

Operator Learning Effect and Interoperator Reproducibility of the Scanning Laser Polarimeter with Variable Corneal Compensation

Shahar Frenkel, MD, PhD, Efrat Slonim, DVM, Amjad Horani, MD, Michal Molcho, PhD, Israel Barzel, Eytan Z. Blumenthal, MD

Purpose: To ascertain operator learning effect, and to quantify the interoperator reproducibility, using the newly introduced GDx VCC (variable corneal compensator) scanning laser polarimeter.

Design: Prospective instrument validation study.

Participants: Three operators with no prior experience in operating the GDx VCC examined one randomly chosen eye of each of 30 randomly ordered subjects (15 glaucoma patients and 15 normal subjects).

Methods: Each study eye was scanned by the 3 operators in a random sequence during a single session. Five GDx parameters were analyzed: TSNIT (temporal, superior, nasal, inferior, temporal) Average, TSNIT Standard Deviation (SD), Superior Average, Inferior Average, and Nerve Fiber Indicator (NFI).

Main Outcome Measures: Retinal nerve fiber layer (RNFL) thickness GDx parameters across operators.

Results: A learning effect was not found for any of the operators for any of the 5 GDx parameters studied. Reproducibility for each GDx parameter was measured as the mean SD value of the measurements taken by the 3 operators. Interoperator reproducibilities for normal and glaucomatous eyes, respectively, were 0.8 and 1.6 μm (TSNIT Average), 1.7 and 2.2 μm (Superior Average), 1.6 and 2.7 μm (Inferior Average), 1.3 and 1.7 μm (TSNIT SD), and 1.4 and 4.6 μm (NFI). The data obtained by each operator were highly correlated with those of the other 2 operators.

Conclusions: With the commercially available GDx VCC, a learning effect was not found for 3 novice operators. In addition, RNFL measurements seem highly reproducible across operators. *Ophthalmology* 2005; 112:257–261 © 2005 by the American Academy of Ophthalmology.

Several different imaging modalities are able to quantify and follow structural changes in the retinal nerve fiber layer (RNFL) and the topography of the optic nerve head.^{1–3} These new technologies assist in the early diagnosis of glaucoma and in the follow-up of patients over time.

The GDx VCC (variable corneal compensator; Laser Diagnostics Technologies, Inc., San Diego, CA) is a scanning laser polarimeter that combines a confocal scanning laser ophthalmoscope and an integrated polarimeter. Polarized light undergoes a phase shift as it passes through polarizing media, such as the cornea, the crystalline lens, and the RNFL. This phase shift has been shown to correlate with the thickness of the RNFL.⁴ Via a 2-step approach, the GDx VCC first compensates for the corneal polarization

component by scanning the foveal area. This initial scan measures the birefringence pattern generated by the Henle fibers, and any nonuniformity detected at the foveal region is attributed to anterior segment (cornea and lens) polarization.⁵ In the second acquisition step, a compensated scan of the peripapillary retina is obtained.

As with any new diagnostic instrument, the validity of the data acquired might be reduced by variability, both interdevice⁶ and interoperator. Quantifying and accounting for such variability is crucial for devising algorithms that analyze progression.

Interoperator variability was previously studied for the optical coherence tomograph (OCT-2000, Humphrey Systems, Dublin, CA) in 10 normal and 10 glaucomatous eyes.³ Two qualified experienced examiners performed the tests. That study showed most of the variance (79%) to be a result of differences between patients, with only a modest proportion attributed to interoperator variability (2%). Kook et al evaluated the intraoperator and interoperator reproducibility of the GDx Nerve Fiber Analyzer (Laser Diagnostics Technologies) with normal Asian subjects. This scanning laser polarimeter lacks the corneal compensation implemented in the newer version, the GDx VCC. They found that interop-

Originally received: June 17, 2004.

Accepted: August 13, 2004.

Manuscript no. 240468.

From the Department of Ophthalmology, Hebrew University–Hadassah Medical Center, Jerusalem, Israel.

Dr Blumenthal has received research support from Laser Diagnostics Technologies, Inc., San Diego, California.

Correspondence to Eytan Z. Blumenthal, MD, Department of Ophthalmology, Hadassah University Hospital, P.O. Box 12000, Jerusalem 91120, Israel. E-mail: eblumenthal@md.huji.ac.il.

erator variability was not statistically significant ($P = 0.075$) and that the difference was even smaller if a single ellipse, drawn by one operator, was used in all measurements ($P = 0.1528$). The authors concluded that the GDx Nerve Fiber Analyzer might provide acceptable interoperator and intraoperator reproducibility.⁷

In the present study, we quantified the learning effect and interoperator variability for 3 operators with no prior experience in operating the GDx VCC in a prospective instrument validation study.

Materials and Methods

Subjects

One randomly selected eye from each of 30 subjects (15 glaucoma and 15 normal subjects) was scanned using one GDx VCC machine. All subjects underwent a complete eye examination, including a detailed medical history, slit-lamp examination, and either a standard full-threshold visual field (VF) analysis (Humphrey Field Analyzer [HFA], Humphrey-Zeiss Systems, Dublin, CA) or a full-threshold VF analysis using FDT (Welch Allyn, Inc., Skaneateles Falls, NY; Humphrey FDT VF, Humphrey-Zeiss Systems).⁸ All glaucoma patients had an HFA examination, whereas normal subjects underwent either an HFA or an FDT VF examination. Informed consent was obtained from all participants, and institutional ethics committee approval was obtained.

Inclusion criteria for normal subjects included a best-corrected visual acuity (VA) of 20/40 or better, a normal slit-lamp examination, an intraocular pressure of ≤ 22 mmHg (with no prior documentation of higher readings), a normal VF (HFA full-threshold or Swedish interactive threshold algorithm, with glaucoma hemifield test results within normal limits, or else an FDT full-threshold C-20 with ≤ 1 locations at the 5% cutoff and no locations showing a probability of 2% or worse on the pattern deviation plot), and no history of ocular surgery or laser treatment. Glaucoma patients were defined as having a standard full-threshold (or Swedish interactive threshold algorithm) VF, in which the glaucoma hemifield test results were outside normal limits, as well as optic disc contour changes typical of glaucoma, defined as excavation, rim thinning, notching, or RNFL wedge-shaped defects. The VF abnormality was repeatable on at least 2 consecutive examinations, with glaucomatous damage (mean \pm standard deviation [SD]) measuring -10.46 ± 8.99 decibels (mean deviation) and 5.89 ± 2.97 decibels (pattern standard deviation).

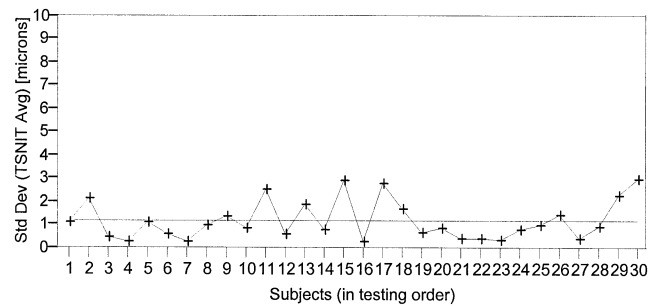


Figure 1. The standard deviation (SD) of the TSNIT (temporal, superior, nasal, inferior, temporal) Average GDx parameter was plotted by subject order. The horizontal line represents the mean SD across subjects. No learning effect was found for the 3 operators.

Operators

For this study, focusing on the examiners operating the GDx VCC machine, we recruited 3 operators with no prior experience in using the GDx VCC device. The operators were (1) an ophthalmic photographer with over 10 years' experience in operating most types of ophthalmic diagnostic and photographic equipment, (2) a glaucoma-trained ophthalmologist, and (3) a newly recruited staff member with no prior experience in operating any diagnostic or other ophthalmic equipment.

Before commencing the study, the 3 operators read the GDx VCC operation manual (Laser Diagnostic Technologies) and spent about 1 hour familiarizing themselves with the device and software. In addition, in the absence of formal training of any type, and the lack of an experienced operator available at any time before or during the study to provide coaching, each operator was allowed to perform 5 practice scans on normal eyes. We limited the practice to only 5 eyes before commencing data collection to allow the operators a minimal understanding of how to operate the device for proper data collection.

GDx VCC

The GDx VCC is the first commercially available scanning laser polarimeter with a variable cornea compensator. It differs from both the fixed cornea compensation GDx,⁷ the Access model, which it resembles externally, and the modified GDx mounted with a manual VCC,⁹⁻¹¹ a noncommercial investigational model, only a few units of which were ever produced.

Table 1. Interoperator Variability

	TSNIT Average	Superior Average	Inferior Average	TSNIT SD	NFI
Total (n = 30)	$F = 0.019$ $P = 0.982$	$F = 0.046$ $P = 0.955$	$F = 0.074$ $P = 0.929$	$F = 0.371$ $P = 0.691$	$F = 0.026$ $P = 0.975$
Normal (n = 15)	$F = 0.052$ $P = 0.950$	$F = 0.035$ $P = 0.966$	$F = 0.027$ $P = 0.973$	$F = 0.480$ $P = 0.622$	$F = 0.026$ $P = 0.975$
Glaucoma (n = 15)	$F = 0.112$ $P = 0.894$	$F = 0.075$ $P = 0.927$	$F = 0.279$ $P = 0.758$	$F = 0.328$ $P = 0.722$	$F = 0.040$ $P = 0.961$

NFI = Nerve Fiber Indicator; SD = standard deviation; TSNIT = temporal, superior, nasal, inferior, temporal. For each of the 5 tested GDx parameters, interoperator variability is presented as a 1-way analysis of variance with the F statistic along with the P value. Data are presented as overall variability and stratified by diagnosis (normal vs. glaucoma).

Table 2. Interoperator Reproducibility

	TSNIT Average	Superior Average	Inferior Average	TSNIT SD	NFI
Total (n = 30)	1.2 (0.3–3.0)	2.0 (0.2–5.4)	2.1 (0.2–5.6)	1.5 (0.3–3.4)	3.0 (0–9.9)
Normal (n = 15)	0.8 (0.3–2.2)	1.7 (0.2–4.3)	1.6 (0.2–4.3)	1.3 (0.5–3.3)	1.4 (0.6–2.6)
Glaucoma (n = 15)	1.6 (0.4–3.0)	2.2 (0.6–5.4)	2.7 (0.5–5.6)	1.7 (0.3–3.4)	4.6 (0–9.9)

NFI = Nerve Fiber Indicator; SD = standard deviation; TSNIT = temporal, superior, nasal, inferior, temporal. For each of the 5 tested GDx parameters, interoperator reproducibility is presented as mean SD (range). Data are presented as overall reproducibility and stratified by diagnosis (normal vs. glaucoma).

GDx VCC Measurements

One eye was randomly chosen for each subject and was scanned by each of the 3 operators in a random sequence, on the same day, during a single session. Only one successful scan per subject was taken by each examiner, as opposed to averaging 3 scans, both to imitate routine clinical practice and to highlight any variability that might have been averaged out. All scans were performed on undilated pupils. The normal and glaucoma subjects were examined in random order. Each scan was composed of a foveal scan (a corneal compensation measurement) and a peripapillary scan. Each operator reset the compensation values when scanning. Thus, compensation settings were not transferred from one operator to the next, but, rather, were reacquired with every operator. Each operator was instructed to achieve a high-quality scan, as judged by his or her subjective impression; to attempt to achieve a quality score, as determined by the machine, of at least 8/10; and to continue rescanning a reasonable number of attempts until such a result was achieved, with the number of scans per examiner noted. The disc margin, outlined by the software, was validated and corrected when needed by each operator during the scanning session. Once a satisfactory high-quality scan was achieved, it was saved, and a hard copy was printed.

GDx Parameters

From the 16 parameters appearing on the GDx printout,¹² we chose the 5 appearing on the first printout page. These parameters are best able to discriminate between normal and glaucomatous eyes and were chosen by the manufacturer to be the most informative: TSNIT (temporal, superior, nasal, inferior, temporal) Average (the average RNFL thickness along the measured ring); TSNIT SD (the SD of the data points along the measured ring); Superior Average (the average of the RNFL thickness values of the points along the ring along the superior 120° section); Inferior

Average (the average of the RNFL thickness values of the points along the ring along the inferior 120° section); and the Nerve Fiber Indicator, which is a support vector machine-derived parameter indicating the likelihood that an eye is glaucomatous.

Statistical Analysis

Lacking a gold standard for true thickness values, the operator learning effect was tested by plotting the SD of the 3 scans of each subject against the order in which the subjects were scanned. A narrowing of the distribution of measurements from the 3 operators, presented as a decrease in the SD against time, would have indicated a learning effect. Interoperator variability was tested for each GDx parameter using a 1-way analysis of variance (ANOVA), whereas interoperator reproducibility was quantified for each parameter by noting the mean, standard error (SE), and 95% confidence interval (CI) for each operator. We decided a priori that if a learning effect was found in the data, then the initial scans, up to the point that the learning effect plateaus, would need to be truncated for the reproducibility part of this study. Data were analyzed using JMP statistical software.¹³

Results

Thirty-two subjects were initially recruited to this study. One subject failed to fixate steadily enough to be scanned, despite good VA (>20/40), and was hence excluded from the study after repeated attempts at scanning by the first operator. A second subject was excluded after being scanned by all 3 operators, when we later discovered that 1 of the examiners forgot to reset the compensator, which led to a scan that relied on the compensation data collected by the previous examiner (the device’s default in repeated scans). These 2 subjects were excluded from the analyses. Of the remaining 30, 16 subjects were men and 14 were women. Mean ages of

Table 3. Operator Variability

	Operator	TSNIT Average	Superior Average	Inferior Average	TSNIT SD	NFI
Mean	A	47.74	56.3	54.64	18.34	40.63
	B	47.29	55.45	53.42	17.18	42.27
	C	47.64	55.4	54.14	17.97	40.9
SE	A	1.7	2.39	2.22	1.04	5.37
	B	1.81	2.34	2.36	0.91	5.47
	C	1.73	2.33	2.19	0.95	5.54
CI	A	3.34	4.69	4.36	2.05	10.52
	B	3.55	4.58	4.62	1.78	10.73
	C	3.39	4.57	4.3	1.86	10.86

NFI = Nerve Fiber Indicator; SD = standard deviation; TSNIT = temporal, superior, nasal, inferior, temporal. Operator variability for the five GDx parameters is presented as mean, standard error (SE) and 95% confidence interval (CI), in microns.

the normal and glaucoma groups were 42.6 ± 14.3 years and 67.1 ± 10.8 years, respectively. The desired scanning quality was achieved with only 1 scanning attempt in 77% of cases (69/90), 2 to 3 attempts were needed in 21% of cases (19/90), and >3 attempts were needed in the remaining 2% (2/90). The machine-determined quality scores obtained were 10 in 8%, 9 in 77%, 8 in 12%, and 7 in 3%. For those who scored only 7, multiple attempts were made to achieve a quality score of ≥ 8 , but this could not be achieved by the operator.

Learning Effect

Having no true gold-standard value against which to compare the data, we decided to test for a learning effect by evaluating the change over time of the SD of the measurements taken by the operators for each subject. Plotting the SD for each subject against the order in which the subjects were examined and fitting a line to this graph for each of the 5 GDx parameters yielded no ostensible learning effect. Figure 1 presents the plot for TSNIT Average as an example.

Interoperator Variability

Table 1 presents the variability among the 3 operators using 1-way ANOVA (*F* test and *P* value). *F*s ranged between 0.019 and 0.371 for the various GDx parameters tested; the differences found between the operators were all nonsignificant. Stratifying the analysis for normal and glaucoma subjects gave similar results—namely, no significant differences between the operators in either of the groups.

Interoperator Reproducibility

Table 2 presents the interoperator reproducibility as the mean SD across the 3 operators for each of the 5 GDx parameters studied. Data are presented as mean (range) and are also stratified by diagnosis (normal vs. glaucoma). Table 3 presents the interoperator reproducibility expressed as mean, SE, and 95% CI provided for each of the 3 operators, for each of the 5 GDx parameters separately. The data presented in Tables 2 and 3 show high reproducibility amid the operators.

Discussion

Initial evaluation of any new diagnostic imaging device must include reproducibility, which is composed of inter-device, interoperator, intrasession, intersession, and short- and long-term variability. We have recently demonstrated, in a study evaluating 13 units, that the GDx VCC device showed high interdevice reproducibility.⁶ In the current study, we evaluated both the operator learning effect and the interoperator reproducibility.

In the absence of a model eye for scanning laser polarimetry, such as the model eye used in standardized ophthalmic echography,¹⁴ we chose to submit a human eye to the test. Underlying this study is the assumption that any changes in RNFL values observed during the course of an hour should be attributed to interoperator variability rather than to true anatomical changes in the RNFL thickness.

Operator learning effect was evaluated by Kook et al⁷ on a previous model of the scanning laser polarimeter. They did not specify what initial training their operators had.

Moreover, each of their operators performed 6 consecutive measurements of each subject before moving on to the next, thus minimizing any variability that may arise in the usual practice of performing one examination per subject. Only normal subjects were tested in their study.

The operators chosen to participate in our study represent the spectrum of operators who may operate such equipment, demonstrating different amounts of past experience in operating other ophthalmic devices. Our results favor the assumption that neither prior experience with ophthalmic imaging devices nor glaucoma training is necessary for operating the GDx VCC. When comparing the spread of measurements for normal eyes to those found for glaucomatous eyes, both with a VA of 20/40 or higher, no significant difference was found among the 3 operators, and the measurements were found to be reproducible. The lack of a learning effect in our study could be attributed to the fact that the operation of the GDx VCC device requires only a very short learning experience, which fell within the 5 prestudy test runs made by each operator.

In conclusion, no learning effect could be demonstrated in the data pooled from these novice operators. In addition, we found high reproducibility among the different operators, independent of their prior experience in operating ophthalmic diagnostic equipment. Therefore, it seems that patient follow-up may be independent of both the GDx VCC operator and the operator's past experience in operating the device.

References

1. Burk RO, Volcker HE. Current imaging of the optic disk and retinal nerve fiber layer. *Curr Opin Ophthalmol* 1996;7:99–108.
2. Funk J. Imaging of the optic disc in glaucoma: which way to go? *Curr Opin Ophthalmol* 1998;9:71–6.
3. Blumenthal EZ, Williams JM, Weinreb RN, et al. Reproducibility of nerve fiber layer thickness measurements by use of optical coherence tomography. *Ophthalmology* 2000;107:2278–82.
4. 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.
5. Garway-Heath DF, Greaney MJ, Caprioli J. Correction for the erroneous compensation of anterior segment birefringence with the scanning laser polarimeter for glaucoma diagnosis. *Invest Ophthalmol Vis Sci* 2002;43:1465–74.
6. Blumenthal EZ, Frenkel S. Inter-device reproducibility of the scanning laser polarimeter with variable corneal compensation. *Eye*. In press.
7. Kook MS, Sung K, Park RH, et al. Reproducibility of scanning laser polarimetry (GDx) of peripapillary retinal nerve fiber layer thickness in normal subjects. *Graefes Arch Clin Exp Ophthalmol* 2001;239:118–21.
8. Sample PA, Bosworth CF, Blumenthal EZ, et al. Visual function-specific perimetry for indirect comparison of different ganglion cell populations in glaucoma. *Invest Ophthalmol Vis Sci* 2000;41:1783–90.
9. Greenfield DS, Knighton RW, Huang XR. Effect of corneal polarization axis on assessment of retinal nerve fiber layer thickness by scanning laser polarimetry. *Am J Ophthalmol* 2000;129:715–22.

10. Weinreb RN, Bowd C, Greenfield DS, Zangwill LM. Measurement of the magnitude and axis of corneal polarization with scanning laser polarimetry. *Arch Ophthalmol* 2002; 120:901–6.
11. Zhou Q, Weinreb RN. Individualized compensation of anterior segment birefringence during scanning laser polarimetry. *Invest Ophthalmol Vis Sci* 2002;43:2221–8.
12. GDx VCC [computer program]. Version 5.1.0. San Diego: Laser Diagnostics Technologies.
13. JMP [computer program]. Version 5.0. Cary, NC: SAS Institute; 2003.
14. Ossoinig KC. Standardized echography: basic principles, clinical applications, and results. *Int Ophthalmol Clin* 1979; 19:127–210.