European Journal of Ophthalmology / Vol. 11 / Suppl. 2, 2001 / pp. S50-S56

Measuring structural changes in the optic nerve head and retinal nerve fibre layer

A. ANTÓN

Institute of Applied Ophthalmobiology, University of Valladolid, Valladolid, Hospital de la Esperanza and Barcelona Ocular, Universidad Autónoma de Barcelona, Barcelona - Spain

PURPOSE. *Assessing structural changes in the optic nerve is fundamental to the diagnosis and follow-up of glaucoma. Many clinical decisions are based on the determination of stability or non-stability of glaucomatous damage. Retinal nerve fibre layer (RNFL) and optic disc photographs are very useful in clinics but their assessment is based on subjective evaluation.*

METHODS. *Confocal scanning laser ophthalmoscope and confocal scanning laser polarimetry are available and allow objective and quantitative analysis of the optic nerve. These instruments are currently being assessed for their use in glaucoma follow-up.*

CONCLUSIONS. *This article summarises their advantages and limitations in performing the difficult task of detecting progression or changes over time.* Eur J Ophthalmol 2001; 11 (Suppl 2): S50-S56

KEY WORDS. *Glaucoma, Structural changes, Progression, Confocal scanning laser ophthalmoscopy, Confocal scanning laser polarimetry*

INTRODUCTION

Glaucomatous optic neuropathy causes characteristic structural changes in the retinal nerve fibre layer and optic nerve head. The identification and description of these findings and the detection of changes over time is a fundamental component of glaucoma diagnosis and follow-up, and is also useful for evaluating the disease and the effects of treatment. The interpretation of clinical findings from the clinical examination of the optic nerve head and retinal nerve fibre layer (RNFL) may be complicated because of the great variability among normal optic disc sizes and shapes, and the presence of other associated diseases or media opacities. Undoubtedly, the most difficult task is the comparison of findings over time in order to establish stability or progression of the disease.

In clinical practice, the evaluation of structural changes is usually performed based on biomicroscopy findings and optic disc photographs. Unquestionably, meticulous subjective assessment using a slit lamp is very useful in the clinic, but this technique lacks objectivity and quantification. Optic nerve photographs are useful for comparing findings over time. However, they require subjective interpretation or colour-based computer analysis, and are highly influenced by media opacity and photographic measurements (exposure, light source, digital processing) that may influence the appearance of retinal structures.

An optimal method of optic nerve assessment should be objective, quantitative, and reproducible. Different imaging devices have been recently developed offering objective, quantitative, three-dimensional evaluation of the optic nerve head and RNFL. Among these are confocal scanning laser ophthalmoscopy and confocal scanning laser polarimetry. Cross-sectional studies have shown that these instruments are sensitive and specific in the diagnosis of glaucoma, although none are clearly superior to results obtained with experienced evaluation of good quality stereoscopic photographs. On the other hand, these instruments

1120-6721/S50-07$03.50/0

are undergoing long-term prospective evaluations with the results obtained to date suggesting that these may be particularly promising for monitoring progression in glaucoma.

The instruments

Heidelberg retina tomograph

The Heidelberg Retina Tomograph (Heidelberg, Germany; HRT) uses a confocal scanning diode laser (670 nm) ophthalmoscope to acquire highly reproducible and accurate three-dimensional topographic images of the optic disc and peripapillary retina. A topographic image is taken as a series of optical sections at 32 consecutive focal planes, each consisting of 256 x 256 pixels (65, 536 picture elements) over a 10° to 20° field of view. A mean of three such images is normally used for all analyses. The optic disc margin is delimited by a contour line placed around the inner margin of the peripapillary scleral ring. Most calculations are referred to a "reference plane", which can be located freely but is set automatically at 50 mm posterior to the mean peripapillary retinal height along the contour line at the temporal sector between 350° and 356°.

The HRT calculates several optic nerve measurements (e.g. disc area, rim area, rim volume) that describe the neuro-retinal rim and the cup. Most optic disc topography measurements are significantly correlated with optic nerve fibre number in experimental glaucoma (1). The comparison of HRT data with a normative database facilitates the classification of the optic nerve head as normal or glaucomatous. Nevertheless, the great inter-individual variability in size (mean ± SD: 2.6 ± 0.7; range: 0.8 to 5.4 mm^2) (2) and shape of the optic nerve head limits the ability of standard HRT measurements to separate normal eyes from damaged eyes. It is the application of a discriminant function (3) that allows for simultaneous evaluation of several measurements and classifies normal and glaucomatous optic nerves with a sensitivity of 87% and a specificity of 84%.

Usually, glaucomatous structural changes initially affect specific areas of the RNFL and the optic nerve head. Even if damage is diffuse, it is not necessarily homogeneous. For this reason, several research groups have proposed the division of the optic nerve head into 36 10° sectors in order to obtain a more detailed evaluation (4). Sector analysis of the optic nerve data and the use of ratio or normalized measurements (transformed into relative values ranging from 0 to 1) have opened up new ways to assess structural damage in glaucoma. Data from each of the 36 sectors can be compared with normal values, structural changes can be detected and precisely located, and the topographical location of optic nerve damage and functional damage can be correlated (5, 6). Individual sector changes over time can also be evaluated.

Comparing serial examinations of the same eye may help to determine whether the appearance of the optic nerve head is stable or whether significant changes have occurred during a given time interval. Moreover, a comparison may allow the calculation of the rate of progression. The detection of progression is probably the most ambitious application of any of these instruments. The HRT already offers several methods for evaluating changes in optic disc topography, although these applications are changing constantly, and only long-term studies will determine their usefulness and precision (Fig. 1).

- First, the contour line from the baseline image can be easily exported and placed automatically by HRT software, enabling comparison of standard measurements, ratio measurements, rank segment deviation curves, and the location and depth of structural defects of two different topographic images of the same eye.
- Second, the HRT can calculate the topographic difference image by normalising and subtracting the follow-up image from the initial (baseline) image. The height change is considered significant if it is greater than twice the local combined (baseline image plus follow-up image) standard deviation.
- Third, Chauhan et al (7–9) have developed a progression analysis based on condensing the 256 x 256 standard pixel map into a 64 x 64 array of "superpixels", and computing a confidence interval map. The difference between the topography values in each baseline condensed pixel and the corresponding one in the follow-up image allows the computation of a p value for significant differences. The p values are then represented in a colour-coded map by the HRT software. This method is independent of contour line or reference plane. The latest soft-

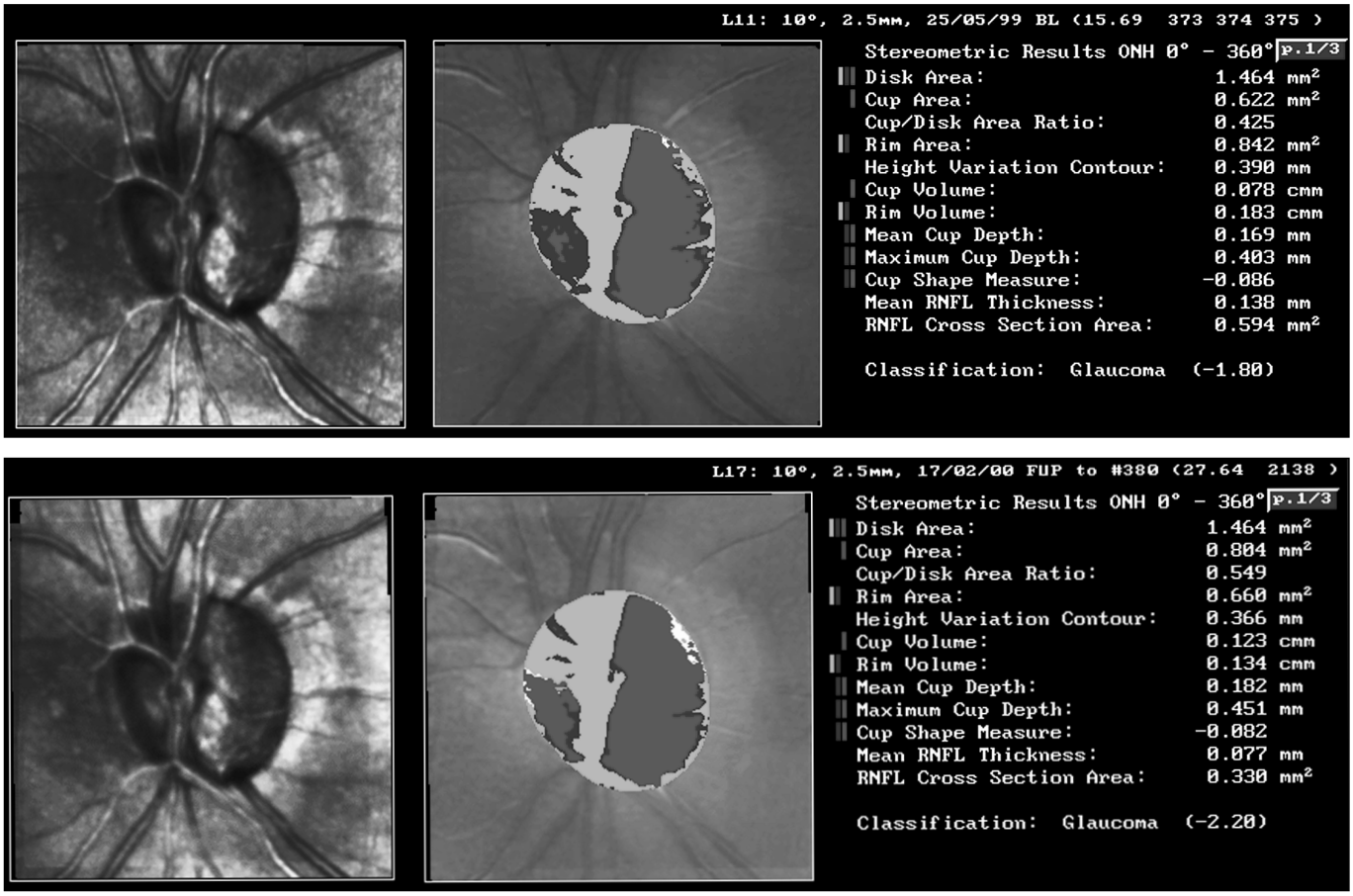

Fig. 1 - *Scanning laser ophthalmoscope images of a left eye that demonstrated progression over a 9-month period. The colour-coded graph shows an increase in the red area representing the area below the reference plane or the cup. Rim area decreased from 0.84 to 0.66 mm² while cup area increased from 0.62 to 0.80 mm².*

ware version implemented in the HRT II applies the most recent findings from Chauhan's group and confirms structural changes evaluating three different follow up images of the same eye.

Despite the properties and advantages, HRT has certain limitations, including the need to calculate most measurements in relation to a reference plane and contour line, and the need for long-term evaluation that is currently under progress.

Nerve fiber analyzer

The Nerve Fiber Analyzer (Laser Diagnostic Technologies, San Diego, CA; NFA) is a scanning laser polarimeter (10) that uses a near-infrared diode laser (wavelength, 780 nm), which is polarised, to illuminate the retina. As light crosses the birefringent nerve fibre layer, a change in the polarisation state of the light occurs. This shift is called "retardation", and is computer-stored and attributed to the parallel orientation of microtubules in the retinal axons. Retardation is measured at 65,536 individual retinal locations (256 x 256 pixels), acquisition time is 0.7 seconds and a compensating device is used to neutralise the birefringence of the anterior segment structures. Retardation has been shown to be linearly related to the thickness of the retinal nerve fibre layer (11).

The instrument offers a screening program and a full test. The full test requires the generation of a measurement ellipse and allows combining images to calculate a mean image. The implemented software calculates multiple global and sector measurements. The latest version, NFA-GDx, offers 14 measurements and a normative database. If any measurement is outside the 80% normal limit, the percentile appears in the print out. The recently developed "modulation para-

meters" provide a comparison between the thickest and thinnest parts of the retinal nerve fibre layer, and seem to be among the best individual measurements for classifying glaucoma and normal eyes (12). Another measurement incorporated in the GDx is the "number", an experimental neural network algorithm that assigns a number from 0 (normal) to 100 (glaucoma) to each eye after analysing all the measurements. Weinreb and co-workers (12) evaluated the diagnostic precision of NFA-GDx and found an area under the receiver operating characteristic (ROC) curve of 0.78, a sensitivity of 82%, and a specificity of 62% for the "number"; for the discriminant function an area under the ROC curve of 0.89, a sensitivity of 74% and a specificity of 92%.

This instrument allows the operator to transfer the ellipse from a patient's visit to the following visit in order to compare all measurements and help in deciding whether there is a change in the optic nerve. The NFA-GDx offers a proprietary follow-up algorithm based on pixel by pixel comparison between two images of the same eye, and shows significant differences in a colour coded map. To the best of my knowledge no report has been published on the evaluation of this algorithm.

The NFA also has certain drawbacks and limitations. Firstly, it gives high retardation measurements in particular areas of peripapillar atrophy or chorio-retinal scars. Secondly, the anterior segment birefringence compensator does not account for all patterns of corneal birefringence and this induces significant measurement errors. As improved version of the biorefringence compensator is currently being developed. Finally, long-term evaluation is needed in order to establish whether this instrument is useful for detecting progression of glaucomatous optic neuropathy.

Variability vs change

The key to detecting change is differentiating real change from variability of the measurements. The latter is related directly to the reproducibility of the instrument and its ability to identify changes. The better the reproducibility, the smaller the change that may be detected by a certain device. Intra-observer variability is the difference between measurements obtained by the same instrument in the same eye by the same observer. Inter-observer variability is the difference between measurements obtained by the same instrument in the same eye by different observers. Variability may be quantified by calculating the mean standard deviation, mean standard deviation equivalent, coefficient of variance, intraclass correlation coefficient or the "limits of agreement" (LA). Standard deviation may be calculated from the relative height or the retardation values obtained at each of the pixels by the scanning lasers. The coefficient of variance (COV) is the ratio between the standard deviation and the mean and it ranges from 0 to 100%, with the lower values indicating better reproducibility. Another useful index is the intraclass correlation coefficient (ICC), defined as the ratio between the variance (V) due to the patient effect and the sum of the variance due to patient and the measurement effect:

$$ICC = V\,(patient) / (V\,(patient) + V\,(error))$$

The values of ICC range from 0 to 1, with 1 representing the best reproducibility. It is generally accepted that reproducibility of measurements is high when the ICC is greater than 0.9 or 90%. Finally, the "limits of agreement", as described by Bland and Altman (13), are based on the difference between two consecutive measurements and the 95% range of this difference. The value is given in the same unit as the measurement assessed and provides the amount of change of a certain measurement needed for it to be statistically significant. All or some of these values have been calculated for each of the instruments in various studies.

Variability of HRT

The variability of HRT raw measurements may be analysed without the need to place the contour line. The relative height of each pixel or small groups of pixels vary in different measurements, and mean standard deviations (MSD) range from 25 to 49 μm (7, 14–16) (with higher MSD representing greater variability). Variability is greater in glaucomatous nerves than in optic nerves (Tab. I).

Reproducibility can also be assessed after the contour line is placed and certain measurements have been calculated. The COV varies with each of the measurements and in different studies. The best values of intra-observer COV are 2%, 4%, and 8% for rim

area, rim volume, and third moment, respectively. As expected, the variability among measurements obtained by different observers is slightly greater than the intra-observer variability (17-19) (Tab. II). Considering that there is great normal variability among normal discs, it may be better to calculate the ICC.

Furthermore, there are different factors that increase the variability of HRT measurements. According to Chauhan (7) and co-workers, variability increases with age, and most studies find greater variability among glaucomatous nerves than among normal nerves. On the other hand, measurements of relative height are less reproducible in areas with greater slope or along the vessels.

Variability of NFA

The reproducibility of the NFA may be evaluated within a pixel or within measurements, and it has improved as subsequent versions of the instrument have become available (NFA I, NFA II, and NFA-GDx). The MSD (and its 95% confidence interval) of average retardation within a 10 pixel width band (9 baseline images) is 0.43° (0.36-0.51°) with a mean COV of 4.2% (3.8-4.5%) according to Zangwill and co-workers (20). Similarly, Hoh (21) studied the total RNFL thickness measurements and observed that mean COV was 4.48 ± 1.76% and 4.92 ± 2.32% for two different operators. Inter-operator reproducibility was high (p = 0.20-0.93)

TABLE I - VARIABILITY OF HRT MEASUREMENTS OF NORMAL AND GLAUCOMATOUS OPTIC NERVES

	Method	Author	Normal	Glaucoma
MSD	Pixel by pixel	Dreher (14)	38–42	41–49
MSD equivalent	64 x 64 pixels	Chauhan (7)	25	31
MSD equivalent	10 x 10 pixels	Cioffi (15)	25	–
MSD	Pixel by pixel	Lusky (16)	30	31

TABLE II - INTRA- AND INTER-OBSERVER VARIABILITY OF DIFFERENT HRT VALUES (17–19)

Measurement	Intra-observer COV COV	Inter-observer COV	Intra-observer ICC ICC	Inter-observer ICC
Disk area	2.4–3.8	4	0.98	0.67
Third moment	8		–0.89	0.92
Rim volume	4–8.5	9	0.99	0.73

TABLE III - REPRODUCIBILITY OF SOME NFA-GDx MEASUREMENTS. INTRACLASS CORRELATION COEFFICIENT (ICC) AND LIMITS OF AGREEMENT (LA) FOR THE MOST REPRODUCIBLE MEASUREMENTS

Measurement	LA (unit)	LA (%)	ICC
Average thickness	5.1 µm	9.3	0.97
Ellipse average	5.2 µm	9.5	0.97
Superior integral	0.015	10.0	0.98
Superior average	5.9 µm	10.4	0.98

Adapted from Colen et al., 2000 (22)

if a single ellipse was applied to the images of both examiners. However, measurement reproducibility worsened if each operator created his own measurement ellipse ($p \leq 0.05$ for 3 out of 5 participants) (21).

Colen (22) evaluated the ICC and the LA for all 14 measurements calculated by the NFA-GDx in a recent study and observed that reproducibility varied considerably across measurements (Tab. III), but did not find consistently better or worse reproducibility in glaucoma patients than in normal individuals. In glaucoma patients, the ICC was lowest for the inferior ratio (0.91) and the maximum modulation (0.88), and highest for the superior average retardation and superior integral (0.98 in both). The LA ranged from 9.3% (5.1 µm) of the average thickness and the 9.5% (5.2 µm) of the ellipse average to the 29% (17.7) and 30% (0.31) of the "number" and the maximum modulation, respectively. The ICC was over 0.90 in 13 of 14 measurements, indicating that the reproducibility of the NFA-GDx in general is high.

Similarities between the HRT and NFA

Several features are common to both the HRT and the NFA.

- Measurement variability is present, as expected, in both devices and needs to be considered when evaluating potential structural changes or progression.
- Both instruments offer good reproducibility and may be particularly useful for evaluating changes in optic nerve structure over time.
- Inter-observer variability is greater that intra-observer variability, particularly when the ellipse (NFA) or the contour line (HRT) are placed independently by each observer.

Although the ellipse and the contour line may be easily transferred from one examination to the next, this final point is an important clinical issue. Most glaucoma patients will be examined over long periods of time (decades) and several operators will probably be involved in their follow-up. The inter-observer variability could probably be decreased with meticulous operation of the instrument, thorough training of the operators, use of disk photographs to help the operators placing the contour line or the ellipse (standard procedure in most research groups), and/or improvement of the software implemented in the instrument.

CONCLUSIONS

1. Instruments are available to provide objective and quantitative data of the optic nerve, including the confocal scanning laser ophthalmoscopy and confocal scanning laser polarimetry.
2. Intra-individual variability is an issue that has to be considered when any technique is used to diagnose glaucoma, and monitor its progression, in clinical practice.
3. Present data regarding HRT and NFA reproducibility suggests that both techniques could be useful in detecting change or progression of glaucomatous optic neuropathy.
4. Algorithms to detect progression or change are under continuous improvement and are currently being prospectively evaluated. These studies will establish the usefulness and limitations of these instruments.

Reprint requests to:
Alfonso Antón, MD
Instituto de Oftalmobiología Aplicada (IOBA)
University of Valladolid, Facultad de Medicina
C/ Ramón y Cajal 7
47005 Valladolid, Spain
aanton@retemail.es

REFERENCES

1. Yucel YH, Gupta N, Kalichman MW, et al. Relationship of optic disc topography to optic nerve fiber number in glaucoma. Arch Ophthalmol 1998; 116: 493-7.
2. Jonas JB, Zack FM, Gusek GC, Naumann GO. Pseudoglaucomatous physiologic large cups. Am J Ophthalmol 1989; 107: 137-44.
3. Mikelberg FS, Parfitt CM, Swindale NV, Graham SL, Drance SM, Gosine R. Ability of the Heidelberg Retina Tomograph to detect early glaucomatous visual field loss. J Glaucoma 1995; 4: 242-7.

4. Asawaphureekorn S, Zangwill L, Weinreb RN. Ranked-segment distribution curve for interpretation of optic nerve topography. J Glaucoma 1996; 5: 79-90.
5. Yamagishi N, Antón A, Sample PA, Zangwill L, Lopez A, Weinreb RN. Mapping structural damage of the optic disc to visual field defect in glaucoma. Am J Ophthalmol 1997; 123: 667-76.
6. Antón A, Yamagishi N, Zangwill L, Sample PA, Weinreb RN. Mapping structural to functional damage in glaucoma with standard automated perimetry and confocal scanning laser ophthalmoscopy. Am J Ophthalmol 1998; 125: 436-46.
7. Chauhan BC, LeBlanc RP, McCormick TA, Rogers JB. Test-retest variability of topographic measurements with confocal scanning laser tomography in patients with glaucoma and normal subjects. Am J Ophthalmol 1994; 118: 9-15.
8. LeBlanc RP, Chauhan BC, Blanchard JW, Hamilton DC. Detecting glaucomatous progression using an analysis of topographic measurements with scanning laser tomography. Invest Ophthalmol Vis Sci 1994; 35 (Suppl): S1730.
9. Blanchard JW, Chauhan BC, Hamilton DC, LeBlanc RP, Field CA. Effect of spatial correlation in detecting changes in the optic nerve head with confocal scanning laser tomography. Invest Ophthalmol Vis Sci 1996; 37 (Suppl): S1091.
10. Weinreb RN, Shakiba S, Zangwill L. Scanning laser polarimetry to measure the nerve fiber layer of normal and glaucomatous eyes. Am J Ophthalmol 1995; 119: 627-36.
11. Weinreb RN, Dreher AW, Coleman A, Quigley H, Shaw B, Reiter K. Histopathologic validation of Fourier-ellipsometry measurements of retinal nerve fiber layer thickness. Arch Ophthalmol 1990; 108: 557-60.
12. Weinreb RN, Zangwill L, Berry CC, Bathija R, Sample PA. Detection of glaucoma with scanning laser polarimetry. Arch Ophthalmol 1998; 116: 1583-9.
13. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1: 307-10.
14. Dreher AW, Tso PC, Weinreb RN. Reproducibility of topographic measurements of the normal and glaucomatous optic nerve head with the laser tomographic scanner. Am J Ophthalmol 1991; 111: 221-9.
15. Cioffi GA, Robin AL, Eastman RD, Perell HF, Sarfarazi FA, Kelman SE. Confocal laser scanning ophthalmoscope. Reproducibility of optic nerve head topographic measurements with the confocal laser scanning ophthalmoscope. Ophthalmology 1993; 100: 57-62.
16. Lusky M, Bosem ME, Weinreb RN. Reproducibility of optic nerve head topography measurements in eyes with undilated pupils. J Glaucoma 1993; 2: 104.
17. Azuara-Blanco A, Harris A, Cantor LB. Reproducibility of optic disk topographic measurements with the Topcon ImageNet and the Heidelberg Retina Tomograph. Ophthalmologica 1998; 212: 95-8.
18. Garway-Heath DF, Poinoosawmy D, Wollstein G, et al. Inter- and intraobserver variation in the analysis of optic disc images: comparison of the Heidelberg retina tomograph and computer assisted planimetry. Br J Ophthalmol 1999; 83: 664-9.
19. Hatch WV, Flanagan JG, Williams-Lyn DE, Buys YM, Farra T, Trope GE. Interobserver agreement of Heidelberg retina tomograph parameters. J Glaucoma 1999; 8: 232-7.
20. Zangwill L, Berry CA, Garden VS, Weinreb RN. Reproducibility of retardation measurements with the nerve fiber analyzer II. J Glaucoma 1997; 6: 384-9.
21. Hoh ST, Ishikawa H, Greenfield DS, Liebmann JM, Chew SJ, Ritch R. Peripapillary nerve fiber layer thickness measurement reproducibility using scanning laser polarimetry. J Glaucoma 1998; 7: 12-5.
22. Colen TP, Tjon-Fo-sang MJ, Mulder PG, Lemij HG. Reproducibility of measurements with the nerve fiber analyzer. J Glaucoma 2000; 9: 363-70.