The Effect of Atypical Birefringence Patterns on Glaucoma Detection Using Scanning Laser Polarimetry with Variable Corneal Compensation

Christopher Bowd, Felipe A. Medeiros, Robert N. Weinreb, and Linda M. Zangwill

PURPOSE. The purpose of this study was to investigate the effect of the presence of atypical birefringence patterns, as measured by the typical scan score (TSS), on the diagnostic accuracy of a scanning laser polarimeter (the GDx VCC; Carl Zeiss Meditec, Inc., Dublin, CA) assessed by receiver operating characteristic (ROC) curves for discriminating between glaucoma and healthy eyes.

METHODS. Two hundred thirty-three glaucomatous eyes (repeatable abnormal visual fields by pattern standard deviation [PSD] and/or glaucoma hemifield test [GHT]) from 153 patients with glaucoma and 104 eyes from 71 healthy participants enrolled in the UCSD Diagnostic Innovations in Glaucoma Study (DIGS) were imaged using the GDx VCC. An ROC regression model was used to evaluate the influence of the covariates TSS; disease severity, defined as standard automated perimetry (SAP) mean deviation [MD]; and age in years on the diagnostic accuracy of the GDx parameters nerve fiber indicator [NFI], TSNIT (temporal, superior, nasal, inferior, temporal) average thickness, superior average thickness, inferior average thickness, and TSNIT standard deviation. Areas under the ROC curve were calculated for specific levels of the covariates according to the results provided by the model.

RESULTS. TSS and SAP MD significantly affected the diagnostic accuracy of each investigated GDx VCC parameter. Low TSSs, indicating the presence of atypical scans, were associated with decreased accuracy. For NFI, ROC curve areas ranged from 0.749 (when TSS = 20) to 0.904 (when TSS = 100). A similar influence of TSS was found for other parameters. In addition, diagnostic accuracy increased with increasing disease severity. For instance, for NFI, ROC curve areas ranged from 0.853 (when SAP MD = −3) to 0.954 (when SAP MD = −15).

CONCLUSIONS. The diagnostic accuracy of GDx VCC parameters is affected by disease severity and is adversely affected by the presence of atypical retardation patterns (i.e., decreasing TSS). GDx VCC scans with atypical scan patterns should be interpreted with caution when used in clinical practice. (Invest Ophtalmol Vis Sci. 2007;48:223–227) DOI:10.1167/iovs.06-0787

From the Hamilton Glaucoma Center, Department of Ophthalmology, University of California, San Diego, California.


Submitted for publication July 11, 2006; revised August 21, 2006; accepted October 11, 2006.

Disclosure: C. Bowd, None; F.A. Medeiros, Carl Zeiss Meditec, Inc. (F); R.N. Weinreb, Carl Zeiss Meditec, Inc. (F, R); Heidelberg Engineering GmbH (F, R); L.M. Zangwill, Carl Zeiss Meditec, Inc. (F, R); Heidelberg Engineering GmbH (F, R)

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Corresponding author: Christopher Bowd, Hamilton Glaucoma Center, Department of Ophthalmology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92039-0946; cbowd@eyecenter.ucsd.edu.

Copyright © Association for Research in Vision and Ophthalmology

223

Scanning laser polarimetry (SLP) provides real-time, objective measurements for assessing retinal nerve fiber layer (RNFL) thickness in glaucoma. Recent studies have shown that using SLP with variable corneal compensation (GDx VCC; Carl Zeiss Meditec, Inc., Dublin, CA) improves diagnostic precision and strengthens cross-sectional structure-to-function associations, compared with using SLP with fixed corneal compensation (FCC).1-5

However, GDx with VCC has been criticized for providing a relatively large number of artifact-laden images, called atypical scans. Atypical scans are scans with an atypical birefringence (i.e., retardance) (ABP) that is not representative of RNFL thickness patterns found histologically (i.e., increased birefringence superiorly and inferiorly, indicating thicker RNFL compared with decreased birefringence temporally and nasally, indicating thinner RNFL). Rather, in addition to high birefringence superiorly and inferiorly, scans with ABP display increased birefringence in the temporal and nasal quadrants in radial patterns centered on and surrounding the entire optic disc. Current GDx VCC software includes an exportable parameter called typical scan score (TSS) that provides a numerical representation of the degree of "typicalness" in each scan, ranging from 1 (extremely atypical) to 100 (typical; see Fig. 1). This parameter is the output of a support vector machine machine-learning classifier trained to identify scans that were subjectively assessed as atypical by instrument developers. Previous studies have suggested cutoff values for this parameter that result in the exclusion of subjectively classified atypical scans.6,7 However, although these cutoffs provide a way to exclude atypical scans, they do not directly address the effect of atypical scans on the diagnostic ability of GDx VCC.

The purpose of the present study was to investigate the effect of TSS on the accuracy of the GDx VCC for classifying eyes as healthy or glaucomatous. We used a receiver operating characteristic (ROC) regression technique to analyze the covarying effects of TSS, glaucoma severity, and age on ROC curves describing the diagnostic performance of several GDx VCC parameters.

METHODS

Participants

One or two eyes were selected from each of 242 participants enrolled in the University of California, San Diego–based longitudinal Diagnostic Innovations in Glaucoma Study (DIGS) for study. All participant eyes had GDx VCC imaging and a reliable visual field test within 6 months. A total of 337 eyes were studied: 104 were classified as healthy and 233 were classified as glaucomatous.

Each study participant underwent a comprehensive ophthalmic evaluation including review of medical history, best corrected visual acuity testing, slit lamp biomicroscopy, IOP measurement with Goldmann applanation tonometry, gonioscopy, dilated fundus examination with a 78-D lens, simultaneous stereoscopic optic disc photography (TRC-SS; Topcon Instruments Corp. of America, Paramus, NJ), and standard automated perimetry (SAP) using the 24-2 Swedish Interactive Threshold Algorithm (SITA; Humphrey Field Analyzer II; Carl Zeiss
Meditec, Inc.). To be included in the study, participants had to have best corrected acuity better than or equal to 20/40, spherical refraction within ±5.0 D and cylinder correction within ±3.0 D, and open angles on gonioscopy.

Participants were excluded if they had a history of intraocular surgery except for uncomplicated cataract or glaucoma surgery. We also excluded all participants with nonglaucomatous secondary causes of elevated IOP (e.g., iridocyclitis, trauma), other intraocular eye disease, other diseases affecting the visual field (e.g., pituitary lesions, demyelinating diseases, HIV or AIDS, or diabetic retinopathy), medications known to affect visual field sensitivity, or problems other than glaucoma that affect color vision.

Glaucomatous eyes were defined as those with repeatable (two consecutive) SAP results outside normal limits by pattern standard deviation (PSD; P < 5%) or glaucoma hemifield test (GHT). The first abnormal SAP was on or before the imaging date. If only one eye of a study participant had repeatable abnormal visual fields, the nonqualifying eye was excluded from study. Average SAP mean deviation (MD) of the glaucomatous eyes within 6 months of GDx VCC imaging was −5.67 dB (median = −3.51, SD = 5.71; range = −31.46 to +1.00) and average PSD was 5.35 dB (median = 3.58, SD = 3.73; range = 1.38 to 15.58). The mean age of the patients with glaucoma (i.e., those with glaucoma in at least one eye, n = 173) was 68.4 years (median = 69.8; SD = 11.6; range = 33.2 to 91.8), 91 (53%) were women, and 146 (84%) were self-reported white. Eighty-two (47%) were pseudophakic.

Healthy eyes were defined as those with healthy-appearing optic discs on clinical examination, SAP results (MD, PSD, and GHT) within normal limits, and no history of IOP > 22 mm Hg. Average SAP MD of the healthy eyes was −0.86 dB (median = −0.75, SD = 1.38; range = −4.78 to +1.75) and average PSD was 1.62 dB (median = 1.54, SD = 0.45; range = 0.99 to 3.61). Both global indices differed significantly and were better in healthy than in glaucomatous eyes (t-tests, P < 0.001). The mean age of the healthy participants (n = 69) was 51.0 years (median = 55.3, SD = 16.4; range = 18.3 to 83.6) and was significantly lower than that of the patients with glaucoma (t-test, P < 0.001). Fifty (72%) healthy participants were women and 56 (81%) were self-reported white. Two of the healthy eyes were pseudophakic.

This research adhered to the tenets of the Declaration of Helsinki and the Health Insurance Portability and Accountability Act (HIPAA).

Instrumentation

All participants’ eyes were imaged with a commercially available scanning laser polarimeter (GDx VCC; software version 5.0.1; Carl Zeiss Meditec, Inc.). Principles of this technology have been provided in detail elsewhere. In general, scanning laser polarimetry measures the retardation of light reflected from the birefringent RNFL fibers and provides an estimated RNFL thickness based on the linear relationship between observed retardation, measured using a prototype instrument, and RNFL thickness determined historically. The GDx VCC employs a variable corneal polarization compensator that allows eye-specific compensation of anterior chamber birefringence. After determining the axis and magnitude of corneal polarization in each eye by macular scanning,3 three appropriately compensated retinal polarization images per eye were automatically obtained and combined to form each mean image used for analysis. Only well focused, evenly illuminated and centered scans with residual anterior segment retardation ≤ 12 nm and SD ≤ 7 μm, determined by the perimeter software, were included (cutoffs suggested by Qienyuan Zhou PhD, Carl Zeiss Meditec, Inc., verbal communication, June 2005).

In this study we examined the effect of the software-determined TSS on the ability of several parameters to discriminate between healthy and glaucomatous eyes. The TSS is a continuous variable ranging from 0 to 100 and is the result of a support vector machine analysis of VCC data labeled for training based on the subjective appearance of each scan (typical versus atypical). TSS is based on the slope, standard deviation, and average magnitude of RNFL thickness measurements from the edge of the optic disc extending outward to 20°. Low TSSs indicate very atypical scans (Fig. 1, leftmost image, TSS = 20) and high typical scan scores indicate very typical scans (Fig. 1, rightmost image, TSS = 100).

RNFL measurement parameters investigated in this study were the nerve fiber indicator (NFI), TSNIT (circumpapillary RNFL thickness measured under the automatically defined a 3.2-mm diameter calculation circle: T, temporal sector; S, superior sector; N, nasal sector; I, inferior sector) average, superior average, inferior average, and TSNIT SD. These parameters were selected because they are those provided on the instrument print-out designed for clinical use. We also examined the effects of the covariates disease severity (defined as SAP MD) and age in an ROC regression model with TSS.

Statistical Analysis

In the present study, we used an ROC regression modeling technique to evaluate the influence of atypical scan patterns on the diagnostic accuracy of the GDx VCC in glaucoma. This modeling approach has been recently applied by Medeiros et al.7 to evaluate the influence of covariates on the performance of diagnostic tests in glaucoma. This methodology allows the evaluation of the influence of covariates on the diagnostic performance of the test, so that ROC curves for specific values of the covariate of interest can be obtained. Also, it allows adjustment for the possible confounding effects of other covariates. Details of the modeling procedure have been described previously.9,10 In brief, the $\text{ROC}_{\text{x},\text{MD},\text{q}}(q)$ is the probability that a subject with diseased eyes with disease-specific covariates $X_1$ (that is, covariates specific to subjects with diseased eyes such as disease severity, for example) and common covariates $X$ (covariates common to both subjects with diseased eyes and healthy subjects) has test results $Y_2$ that are greater than or equal to the $q$th quantile of the distribution of tests results from subjects with nondiseased eyes. That is, when the specificity of the test is $1 - q$, the sensitivity is $\text{ROC}_{\text{x},\text{MD},\text{q}}(q)$.

The general ROC regression model can be written as

$$\text{ROC}_{\text{x},\text{MD},\text{q}}(q) = \Phi(\alpha_1 + \alpha_2 \Phi^{-1}(q) + \beta X + \beta_0 X_0)$$

where the coefficients $\alpha_1$ and $\alpha_2$ are the intercept and slope of the ROC curve, respectively, and $\Phi$ is the normal cumulative distribution function. If the coefficient for a specific variable $X(B)$ is greater than 0,
RESULTS

The estimates of the coefficients of the ROC regression model for the parameter NFI are shown in Table 1. To provide more meaningful values, the variables severity and age were centered on their approximate mean values in glaucomatous eyes (−5.0 dB and 65 years, respectively) when estimating the coefficients. The presence of atypical scan patterns had a significant influence on the diagnostic performance of the NFI, as indicated by the statistically significant value attributed to the coefficient representing TSS (P = 0.004). Lower TSSs (indicating more atypical scans) were associated with decreased diagnostic ability. There was no statistically significant influence of the TSS on the shape of the ROC curve, that is, the influence of TSS was similar throughout the range of false positive (i.e., 1 − specificity) values, as indicated by the non-significant coefficient associated with the interaction term $TSS \times \Phi^{-1}(q)$ ($P = 0.073$).

The severity of disease, as measured by MD, also had a significant influence on the diagnostic performance of the parameter NFI, as indicated by the statistically significant value attributed to the coefficient representing severity ($P = 0.021$). As expected, lower values of MD were associated with increased accuracy to distinguish between healthy and glaucomatous eyes. Age did not influence the performance of the NFI ($P = 0.479$).

Similar results were found for ROC regression models obtained for the parameters TSNIT average, superior average, inferior average, and TSNIT SD. That is, lower values of TSS were associated with decreased performance for all these parameters, whereas lower values of SAP MD were associated with increased performance. Based on the results provided by the regression models for each parameter, we calculated areas under the ROC curves for each parameter at specific levels of the covariate TSS. Table 2 shows ROC curve areas for each investigated GDx VCC parameter at arbitrary TSSs of 20, 50, 70, and 100. Figure 1 shows examples of scans associated with these cutoffs. Figure 2 shows ROC curves describing the diagnostic accuracy (in terms of sensitivity and 1 − specificity) for the GDx NFI parameter at the same arbitrary TSS cutoffs. Table 3 shows ROC curve areas for each investigated GDx VCC parameter at arbitrary SAP MD values of −15, −10, −5 and −3 dB.

DISCUSSION

The present study demonstrates the significant affect of atypical birefringence patterns (ABP) on the diagnostic accuracy of
glaucoma detection using GDx VCC RNFL measurements. The more atypical the birefringence pattern (indicated by decreased TSS), the less accurate the glaucoma-versus-normal classification. We used an ROC covariance regression model that allows the degree of typicality to function as a continuous variable. This method eliminated the need to choose discrete “bins” that may not contain sufficient data points to allow between-group analyses (e.g., ANOVA). In addition, we confirmed the significant effect of disease severity (defined by SAP MD) on the diagnostic accuracy of GDx VCC measurements demonstrated previously. In a recent study investigating the effect of ABP on VCC measurements, Bagga et al. demonstrated a decreased association between average RNFL measurements obtained with VCC and those obtained with optical coherence tomography (OCT) for atypical scans (defined subjectively) compared with typical scans ($R^2 = 0.42, 0.71$, respectively). A decreased association with OCT measured rim area also was reported ($R^2$ for atypical scans = 0.27; $R^2$ for typical scans = 0.76). These results suggest that measurements from atypical scans are less accurate than those from typical scans, because a strong association among structural measurements is expected.

In a study comparing the diagnostic precision of GDx VCC measurements with and without ABP for discriminating between healthy and glaucomatous eyes, Da Pozzo et al. reported much better discrimination using typical scans. They reported ROC curves for discriminating between healthy eyes with typical birefringence patterns (TBP, defined as scans with TSS $\geq 70$) and glaucomatous eyes with TBP and between the same healthy eyes and glaucomatous eyes with ABPs (TSS $< 70$). For the former comparison, ROC curve areas $\geq 0.90$ were reported for 9 of 14 GDx VCC parameters (range = 0.902 for TSNIT SD to 0.980 for NFI) and for the latter comparison ROC curve areas $\geq 0.90$ were

**Table 3.** ROC Curve Areas for Each GDx VCC Parameter at Four Arbitrarily Chosen Levels of SAP MD

<table>
<thead>
<tr>
<th>Parameter</th>
<th>MD</th>
<th>MD</th>
<th>MD</th>
<th>MD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$-15$</td>
<td>$-10$</td>
<td>$-5$</td>
<td>$-3$</td>
</tr>
<tr>
<td>NFI</td>
<td>0.954</td>
<td>0.922</td>
<td>0.876</td>
<td>0.853</td>
</tr>
<tr>
<td>TSNIT average</td>
<td>0.936</td>
<td>0.900</td>
<td>0.852</td>
<td>0.828</td>
</tr>
<tr>
<td>Superior average</td>
<td>0.948</td>
<td>0.890</td>
<td>0.796</td>
<td>0.748</td>
</tr>
<tr>
<td>Inferior average</td>
<td>0.921</td>
<td>0.878</td>
<td>0.821</td>
<td>0.794</td>
</tr>
<tr>
<td>TSNIT SD</td>
<td>0.904</td>
<td>0.839</td>
<td>0.752</td>
<td>0.711</td>
</tr>
</tbody>
</table>

Continuous variables TSS and age were set at approximate mean values (for glaucomatous eyes) of 80 and 65 years, respectively. This modeling approach allows the evaluation of the influence of covariates on the diagnostic performance of the test, so that ROC curves for specific values of the covariate of interest can be obtained. Also, it allows adjustment for the possible confounding effects of other covariates.
reported for 0 of 14 parameters. However, four parameters were able to discriminate between healthy eyes and glaucomatous eyes with ABP with ROC areas of approximately 0.85: inferior ratio, superior ratio, maximum modulation, and NFI (although ROC area for NFI decreased significantly when ABPs were included). In the present study, we reported an NFI ROC curve area of 0.86 for eyes with TSS of 70, similar to results reported by Da Pozzo et al. We did not investigate ratio parameters, but it is not surprising that their diagnostic precision was quite good when ABPs were involved. GDx VCC superior and inferior thickness ratio parameters are defined as the ratio of superior and inferior RNFL thickness (respectively) to temporal RNFL thickness. Because ABPs are more likely to result in increased temporal RNFL thickness measurements, compared with increased superior and inferior measurements (that are already quite thick), it is likely that in glaucomatous eyes with ABPs, superior and inferior ratios are decreased compared with those in glaucomatous eyes with TBPs. In fact, Da Pozzo et al. showed decreased ratio measurements and increased ROC curve areas in glaucomatous eyes with ABPs compared with those with TBPs. Although discrimination was good with these parameters, we disagree with these investigators, who suggest that ratio parameters should be used as diagnostic indicators in eyes with ABPs, because, despite acceptable discrimination, the discrimination task is based on artifactual measurements.

In the present study, 82 (35%) of 233 glaucomatous eyes were pseudophakic compared with 2 of 102 healthy eyes. Previous studies have shown that cataract removal and subsequent lens replacement results in increased polarimetry-measured RNFL thickness, although this effect may depend on the material composing the replacement lens and the type of cataract initially present and may be caused by changes in lens birefringence. If cataract removal and lens replacement increased RNFL measurements in the present study, then differences between glaucoma and normal eyes may be underestimated. Considering the glaucomatous eyes only, patients with pseudophakic eyes were older than individuals with phakic eyes (73.5 and 65.8 years, respectively: t-test, \( P < 0.001 \)). In addition, patients with glaucoma with pseudophakic eyes had lower TSS than did phakic eyes (74.4 and 86.0, respectively: t-test, \( P < 0.001 \)). We adjusted ROC curve areas for age and TSS so that differences in these variables between experimental groups have been accounted for in our analyses.

Recently, a software upgrade has been introduced that corrects GDx VCC scans with ABPs. This technique, called "enhanced corneal compensation" (ECC) assumes that ABPs are the result of a low signal-to-noise ratio in certain eyes that results from increased light scattering. ECC superimposes a known large birefringence onto the RNFL birefringence to boost the signal-to-noise ratio. After image acquisition the birefringence bias is subtracted from the total measured birefringence to yield the "true" RNFL birefringence. Using this technique, decreased RNFL thickness measurements and increased TSSs in eyes with TBP compared with eyes with ABP were reported. In addition the subjective appearance of ABP was eliminated. Preliminary studies suggest that diagnostic accuracy for glaucoma detection was improved using GDx ECC compared with GDx VCC (Kim KH, et al. IOVS 2006;47:ARVO EAbstract 3352) and that structure-function associations (associations between GDx VCC-measured RNFL thickness and SAP-measured visual sensitivity) were somewhat stronger using GDx ECC (Kook MS, et al. IOVS 2006;47:ARVO EAbstract 3670).

In conclusion, the current results and those reported previously suggest that the diagnostic accuracy of GDx VCC RNFL thickness measurements for detecting glaucoma is decreased significantly when atypical birefringence patterns are present. We therefore suggest scans with ABP should be interpreted with caution when used in clinical practice. It is unclear whether longitudinal analyses using atypical GDx VCC scans will be affected.

References


