method resulting in low Ppos. Measures of agreement showed a trend toward better agreement between photos and VF than by photos and SD-OCT. SD-OCT parameters reflected a tendency toward better agreement with stereo photos than with VF. Finally, higher levels of agreement between tests were obtained when

analyzing with OCT trend parameters compared to OCT event parameters.

Acknowledgments

Disclosure: **S.A. Banegas**, None; **A. Antón**, Alcon (F, C, R), MSD (C, R), Santen (C, R), Transcend (C, R), Thea (C, R), Bausch (C, R), Pfizer (C, R); **A. Morilla-Grasa**, None; **M. Bogado**, None; **E.M. Ayala**, None; **J. Moreno-Montañes**, None

References

1. Leung C, Cheung C, Weinreb R, et al. Evaluation of retinal nerve fiber layer progression in glaucoma: a study on optical coherence tomography guided progression analysis. *Invest Ophthalmol Vis Sci*. 2010;51:217–222.
2. European Glaucoma Society. *Terminología y Pautas Para el Glaucoma*. Segunda ed. Savona, Italy: Editrice Dogma S.r.l.; 2003: 1–30.
3. Strouthidis NG, Scott A, Peter NM, Garway-Heath DF. Optic disc and visual field progression in ocular hypertensive subjects: detection rates, specificity, and agreement. *Invest Ophthalmol Vis Sci*. 2006;47:2904–2910.
4. Xin D, Greenstein V, Ritch R, et al. A comparison of functional and structural measures for identifying progression of glaucoma. *Invest Ophthalmol Vis Sci*. 2011;52:519–526.
5. Alencar LM, Zangwill LM, Weinreb RN, et al. Agreement for detecting glaucoma progression with the GDx guided progression analysis, automated perimetry, and optic disc photography. *Ophthalmology*. 2010;117:462–470.
6. Breusegem C, Fieuws S, Stalmans I, et al. Agreement and accuracy of non-expert ophthalmologists in assessing glaucomatous changes in serial stereo optic disc photographs. *Ophthalmology*. 2011;118:742–746.
7. Maddess T. The influence of sampling errors on test-retest variability in perimetry. *Invest Ophthalmol Vis Sci*. 2011;52:1014–1022.
8. Wall M, Woodward K, Doyle C, et al. Repeatability of automated perimetry: a comparison between standard automated perimetry with stimulus size III and V, matrix, and motion perimetry. *Invest Ophthalmol Vis Sci*. 2009;50:974–979.
9. Vizzeri G, Weinreb RN, Gonzalez-Garcia AO, et al. Agreement between spectral-domain and time-domain OCT for measuring RNFL thickness. *Br J Ophthalmol*. 2009;93:775–781.
10. Leung CK, Cheung CY, Weinreb RN, et al. Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: a variability and diagnostic performance study. *Ophthalmology*. 2009;116:1257–1263.
11. Mwanza J-C, Chang R, Budenz D, et al. Reproducibility of peripapillary retinal nerve fiber layer thickness and optic nerve head parameters measured with Cirrus HD-OCT in glaucomatous eyes. *Invest Ophthalmol Vis Sci*. 2010;51:5724–5730.
12. Medeiros FA, Alencar LM, Zangwill LM, et al. Detection of progressive retinal nerve fiber layer loss in glaucoma using scanning laser polarimetry with variable corneal compensation. *Invest Ophthalmol Vis Sci*. 2009;50:1675–1681.
13. Vizzeri G, Weinreb RN, Martinez de la Casa JM, et al. Clinicians agreement in establishing glaucomatous progression using the Heidelberg retina tomograph. *Ophthalmology*. 2009;116:14–24.
14. Alencar LM, Zangwill LM, Weinreb RN, et al. A comparison of rates of change in neuroretinal rim area and retinal nerve fiber layer thickness in progressive glaucoma. *Invest Ophthalmol Vis Sci*. 2010;51:3531–3539.
15. Saarela V, Falck A, Airaksinen PJ, et al. The sensitivity and specificity of Heidelberg retina tomograph parameters to glaucomatous progression in disc photographs. *Br J Ophthalmol*. 2010;94:68–73.
16. Carl Zeiss Meditec., Inc. User manual: Cirrus HD-OCT 4.0. Jena, Germany: Carl Zeiss Meditec, Inc.; 2009.
17. Wollstein G, Schuman JS, Price LL, et al. Optical coherence tomography longitudinal evaluation of retinal nerve fiber layer thickness in glaucoma. *Arch Ophthalmol*. 2005;123:464–470.
18. Leung CK, Chiu V, Weinreb RN, et al. Evaluation of retinal nerve fiber layer progression in glaucoma: a comparison between spectral-domain and time-domain optical coherence tomography. *Ophthalmology*. 2011;118:1558–1562.
19. Lopez-Peña MJ, Ferreras A, Polo V, et al. Relationship between standard automated perimetry and retinal nerve fiber layer parameters measured with laser polarimetry. *Arch Soc Esp Oftalmol*. 2010;85:22–31.
20. Caprioli J, Nouri-Mahadavi K, Law SK, et al. Optic disc imaging in perimetrically normal eyes of glaucoma patients with unilateral field loss. *Trans Am Ophthalmol Soc*. 2006;104:202–211.
21. Bengtsson B, Heijl A. A visual field index for calculation of glaucoma rate progression. *Am J Ophthalmol*. 2008;145:343–353.
22. Abramson JH. WINPEPI updated: computer programs for epidemiologists, and their teaching potential. *Epidemiol Perspect Innov*. 2011;8:1.
23. King JE. Software solutions for obtaining a kappa-type statistic for use with multiple raters. Paper presented at: Annual Meeting of the Southwest Educational Research Association; February 5–7, 2004; Dallas, TX.
24. Jampel HD, Friedman D, Quienley H, et al. Agreement among glaucoma specialists in assessing progressive disc changes from photographs in open-angle glaucoma patients. *Am J Ophthalmol*. 2009;147:39–44.
25. Heijl A, Leske MC, Bengtsson B, et al. Reduction of intraocular pressure and glaucoma progression: results from the Early Manifest Glaucoma Trial. *Arch Ophthalmol*. 2002;120:1268–1279.
26. Moon BG, Sung KR, Cho JW, et al. Glaucoma progression detection by retinal fiber layer measurement using scanning laser polarimetry: event and trend analysis. *Korean J Ophthalmol*. 2012;26:174–181.
27. Kerrigan-Baumrind L, Quigley H, Pease E, et al. Number of ganglion cells in glaucoma eyes compared with threshold visual field tests in the same persons. *Invest Ophthalmol Vis Sci*. 2000;41:741–748.
28. Suh MH, Park KH, Kim H, et al. Glaucoma progression after first-detected disc hemorrhage by optical coherence tomography. *J Glaucoma*. 2012;21:358–366.
29. Bowd C, Zangwill L, Medeiros FA, et al. Structure-function relationships shown by CSLO, OCT, and SLP. *Invest Ophthalmol Vis Sci*. 2006;47:2889–2895.
30. Hood DC, Anderson SC, Wall M, et al. Structure versus function in glaucoma: an application of a linear model. *Invest Ophthalmol Vis Sci*. 2007;48:3662–3668.
31. Ernest PJ, Schouten JS, Beckers HJ, et al. Prediction of glaucomatous visual field progression using baseline clinical data [published online ahead of print September 26, 2014]. *J Glaucoma*. doi:10.1097/IJG.0000000000000146
32. Nassiri N, Moghimi S, Coleman A, et al. Global and pointwise rates of decay in glaucoma eyes deteriorating according to pointwise event analysis. *Invest Ophthalmol Vis Sci*. 2013;54:1208–1213.
33. Na JH, Sung KR, Lee KS, et al. Detection of glaucomatous progression by spectral-domain optical coherence tomography. *Ophthalmology*. 2013;120:1388–1395.
34. Wessel JM, Horn FK, Tornow RP, et al. Longitudinal analysis of progression in glaucoma using spectral-domain optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2013;54:3613–3620.

Investigative Ophthalmology & Visual Science