CHAPTER 3

Reproducibility of measurements with the Nerve Fiber Analyzer

Colen TP and Lemij HG
Journal of Glaucoma 2000;9(5):363-370

Abstract

Purpose To determine the reproducibility of measurements with the Nerve Fiber Analyzer (NFA), a scanning laser polarimeter designed for quantifying glaucoma, in both normal and glaucomatous subjects. We also assessed the variance of measurements between instruments.

Methods Measurements were made with the third generation NFA, the GDx. The study consisted of three parts. In the first part, we measured right eyes of 10 healthy volunteers on 5 consecutive days. In the second part, 45 glaucoma patients underwent NFA measurements of one randomly selected eye on two separate days within a 5-week period. For all 14 available parameters, reproducibility of measurements was expressed in terms of 95% limits of agreement, and as the intraclass correlation coefficient. The NFA software has an option of creating a mean image from a selection of single images; for both parts of the study, the reproducibility of measurements was calculated for a 'single image', and a 'mean of 3' image. In the third part of the study, 17 volunteers underwent repeated NFA measurement sessions on each of 3 different instruments. Using MANOVA, we determined the variance of measurements between instruments.

Results The reproducibility of measurements varied considerably across parameters. Limits of agreement in mean images for superior maximum and inferior maximum were 7.2μ and 7.7μ, respectively in our healthy subjects, and 8.7μ and 7.9μ, respectively in our glaucoma patients. For normal subjects, the intraclass correlation coefficient was >90% in 10 out of 14 parameters. In glaucomatous subjects the intraclass correlation coefficient was >90% in 13 out of 14 parameters. Some parameters reproduced better in a mean than in a single image; these differences, however, were small, and in general not statistically significant. The between instruments component also varied across parameters, being highest in ratio-based parameters.

Conclusions 1. The reproducibility of measurements varied across parameters. 2. In general, the reproducibility of measurements with the NFA was high. 3. The reproducibility of measurements was similar between normal and glaucomatous subjects. 4. Any measured change in nerve fiber layer thickness would be statistically significant if it exceeded about 7-8μ in the superior maximum or inferior maximum parameter in normal subjects. 5. Reproducibility of measurements hardly differed between single images and mean images. 6. The reproducibility of measurements between the 3 instruments we used was highest for straight parameters.

Introduction

The Nerve Fiber Analyzer (NFA; Laser Diagnostic Technologies, San Diego, CA) is a scanning laser polarimeter, designed for the detection and follow-up of glaucoma. It uses a 780-nm diode laser to scan the retina, and to assess the retinal nerve fiber layer (NFL) thickness in the peripapillary region, by measuring changes in polarization (i.e., retardation) of the scanning laser beam. Results from previous studies with the NFA[1,2] have shown a high sensitivity and specificity in discriminating between a normal and glaucomatous NFL.

Knowing the reproducibility of measurements of a method provides an insight into its precision, which is crucial to a meaningful assessment of changes over time. Many studies into the reproducibility of measurements of the NFA have been published. Some results were obtained with the NFA I[3-7] and later studies were carried out with the NFA II.[8-10] No results obtained with the latest NFA version (the GDx), have yet been reported. The GDx has identical hardware as the NFA II, but yields a different set of parameters (only 6 of the 14 parameters available on the GDx were available on the NFA II).

In some studies, the 'within-pixel' reproducibility was calculated, rather than the reproducibility within a parameter. In the within-pixel procedure, several retardation images were aligned, and for every pixel, a standard deviation was calculated. The standard deviations of all 65536 pixels (or only of those outside the peripapillary circle) were averaged. This approach gives the clinician an overall impression of the reproducibility of the measurements in a particular patient. In addition, within pixel reproducibility of measurements can be used for determining statistically significant, localized, change over time. To that end, the software can digitally subtract consecutive images, highlighting pixels where a statistically significant change has occurred, all based on the location specific within pixel reproducibility of a given eye. It is yet unclear how sensitive, specific, and meaningful this approach is.

The reported overall within-pixel reproducibility in the literature usually looks extremely good, being around 3-5μ. Although it demonstrates the high precision of the instrument, it probably has little intuitive clinical meaning. The GDx provides 14 parameters that may be used clinically for the diagnosis and follow-up of patients. Clinicians might therefore be more interested in how well these parameters reproduce.

The intraclass correlation coefficient (ICC) and the coefficient of variation (CV) are common ways to express the reproducibility of measurements per parameter. We think, however, that these measures are not very intuitive to the clinician. We therefore looked for a method that expresses the reproducibility of measurements per parameter, that is intuitive and has direct clinical meaning. The 'limits of agreement', as described by Bland and Altman in the Lancet in 1986[11] is such a method. It is based on the difference between two consecutive measurements, and the 95% range of this difference. It has the same unit of measure as the parameter itself and provides the amount of change of a parameter that is needed for it to be statistically significant. Limits of agreement may be useful to the clinician to assess the significance of any measured change over time.

This study consisted of three parts. In the first part, we determined the reproducibility of measurements in healthy subjects. In the second part, it was assessed in glaucoma patients. These two parts addressed the variance of measurements within a single instrument. In the third part of this study, we assessed this variance across instruments.

Methods

The Nerve Fiber Analyzer A detailed description of the NFA has been published elsewhere.[3,12,13] In short, a 780 nm diode laser scans the peripapillary retina with a scanning angle of 15x15 degs. The laserlight is polarized and as it passes through the re-

tinal nerve fiber layer (NFL), a phase shift called retardation occurs, attributed to the parallel orientation of the microtubules inside the retinal axons. This phase shifting property of the NFL is called birefringence. In the backscattered light, the amount of retardation is thought to linearly reflect the retinal NFL thickness, as has been shown in a monkey model.[14] The retinal NFL thickness is then displayed in a color-coded 256x256 pixels image.

In this image, the operator will typically position a circle or ellipse on the margin of the optic nerve head. A second one with 1.75 times the diameter of the first is displayed automatically, allowing the analysis of only those pixels that sit peripherally to it. Areas of blood vessels are a source of noise[6] and are therefore automatically excluded for analysis by the software. The entire image is divided into 4 segments centered on the optic disc: superior 120°, inferior 120°, nasal 50° and temporal 70°. For a quantitative approach, several parameters are available.

The GDx is a third generation NFA, with a built-in normative database. For all 14 parameters that are available to the user, the percentile in the normal range is computed. When it is below 10% or over 90%, the percentile is displayed on the printout.

Measurement procedures In our study, a measurement session (for brevity a 'measurement') for one eye always consisted of obtaining a minimum of 3 single images of high quality each (i.e. well focussed, the optic disc well centered in the image, equal and total illumination in all segments and no detectable eye movement during image acquisition). Typically, this required about 6 tries. The software of the instrument allows the operator to select any amount of images that may then be aligned and converted into one 'mean image'. In the first two parts of our study, we assessed the reproducibility of measurements for single images (always the first of our 3 images of high quality) and for a 'mean of 3' image. In the third part of the study, only the 'mean of 3' image was considered. In selecting the 3 images of high quality, the operator was masked to the results of the previous measurements. During all measurements, we saw to it that patients had their heads as upright as possible. All three GDx instruments were located in similar, adjacent rooms with ambient lights left on. The software and hardware configuration of all three instruments was identical. All measurements were carried out by the same operator who also manually positioned the ellipse in all images.

Subjects In the first part of the study, measurements were taken on 5 consecutive days of right eyes of 10 healthy volunteers. Their ages ranged from 22 to 61 years (mean: 36 years). None of them had any ocular history; all had an IOP below 21 mm Hg, normal looking optic discs and normal visual fields (GHT within normal limits) on the Humphrey Field Analyzer (24-2 full threshold program). The refractive error of all subjects was between +7 and –7 D.

In the second part, we measured one eye of 45 glaucoma patients on two separate days within 5 weeks. The age of the patients was on average 63 years (range from 39 to 77 years); the mean defect (MD) of the patients was on average –10.7dB (stan-

dard deviation = 7.04). If a patient had glaucoma in one eye only, that eye was considered. If a patient had glaucoma in both eyes, one eye was selected randomly. Patients were recruited from the Glaucoma Service of our hospital, and all had glaucomatous visual field defects with glaucomatous appearing optic discs. Patients with any coexisting ocular disease, including pseudophakia, or systemic diseases with possible ocular involvement, such as diabetes, were excluded from the study.

In the third part of the study, 17 volunteers (4 glaucoma patients and 13 healthy subjects without any ocular history, aged between 27 and 62 years) had their right eye measured twice on each of 3 different instruments. Again, a measurement consisted of a mean of 3 single images of high quality.
Informed consent was obtained from all participating subjects after the nature of the procedure had been fully explained

Statistical Analysis For all statistical tests, the level of statistical significance was set at α=0.05. There was no suspicion of non-normality for the within-patient variability of all parameters in both the normal subjects and glaucoma patients.

In the first two parts of the study, the reproducibility of measurements was expressed as 95% limits of agreement (LA) and in terms of the intraclass correlation coefficient (ICC). Calculations were performed for all 14 parameters available to the user. The definitions of all 14 parameters have been presented in the appendix. Some parameters (what we call straight parameters) relate to direct measurements, whereas others are obtained more indirectly by division of direct measures (these are so called ratio-parameters marked by (r) in Tables 1-3) or by a neural network analysis (The Number). LA relate to the agreement of measurements on two consecutive days within the same subject, and are defined, in this article, as 1.96 times the square root of two times the within-subject variance.[11] LA are also called 'reproducibility' and represent a critical value that will not be exceeded with 95% probability by the absolute difference of two single measurements within the same subject on two consecutive days under the same conditions, assuming a normally distributed within-subject variability with the same variance across all subjects. In the methods described by Bland and Altman, the LA are calculated from two repeated measurements. If one has more than two repeated measurements, the within-subject variance can also be estimated, with greater precision than with only two repeated measurements. This was done in our normal group, where we used a series of 5 repeated measurements per subject.

In addition to the traditional LA, we also expressed the LA of a parameter as a percentage of the mean of the between-subjects distribution of that variable (LA%). This ad-hoc expression equals 1.96 * $\sqrt{2}$ times the coefficient of variation, and is only presented here in order to facilitate the mutual comparison of the LA of different variables. In addition, means and total standard deviations have been presented per parameter (Tables 3-1 and 3-2).

Our second measure to express the reproducibility of measurements was the ICC; we estimated the within-subject and between-subject variances, again assuming that the within-subject variance was the same across the volunteers. The ICC was

obtained by expressing the between-subject variance as a percentage of the total (within- plus between subject) variance. The ICC represents the power by which subjects can be distinguished from each other by their measurement outcome. It is generally accepted that reproducibility of measurements is high when the ICC is >90%.

In the normal subjects, differences in within-subject standard deviation between measurements based on single images and mean images were compared within the same subject with a Wilcoxon matched pairs signed ranks test after logarithmic (ln) transformation of the standard deviation. In the glaucoma group, a sign test was used because of too many zero within-subject standard deviations. We have presented p-values related to the statistical significance of differences in LA between single images and mean images (Table 3-1 and 3-2).

For the third part of the study, a random effects model with subject and instrument as random factors, was used to estimate three variance components: between-subject variance, between instrument variance, and error (or within-instrument) variance. Because of imbalance in the data of this experiment a maximum likelihood method was used ('PROC MIXED' in the SAS statistical package).

Results

Reproducibility of measurements in normal subjects For all available parameters, the mean value across subjects, the standard deviation (SD), the 95% Limits of Agreement (LA), the 95% Limits of Agreement as a percentage of the mean value (LA%), and the Intraclass correlation coefficients (ICC) for normal subjects have been summarized in Table 3-1. The reproducibility of measurements varied considerably between parameters. In general, the straight parameters were more robust than the ratio-parameters. For example, in mean images, superior maximum had LA(%) of 8.5% of the mean values, whereas the LA(%) was 29.2% for maximum modulation. In mean images in healthy subjects, the parameters with LA of 10% and lower were: superior maximum, inferior maximum, symmetry, average thickness, superior average, inferior average and ellipse average. The superior/nasal parameter reproduced better than the superior ratio (which is a superior/temporal ratio).

In mean images, the LA of the superior maximum and inferior maximum parameter were 7.2μ and 7.7μ, respectively. In some parameters, the LA were lower in mean images than in single images. In others, it was vice versa. However, these differences were small, and only one out of 14 p-values that reflected the statistical significance of this difference, was <0.05 (superior maximum; p=0.007).
The differences in ICC between single images and mean images were small, and probably insignificant. In mean images, the intraclass correlation coefficient (ICC) was over 90% except for superior ratio, inferior ratio, maximum modulation and ellipse modulation.

Reproducibility of measurements in glaucoma patients As for normal subjects, some parameters had higher LA than others (table 3-2). Reproducibility of measurements was not consistently better or worse in glaucoma patients as compared to normals. In single images, LA were lower for 7 parameters in glaucoma patients as com-

pared to healthy subjects. In the remaining 7 parameters, the LA were higher. For mean images, the same was observed.
All parameters, except the number, had lower LA in mean images than in single images. However, in only one of the 14 parameters was this difference statistically significant (superior integral; p=0.021).

The differences in ICC between single images and mean images were also small, and probably insignificant. The ICC for mean images was over 90% in all parameters except for maximum modulation.

Variability between instruments The results of the analysis of variance have been presented in Table 3-3. All components that contribute to the variance of a repeated measurement have been expressed in percentages of the total variance of 100%.

The reproducibility of measurements varied across parameters. The *between patients* component was by far the largest one. In general, the straight parameters showed much less variation than the ratio parameters. Of the straight parameters the between instrument variation was on the same order of magnitude as the *within instrument* variation (typically 2-5% of the total variation). Of the ratio parameters, both the *between instrument* and the *within instrument* (or error) component could be as high as 31.1%.

Discussion

We have shown that the reproducibility of measurements of the GDx varies considerably across parameters. For example, 95% limits of agreement (LA) in normal subjects varied from 7.1% for ellipse average to 29.2% for maximum modulation. We therefore think it is more meaningful to speak of 'the reproducibility of a parameter', rather than of 'the reproducibility of a technology'.

The 3 parameters similar to those that have shown a high sensitivity and specificity for detecting glaucoma with the NFA I,[1] (superior maximum, inferior maximum and symmetry) were among the most robust parameters (LA of each approximately 9% and the ICC well over 90%). These three parameters, however, did not discriminate very well between normal and glaucoma in the study by Weinreb et al.[2] Three parameters that discriminated well in their study, and reproduced well in our study, were ellipse average, inferior average and average thickness.

It is generally accepted that the reproducibility of measurements is high when the ICC is over 90%. This was true for most of the parameters that we have investigated, both in glaucoma patients and normals. We therefore conclude that the reproducibility of measurements with the GDx, in general, is high.

We have also found that the LA are not consistently better or worse in glaucoma patients than in normal subjects. Also, the ICC in normal subjects was over 90% in 10 out of 14 parameters; in glaucoma patients the ICC was over 90% in 13 out of 14 parameters. These conclusions are in agreement with those of Holló *et al.*,[9] who reported coefficients of variation that were similar for glaucoma patients and control subjects with the NFA II.

We have presented LA for all available GDx parameters. The tenet of these LA is that they are a clinically meaningful and intuitive parameter: as soon as a change in a certain parameter exceeds these limits, it reflects a statistically significant change. Any change within these limits is, with 95% probability, a coincidence. All parameters have their own characteristic LA. We believe that LA provide, in principle, clinically useful information to assess the significance of any change measured over time. How meaningful this approach is, has yet to be shown in long-term studies. The Number is unlikely to be useful for detecting change over time, since it reflects the likelihood of glaucoma, and not the severity of the disease.

Our results show that some parameters tend to be slightly more robust in mean images than in single images, especially in glaucoma patients. However, the present differences were small and statistically significant for only one in 14 parameters (both in the group of normals and in the glaucoma group). This one p-value, however, probably has little meaning in a series of multiple comparisons. One might expect that, in case of no differences between 20 given test variables, one in every 20 p-values will be below 0.05. We therefore conclude that LA are generally slightly lower in mean images than in single images, but that these differences do not reach statistical significance. In general, one would expect a mean image to have a better reproducibility than a single image, since one is averaging over more pixels of information, thus reducing the effect of random noise. Fortunately, this difference was not statistically significant. In a previous study with the NFA II we did find a statistical significance between a mean of three and a single image.(Colen et al., IOVS 1998;89(suppl):S-3223) It is unclear why we found no such statistical significance with the GDx.

One might conclude from our study that a single image is just as good as a mean image since the LA of single images does not differ statistically significantly from that of mean images. However, using only one image is not the same as taking only one image. We always take a minimum of 6 images per eye. This makes it easier for us to select those images of highest quality (fewest motion artifacts, best illumination etc.)

Finally, we have demonstrated that reproducibility of measurements across three different instruments is highest for straight parameters. Ratios are less robust, obviously because both the numerator and the denominator will have their variability. Notably if the denominator is relatively small to its variability, the variability of the ratio will be large. Not surprisingly, symmetry was more robust than the other ratio parameters in normal subjects (relatively large denominator), but not as robust in glaucoma patients, where small denominators were to be expected in individual cases.

In conclusion, we have explored the reproducibility of measurements with the Nerve Fiber Analyzer/GDx. Some parameters are promisingly robust to serve in follow-up of patients, and we are currently investigating the significance of the more robust parameters in long-term follow-up studies.

Conclusions

The reproducibility of measurements varied across parameters. In mean images in healthy subjects the parameters with LA of 10% and lower were: superior maximum, inferior maximum, symmetry, average thickness, superior average, inferior average and ellipse average. The ICC indicated that the reproducibility of measurements with the NFA in general is high. The LA in normal and glaucomatous subjects were similar. The tenet of the LA is that, for example, in normals, any measured change in nerve fiber layer thickness is statistically significant when it exceeds about 7-8μ in the superior maximum or inferior maximum parameter. Differences between single and mean images were small and not statistically significant. Finally, reproducibility of measurements across three different instruments was highest for straight parameters.

Acknowledgement: *We thank Laser Vision for courteously lending us a GDx for the third part of our study.*

Appendix

GDx Parameters

Superior Maximum: This is the average of the 1500 thickest pixels in the superior quadrant.

Inferior Maximum: This is the average of the 1500 thickest pixels in the inferior quadrant.

Symmetry: This is the ratio of the average of the 1500 thickest pixels in the superior quadrant over the average of the 1500 thickest pixels in the inferior quadrant.

Superior Ratio: This is the ratio of the average of the 1500 thickest pixels in the superior quadrant over the average of the 1500 median pixels in the temporal quadrant.

Inferior Ratio: This is the ratio of the average of the 1500 thickest pixels in the inferior quadrant over the average of the 1500 median pixels in the temporal quadrant.

Superior/Nasal: This is the ratio of the average of the 1500 thickest pixels in the superior quadrant over the average of the 1500 median pixels in the nasal quadrant.

"The Number": This is an experimental number currently under evaluation. A neural network has been trained to look at all values obtained when an image is acquired. The neural network then assigns a number between 0 and 100 to each patient according to the following scale: 0 = totally normal, 100 = glaucoma. Early evaluation of the number indicates that patients who score between 0-30 are normal; patients scoring over 70 are glaucomatous; and those scoring between 30-70 are glaucoma suspects. Patients in this last category may prove to be "borderline" or "outside normal limits" on some GDx parameters but not exhibit any visual field loss or other indications of glaucoma.

Max Modulation: Provides an indication of the difference between the thickest parts of the nerve fiber layer and the thinnest parts. First, the average is calculated for 1) the 1500 thickest points in the superior quadrant, 2) the 1500 thickest points in the inferior quadrant, 3) the 1500 median points in the nasal quadrant and 4) the 1500 median points in the temporal quadrant. Once an average number is derived for each quadrant, the lowest number is subtracted from the highest number. The resulting number is then divided by the lowest number.

Average Thickness: The average thickness of all pixels in the image; evaluates all 65,536 points used to create an image. The Average is calculated by adding up all of the values of the usable pixels outside of the optic nerve head (as designated by the operator) and dividing by the number of pixels used.

Ellipse Modulation: Like "Max Modulation", Ellipse Modulation is an indication of the difference between the thickest parts of the nerve fiber layer and the thinnest parts. The difference is that, rather than using all of the points in the image, Ellipse Modulation uses the points in the ellipse surrounding the optic nerve. Ellipse Modulation is calculated by taking the thickest pixel along the ellipse, subtracting the thinnest pixel along the ellipse, and dividing the total by the value of the thinnest pixel.

Ellipse Average (measurement is in microns)**:** The average thickness of the nerve fiber layer around the ellipse surrounding the optic nerve.

Superior Average (measurement is in microns)**:** The average thickness of the nerve fiber layer along the portion of the ellipse surrounding the optic nerve in the superior quadrant.

Inferior Average (measurement is in microns)**:** The average thickness of the nerve fiber layer along the portion of the ellipse surrounding the optic nerve in the inferior quadrant.

Superior Integral (measurement is in millimeters squared)**:** The total area under the curve (or total volume) of the nerve fiber layer along the superior portion of the ellipse surrounding the optic nerve.

	SINGLE IMAGE						MEAN OF 3					
	unit	mean	SD	LA	LA (%)	ICC (%)	mean	SD	LA	LA (%)	ICC (%)	p-value
superior maximum	µ	85.0	14.45	9.4	11.0	94.6	84.6	14.7	7.2	8.5	96.9	.007
inferior maximum	µ	90.5	15.45	9.0	10.0	95.6	89.3	15.6	7.7	8.6	96.8	.139
symmetry (r)		0.95	0.12	0.11	11.9	89.4	0.96	0.13	0.09	9.3	94.3	.799
superior ratio (r)		2.29	0.45	0.43	18.8	88.1	2.34	0.50	0.47	20.0	88.8	.333
inferior ratio (r)		2.41	0.33	0.51	21.1	69.1	2.43	0.35	0.50	20.3	73.6	.285
superior/nasal (r)		2.12	0.35	0.26	12.2	93.2	2.13	0.38	0.32	15.1	90.8	.721
the number		18.5	13.2	11.7	63.2	89.8	18.7	13.1	10.4	55.6	91.9	.327
maximum modulation (r)		1.60	0.34	0.42	26.0	80.4	1.64	0.39	0.48	29.2	80.0	.114
average thickness	µ	62.0	7.8	4.7	7.6	95.1	60.8	7.7	4.8	8.0	94.9	.678
ellipse modulation (r)		2.62	0.33	0.55	21.1	64.5	2.75	0.37	0.60	21.7	65.7	.646
ellipse average	µ	65.4	7.7	5.1	7.8	94.3	64.3	7.6	4.6	7.1	95.3	.260
superior average	µ	75.7	11.1	8.1	10.7	93.1	74.2	10.9	7.8	10.5	93.4	.515
inferior average	µ	77.0	11.3	7.1	9.2	94.9	76.0	10.4	6.7	8.9	94.6	.799
superior integral	mm2	0.19	0.035	0.025	12.7	93.4	0.19	0.034	0.026	13.7	92.4	.333

Table 3-1. Reproducibility of measurements in healthy subjects

Reproducibility of all 14 available parameters has been calculated for single images and 'mean of 3' images. Parameters marked with an (r) are ratio-based parameters. The unit of measurement has been given; some parameters are without dimension. The mean value across subjects (and the standard deviation SD) has been given for every parameter. Limits of Agreement (LA) have been presented in the same dimension as the parameter itself. To compare between parameters, LA have also been expressed as a percentage of their mean value (LA%). In addition, the Intraclass Correlation Coefficient (ICC) has been given for every parameter. The final column represents p-values for statistical significance of differences in LA between single images and mean images.

Table 3-2. Reproducibility of measurements in glaucoma patients

	SINGLE IMAGE					MEAN OF 3					
	unit	mean	SD	LA	LA (%)	ICC (%)	mean	SD	LA	LA (%)	ICC (%)	p-value
superior maximum	μ	66.1	16.9	14.4	21.8	90.2	65.7	17.3	8.7	13.3	96.7	.449
inferior maximum	μ	73.9	17.6	9.3	12.6	96.3	72.8	18.9	7.9	10.9	97.7	.720
symmetry (r)		0.91	0.17	0.15	16.4	89.4	0.92	0.17	0.12	12.6	94.1	.432
superior ratio (r)		1.61	0.40	0.31	19.0	92.2	1.61	0.41	0.28	17.6	93.6	.613
inferior ratio (r)		1.77	0.38	0.35	19.6	88.6	1.76	0.36	0.28	15.8	91.9	.442
superior/nasal (r)		1.52	0.28	0.17	10.9	95.5	1.51	0.29	0.17	11.1	95.7	.142
the number		55.2	24.6	16.9	30.6	93.8	59.4	24.4	17.7	29.8	93.0	.868
maximum modulation (r)		1.0	0.34	0.33	33.4	86.5	0.99	0.33	0.31	30.9	88.4	.224
average thickness	μ	55.3	10.7	6.5	11.7	95.2	54.3	12.0	5.1	9.3	97.7	.064
ellipse modulation (r)		1.63	0.55	0.49	29.8	89.4	1.68	0.54	0.33	19.4	95.3	.165
ellipse average	μ	55.4	11.0	6.2	11.1	95.9	54.5	12.5	5.2	9.5	97.7	.424
superior average	μ	57.2	13.4	7.1	12.3	96.4	56.6	14.9	5.9	10.4	98.0	1.00
inferior average	μ	63.2	15.5	8.0	12.7	96.5	62.2	16.8	7.1	11.3	97.7	.248
superior integral	mm^2	0.15	0.034	0.019	12.7	95.8	0.15	0.039	0.015	10.0	98.0	.021

Legend: as in 3-1.

Table 3-3. Variability between instruments

	between patients (%)	between instruments (%)	within instruments (%)	total (%)
superior maximum	92.9	2.9	4.2	100
inferior maximum	94.2	2.0	3.8	100
symmetry (r)	83.2	5.3	11.5	100
superior ratio (r)	60.4	12.1	27.5	100
inferior ratio (r)	46.2	22.7	31.1	100
superior/nasal (r)	65.3	20.8	13.9	100
the number	84.7	3.7	11.6	100
maximum modulation (r)	68.7	16.1	15.2	100
average thickness	91.9	5.5	2.6	100
ellipse modulation (r)	52.1	29.0	18.9	100
ellipse average	93.1	4.0	2.9	100
superior average	93.1	4.5	2.4	100
inferior average	95.8	0.7	3.5	100
superior integral	87.0	8.1	4.9	100

Analysis of Variance showing the three components that make up the total amount of variance.

References

1. Tjon-Fo-Sang MJ, Lemij HG. The sensitivity and specificity of nerve fiber layer measurements in glaucoma as determined with scanning laser polarimetry. Am J Ophthalmol 1997;123:62-9.
2. Weinreb RN, Zangwill L, Berry CC, et al. Detection of glaucoma with scanning laser polarimetry. Arch Ophthalmol 1998;116:1583-9.
3. 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.
4. Chi Q, Tomita G, Inazumi K, et al. Evaluation of the effect of aging on the retinal nerve fiber layer thickness using scanning laser polarimetry. J Glaucoma 1995;4:406-13.
5. Junghardt A, Schmid MK, Schipper I, et al. Reproducibility of the data determined by scanning laser polarimetry. Graefes Arch Clin Exp Ophthalmol 1996;234:628-32.
6. Tjon-Fo-Sang MJ, van Strik R, de Vries J, Lemij HG. Improved reproducibility of measurements with the nerve fiber analyzer. J Glaucoma 1997;6:203-11.
7. Waldock A, Potts MJ, Sparrow JM, Karwatowski WS. Clinical evaluation of scanning laser polarimetry: I. Intraoperator reproducibility and design of a blood vessel removal algorithm. Br J Ophthalmol 1998;82:252-9.
8. Zangwill L, Berry CA, Garden VS, Weinreb RN. Reproducibility of retardation measurements with the nerve fiber analyzer II. Journal of Glaucoma 1997;6:384-9.
9. Holló G, Suveges I, Nagymihaly A, Vargha P. Scanning laser polarimetry of the retinal nerve fibre layer in primary open angle and capsular glaucoma. Br J Ophthalmol 1997;81:857-61.
10. Hoh ST, Ishikawa H, Greenfield DS, et al. Peripapillary nerve fiber layer thickness measurement reproducibility using scanning laser polarimetry. J Glaucoma 1998;7:12-5.
11. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;307-10.
12. Dreher AW, Reiter K. Retinal laser ellipsometry: a new method for measuring the retinal nerve fiber layer thickness distribution. Clin Vis Sci 1992;7:481-8.
13. Tjon-Fo-Sang MJ, de Vries J, Lemij HG. Measurement by nerve fiber analyzer of retinal nerve fiber layer thickness in normal subjects and patients with ocular hypertension. Am J Ophthalmol 1996;122:220-7.
14. Weinreb RN, Dreher AW, Coleman A, et al. Histopathologic validation of Fourier-ellipsometry measurements of retinal nerve fiber layer thickness. Arch Ophthalmol 1990;108:557-60.