

**FHS PUBLIC ACCESS**

Author manuscript

*J Glaucoma*. Author manuscript; available in PMC 2016 October 01.

Published in final edited form as:

*J Glaucoma*. 2015 ; 24(8): 561–567. doi:10.1097/IJG.0000000000000065.

## Combining Frequency Doubling Technology Perimetry and Scanning Laser Polarimetry for Glaucoma Detection

**Jean-Claude Mwanza, MD, MPH, PhD<sup>1</sup>, Joshua L. Warren, PhD<sup>2</sup>, Jessica T. Hochberg, BS<sup>1</sup>, Donald L. Budenz, MD, MPH<sup>1</sup>, Robert T. Chang, MD<sup>1</sup>, and Pradeep Y. Ramulu, MD, PhD<sup>1</sup>**<sup>1</sup>Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, Miami, FL<sup>2</sup>Department of Biostatistics, Gillings School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC

### Abstract

**Purpose**—To determine the ability of frequency doubling technology (FDT) and scanning laser polarimetry with variable corneal compensation (GDx-VCC) to detect glaucoma when used individually and in combination.

**Methods**—One hundred and ten normal and 114 glaucomatous subjects were tested with FDT C-20-5 screening protocol and the GDx-VCC. The discriminating ability was tested for each device individually and for both devices combined using GDx-NFI, GDx-TSNIT, number of missed points of FDT, and normal or abnormal FDT. Measures of discrimination included sensitivity, specificity, area under the curve (AUC), Akaike's information criterion (AIC), and prediction confidence interval lengths (PIL).

**Results**—For detecting glaucoma regardless of severity, the multivariable model resulting from the combination of GDx-TSNIT, number of abnormal points on FDT (NAP-FDT), and the interaction GDx-TSNIT \* NAP-FDT (AIC: 88.28, AUC: 0.959, sensitivity: 94.6%, specificity: 89.5%) outperformed the best single variable model provided by GDx-NFI (AIC: 120.88, AUC: 0.914, sensitivity: 87.8%, specificity: 84.2%). The multivariable model combining GDx-TSNIT, NAP-FDT, and interaction GDx-TSNIT\*NAP-FDT consistently provided better discriminating abilities for detecting early, moderate and severe glaucoma than the best single variable models.

**Conclusions**—The multivariable model including GDx-TSNIT, NAP-FDT, and the interaction GDx-TSNIT \* NAP-FDT provides the best glaucoma prediction compared to all other multivariable and univariable models. Combining the FDT C-20-5 screening protocol and GDx-VCC improves glaucoma detection compared to using GDx or FDT alone.

Correspondence: Donald L. Budenz, MD, MPH, Department of Ophthalmology, 5151 Bioinformatics Bldg, CB #7040, 130 Mason Farm Road, Chapel Hill, NC 27599. [donald\\_budenz@med.unc.edu](mailto:donald_budenz@med.unc.edu).

Present address: Drs. Jean-Claude Mwanza and Donald L. Budenz, Department of Ophthalmology, University of North Carolina at Chapel Hill, Chapel Hill, NC; Dr. Robert T. Chang, Department of Ophthalmology, Stanford University, Palo Alto, CA; Dr. Pradeep Y. Ramulu, Wilmer Eye Institute, Johns Hopkins University, Baltimore, MD.

Disclosure: The authors declare no conflict of interest.

Presented at the Association for Research on Vision and Ophthalmology, May 1-5, 2011, Fort Lauderdale, FL, USA.

## INTRODUCTION

Glaucoma is an ocular disease of public health concern due to its high prevalence and significant morbidity worldwide. Because early treatment is effective in delaying the onset of glaucoma<sup>1</sup> or the progression of glaucomatous damage,<sup>2</sup> it is critical from a public health perspective to develop simple, quickly administered, easily interpreted, sensitive and specific methods to identify individuals with glaucoma. Also, finding methods that use non-physicians in the screening setting would represent a better use of limited resources. Frequently used methods in these groups of individuals include measurement of intraocular pressure (IOP) and ophthalmoscopy estimation of cup-to-disc ratio (CDR). However, these methods have low sensitivity and specificity in distinguishing normal eyes from those with glaucoma.<sup>3,4</sup> Standard automated perimetry (SAP), another important glaucoma diagnostic test, is difficult to use in population screening due to its non-portability, long testing time, fatigue-related artifacts, and subject's response variability. In recent years, other technologies have been developed to detect glaucomatous visual field (VF) loss<sup>5</sup> and damage to the retinal nerve fiber layer (RNFL).<sup>6</sup>

Given that glaucomatous damage affects both optic nerve structure and function, it is advantageous to consider both aspects for glaucoma diagnosis. Indeed, it has been shown that combining structure and function tests improves the diagnostic detection compared to either test used individually.<sup>7</sup> With regard to frequency doubling technology (FDT) and scanning laser polarimetry (SLP), two earlier studies screened for glaucoma after combining these two instruments. One of them found that combining these two instruments improves the sensitivity to detect glaucoma as compared to using either instrument individually,<sup>8</sup> whereas the other did not find any improvement in diagnostic performance after combining the two instruments.<sup>9</sup> However, both of these studies used SLP with fixed corneal compensation (FCC), which had the drawback of sometimes providing erroneous measurements inherent to incomplete compensation of corneal birefringence. The variable corneal compensation (VCC) feature later added to upgrade the GDx allows good visualization of the RNFL<sup>10</sup> and better differentiation between healthy and glaucomatous eyes compared to GDx-FCC.<sup>11, 12</sup> On the other hand, studies have reported that FDT testing results strongly correlate with those of conventional Humphrey Visual Field (HVF)<sup>13, 14</sup> and that FDT may detect glaucomatous functional loss earlier than SAP.<sup>15</sup> One concern about FDT as a screening tool has been its high false positive rate,<sup>16</sup> which has led to the recommendation that FDT should not be used alone as a screening test for glaucoma.<sup>17</sup> The aim of the present study was to evaluate the performance of FDT and GDx-VCC in detecting glaucoma when used individually and in combination.

## METHODS

### Subjects

One randomly selected eye from each of 110 normal and 114 glaucomatous subjects was included in this study. Written informed and Health Insurance Portability and Accountability Act consents were obtained from all participants after the study was approved by the Human Subject Research Office Committee of the Miami Miller School of Medicine. All participants were recruited among outpatients attending comprehensive ophthalmology,

optometry and glaucoma clinics of the Anne Bates Leach Eye Hospital, Bascom Palmer Eye Institute, Miller School of Medicine, University of Miami. All subjects underwent an eligibility screening eye examination. The examination included measurement of visual acuity and IOP, slit-lamp examination and fundus ophthalmoscopy and a review of previous HVF (Carl Zeiss Meditec Inc, Dublin, California) tests. Subjects were classified as normal if they had best-corrected visual acuity 20/40; IOP 21 mm Hg; normal fundus examination with cup-to-disc ratio (CDR) up to 0.6, but without evidence of optic nerve or macular disease, or interocular asymmetry in CDR 0.2, focal thinning of the optic disc rim, optic disc drusen, pallor or hemorrhage, age-related macular degeneration, or diabetic retinopathy. Subjects with glaucoma were included if they had a diagnosis of glaucoma as confirmed by a glaucoma specialist based on glaucomatous optic nerve head (ONH) changes with accompanying glaucomatous VF loss in at least one eye. At least two reliable SITA standard 24-2 HVF tests were required, with the most recent of them performed within one year of the enrollment date. VF defects were considered glaucomatous if they met the minimum criteria for a field defect: the Glaucoma Hemifield Test (GHT) was outside normal limits, the pattern standard deviation had  $P$  values  $< 5\%$ , or if there was a cluster of 3 or more points in the pattern deviation plot in a single hemifield (superior or inferior) with  $P$  values  $< 5\%$ , one of which must have a  $P$  value  $< 1\%$ . The severity of glaucoma was defined based on the Hodapp-Parrish-Anderson VF severity grading scale.<sup>18</sup> Exclusion criteria for both normal and glaucomatous subjects included age  $< 18$  years, best corrected VA worse than 20/40 in both eyes, a history of retinal disease (macular degeneration, diabetic retinopathy) or optic nerve disease (non-glaucomatous optic neuropathy).

### Frequency Doubling Technology Perimetry

FDT is a portable lightweight instrument resistant to refractive blur up to 6 diopters, thus not requiring correction of refractive errors.<sup>19</sup> Testing was therefore completed with the subject's habitual correction. All participants were instructed on how to complete the FDT test using an instruction card and through simulated testing in demonstration mode. All selected eyes were evaluated with the FDT perimeter (software version 4.00.0, Welch Allyn, Humphrey Systems, Carl Zeiss Meditec Inc., Dublin, California) with undilated pupils using the C-20-5 screening protocol. Details about the functioning and data acquisition have been extensively presented in the review by Anderson and Johnson.<sup>20</sup> If the first testing of an eye showed one or more abnormal points, a second confirmatory test was performed in that eye. To be considered abnormal, one or more locations had to be identified at least at  $P < 5\%$  at the same location on a second test. The FDT was considered reliable if false positives, false negatives, and fixation losses were all less than 33%. Patients with at least one reliability parameter exceeding 33% were excluded. Unreliable tests consisting of fixation losses and/or false positives were not considered for data analysis.

### Scanning Laser Polarimetry

All eyes were imaged with the GDx-VCC (software version 5.5.1, Carl Zeiss Inc., Dublin, California). The spherical equivalent of the refractive error of each eye was entered into the instrument prior to scanning. All participants were familiarized with the testing procedures prior to actual testing. Testing was performed with pupils undilated. Only good quality scans (quality score 7, evenly illuminated images with perfectly centered optic disc, without

motion artifacts, and no atypical birefringence pattern of RNFL thickness pattern in the four quadrants) were retained for analysis. From all GDx-VCC output measures provided, only the nerve fiber indicator (NFI) and the Temporal-Superior-Nasal-Inferior-Temporal (TSNIT) were considered for statistical analysis in the present study. The NFI is a computer-driven vector based on an advanced neural network algorithm and trained to differentiate normal from glaucomatous eyes. Potential NFI scores range from 0 to 100. The TSNIT average is a summary measure based on RNFL thickness values within the calculation circle around the ONH. It is automatically compared to the normative database and is quantified in terms of probability of normality. Normal values are displayed in green, abnormal values are color-coded based on their probability of normality so that dark blue indicates a 5% likelihood of being normal, light blue indicates the 2% level, yellow 1%, and red 0.5%. Patients with poor quality images (poorly or unevenly illuminated reflectance, quality score < 7, or atypical retardation pattern were excluded). Scans with normal or atypical retardation pattern were identified by visual inspection. Images with normal retardation pattern were defined as those with retardation maps with highest retardation superiorly and inferiorly indicating thicker RNFL and low retardation nasally and temporally indicating thinner RNFL. Images with abnormal retardation pattern were defined as those having alternating areas of low and high retardation arranged in a spoke-like manner in the peripapillary area, or those with high retardation in the nasal and temporal sectors.

### Statistical Analysis

The study sample population was divided into two thirds for modeling and one thirds for validation. This was done to account for the over-fitting principle in modeling in which evaluating an index with the same sample used to create it may result in overestimation of that index efficacy. This was achieved by stratification so that all four groups (normal, mild, moderate, and severe glaucoma subjects) were balanced in both the modeling and validation sets. This procedure yielded 150 cases (76 normal, 33 early glaucoma, 24 moderate glaucoma, and 17 severe glaucoma) with which to build a discrimination model. Univariable and multivariable prediction models were generated using simple and backward selection logistic regression, respectively, with glaucoma status as the outcome variable and NFI, TSNIT, number of abnormal points on FDT (NAP-FDT), and normal or abnormal FDT as candidate variables. A linear predictor score was created from the logistic regression coefficients and then submitted to a receiver operating characteristic (ROC) curve analysis, after which linear predictor scores were calculated for the cases in the validation set comprising 34 normal subjects and 40 patients with glaucoma (22 early, 8 moderate, 10 severe). The logistic regression models provide predicted probabilities of glaucoma status based on the estimated model parameters. The prediction formula for an individual is given as

$$p = \frac{e^{\beta_0 + \beta_1 \chi_1 + \beta_2 \chi_2 + \beta_3 \chi_1 \chi_2}}{1 + e^{\beta_0 + \beta_1 \chi_1 + \beta_2 \chi_2 + \beta_3 \chi_1 \chi_2}}$$

Where  $\chi_1$  and  $\chi_2$  are the individual's GDx-TSNIT and NAP-FDT measurements, respectively,  $\chi_1 \chi_2$  is the product of the two measurements (interaction), and  $\beta_0$ ,  $\beta_1$  and  $\beta_2$  are the logistic regression model parameters which are unknown but estimated using the data.

Akaike's information criterion (AIC), area under the curve (AUC) of the ROC, sensitivity, and specificity were determined for both univariable and multivariable models. In addition, the proportion of correctly classified subjects and the median prediction interval length (PIL) were determined in the validation set. All statistical analyses were performed with SAS version 9.2 (SAS, Cary, NC, USA). A  $P$  value  $< 0.05$  was considered statistically significant.

## RESULTS

### Demographic and Clinical Characteristics of Participants

A total of 124 normal and 155 glaucomatous eyes were initially enrolled in the study. After excluding 14 normal eyes (1 for unreliable FDT, 4 for GDx with low quality score, 9 for abnormal retardation pattern on GDx) and 41 glaucomatous eyes (21 for unreliable FDTs, 5 for GDx scans with quality score, 15 for GDx scans with abnormal retardation pattern), 110 normal and 114 glaucomatous eyes (55 mild, 32 moderate, and 27 severe) were available for analysis. The right eye was selected as study eye 55.4% of the time (124/224), and the left eye 44.6% of the time (100/224). Females represented 61.2% (137/224) of subjects. Unpaired Student  $t$ -test comparisons revealed that normal and glaucomatous eyes differed significantly with regard to age ( $58.0 \pm 10.7$  years vs.  $69.1 \pm 12.2$  years), CDR ( $0.3 \pm 0.1$  vs.  $0.7 \pm 0.2$ ), GDx NFI ( $20.7 \pm 9.7$  vs.  $54.0 \pm 24.6$ ), GDx-TSNIT ( $52.1 \pm 5.5$   $\mu\text{m}$  vs.  $40.2 \pm 8.2$   $\mu\text{m}$ ), and the NAP-FDT ( $1 \pm 3$  vs.  $8 \pm 5$ ) (all  $P < 0.001$ , unpaired Student  $t$ -test, means and standard deviations presented), but not GDx scan quality score ( $8.6 \pm 0.7$  vs.  $8.4 \pm 0.9$ ,  $P = 0.06$ ). Analysis of variance showed that eyes with mild glaucoma significantly differed from those with moderate and severe glaucoma in CDR ( $P = 0.023$  and  $0.008$ ), GDx-NFI (both  $P < 0.001$ ), GDx-TSNIT (both  $P < 0.001$ ), and NAP-FDT (both  $P < 0.001$ ). The comparison between moderately and severely affected eyes did not reach significance levels in any of the parameters ( $P = 0.17 - 0.88$ ). As a sensitivity analysis, we refit each of the statistical models while controlling for age as a continuous variable.

### Glaucoma Detection Using Single and Combined Devices

The results of both the univariable and multivariable models are shown in Table 1 and Table 2. Controlling for age did not affect the outcomes of the analyses as the age-adjusted and unadjusted results differed only slightly. As a result, we choose to present the unadjusted results. For diagnosing glaucoma using a single instrument, GDx-NFI provided the best single variable model for discriminating normal subjects and subjects with glaucoma regardless of severity (AIC: 120.88, AUC: 0.914, sensitivity: 87.8%, specificity: 84.2%), early glaucoma (AIC: 102.26, AUC: 0.872, sensitivity: 87.9%, specificity: 80.3%), and moderate glaucoma (AIC: 44.16, AUC: 0.961, sensitivity: 95.8%, specificity: 92.1%). NAP-FDT was the best model for severe glaucoma (AIC: 36.48, AUC: 0.976, sensitivity: 94.1%, specificity: 97.4%). The backward selection multivariable fitting procedure consistently identified the combination GDx-TSNIT, NAP-FDT, and the interaction GDx-TSNIT \* NAP-FDT as the best discriminating model between normal controls and glaucoma patients regardless of severity (AIC: 88.28, AUC: 0.959, sensitivity: 94.6%, specificity: 89.5%) and between normal subjects and subjects with mild (AIC: 76.06, AUC: 0.930, sensitivity: 87.9%, specificity: 89.5%), moderate (AIC: 39.37; AUC: 0.983, sensitivity: 100%,

specificity: 88.2%), and severe glaucoma (AIC: 30.62, AUC: 0.986, sensitivity: 100%, specificity: 92.1%). This multivariable model combining GDx and FDT parameters always outperformed the best single variable models (Table 1, Table2, and Figure). To determine whether multivariable models including GDx-NFI were better than the model based on GDx-TSNIT, both stepwise and forward selection logistic regression analyses were performed. The performance of the model combining GDx-TSNIT, NAP-FDT, and interaction GDx-TSNIT \* NAP-FDT was consistently slightly better than the combination GDx-NFI, NAP-FDT, and interaction GDx-NFI \* NAP-FDT (AIC: 90.79 for glaucoma regardless of severity and 79.63 for early glaucoma) and the combination GDx-NFI and NAP-FDT (AIC: 40.25 for moderate glaucoma and 32.53 for severe glaucoma).

Table 3 displays the estimates for the glaucoma regardless of severity, early, moderate, and severe glaucoma models. Along with a predicted probability of glaucoma, the logistic regression model also provides a 95% confidence interval for the prediction (prediction interval). We prefer models with shorter prediction intervals since models that produce large intervals indicate that there is a large amount of uncertainty associated with the predictions.

## DISCUSSION

The issue of implementing population screening has constantly been hampered by controversy over cost effectiveness and the lack of an ideal screening test.<sup>21, 22</sup> Since the approach based on detecting early cases is costly and time-consuming for population-based screening, a different approach would be to target only people with moderate to advanced glaucoma. This would be advantageous in terms of reducing the burden of false positive cases that might potentially overwhelm the healthcare system. The use of FDT and GDx in this study conforms with the philosophy of the World Glaucoma Association,<sup>23</sup> according to which an ideal screening test for OAG should be safe, easy to administer and interpret, portable, quick, acceptable to the people who are tested, able to obtain results in the majority of tested individuals and sufficiently valid to distinguish between those who do and those who do not have OAG. Both FDT and GDx-VCC fulfill these criteria and may be good candidates for population screening of glaucoma.

The concept of combining structural and functional tests to diagnose glaucoma and monitor its progression is based on the fact that glaucoma is characterized by structural and functional damage. The usefulness of this approach has been the subject of several studies.<sup>7-9, 24-27</sup> Only a very limited number of studies have assessed the diagnostic capability of combined GDx-VCC and FDT.<sup>8, 9</sup> Methodologically, the uniqueness of the present study lies in 1) the use of a modeling set and a validation set (known to be a more objective way of measuring the performance of various models that have been fit to the training set), 2) the use of the absolute values of NFI and TSNIT and the NAP-FDT as candidate variables in the multivariable analysis, and 3) the use of AIC and PIL as diagnostic performance measures for the first time, in addition to sensitivity, specificity, and AUC.

The results of our study show that the multivariable model combining GDx-TSNIT, NAPFDT, and the interaction GDX-TSNIT \* NAP-FDT consistently outperformed the best

single variable models of either GDx or FDT alone for discriminating normal and glaucoma status regardless of glaucoma severity or normal and any stage of the disease. Heeg et al.<sup>9</sup> compared the performance of FDT total deviation probability plot of the full mode when used alone and in combination with GDx-NFI with a cutoff point >29 in 237 normal and 452 glaucomatous individuals. This combination resulted in a sensitivity decrease from 90% to 83% and a specificity gain from 86% to 94%. The combined sensitivity increased to 99% after excluding patients with early glaucoma. In the study by Horn and colleagues,<sup>8</sup> FDT screening protocol (C-20-5) and GDx-NFI used individually in 252 patients had sensitivities of 85% and 64%, respectively, at a predefined specificity of 95%. When the two devices were combined, the sensitivity increased to 92%. In pre-perimetric glaucoma, FDT and GDx had the same sensitivity of 25%, which significantly increased to 44% after combination of the two tests. Similarly, Shah et al.<sup>7</sup> observed a significant increase (from 42% to 63%) in GDx-NFI sensitivity with negligible decrease in specificity (from 98% to 97%) following combination with FDT-PSD from N-30 thresholding protocol. Toth et al.<sup>27</sup> performed glaucoma screening using GDx-VCC and Matrix FDT (MFDT). Used individually, GDx NFI showed a low sensitivity of 26%, a specificity of 97%, and an AUC of 0.89. Combining abnormal NFI with MFDT screening mode (>2 points with  $P < 0.05$ ) further significantly decreased the sensitivity to 12% with corresponding specificity of 100% and AUC of 0.91. Alternatively, a combination of MFDT (1 point with  $P < 5\%$ ), NFI, and nerve fiber bundle defect on the GDx-VCC deviation map, with at least two of these parameters being abnormal, only increased the sensitivity to 42%. Although there are discrepancies in diagnostic performance measures that may mostly be ascribed to differences in methodology (i.e. NFI cutoffs, FDT protocol) and study participants (i.e. sample size, disease severity studied), there is a general agreement that combining FDT and GDx-VCC improves the glaucoma diagnostic performance of either device used individually.

We also compared the performance of the multivariable model to that of CDR as estimated by ophthalmoscopy (Tables 1 and 2) based on the fact that CDR is not a good metric for detecting glaucoma, particularly in early stages. Based on the AIC analysis, it is clear that the multivariable model outperforms the CDR model for “All Cases”, “Moderate Cases”, and “Severe Cases”. For “Early Cases”, the AIC values are more similar though the multivariable model value is still lower. Interestingly, the AUC values are always higher for the multivariable model as well. Although the confidence intervals overlap, the estimated sensitivities at fixed levels of specificity are also larger in each case for the multivariable model.

The interactive effect of GDx-TSNIT and FDT-NAP reported herein set this study apart from prior studies. From the practical standpoint, this finding signifies for example that while a moderately low GDx-TSNIT and moderately high FDT-NAP may individually have glaucoma diagnostic value, beyond these simple effects, the multiplicative GDx-TSNIT \* NAP-FDT effect of both increases the probability of glaucoma. It is important to note that clinicians often look for interaction between variables mentally when deciding if glaucoma is present, and give “extra value” when multiple streams of data tend to suggest the presence of glaucoma. Thus, this interaction is clinically relevant because it increases the interpretability of test results.

Akaike's information criterion (AIC) is an indicator of the goodness of fit of a proposed model. It balances model fit with complexity by penalizing models with an increased number of parameters, discouraging overfitting. Models with lower AIC values are preferred with differences of four or more typically taken to be meaningful. AIC values can only be compared across models fit to the same dataset and do not indicate that any of the proposed models are necessarily adequate, as only comparisons between the models are relevant. Use of AIC is commonly used to determine the preferred model among a set of candidate models.<sup>28</sup> The PIL is another useful measure in comparing the predictive ability of competing models and should not be used without also considering the other measures (AIC, AUC, sensitivity, specificity). A predictive model that is estimating probabilities near 0.5 will likely have a larger confidence interval associated with the estimate than a model that is estimating probabilities closer to 0 or 1. Models that are able to clearly distinguish between people with and without glaucoma are preferable. A prediction near 0.5 is clearly not as informative as a prediction closer to 0 or to 1. The PIL is able to penalize these models that are unable to clearly differentiate these people. If a model is providing predicted probabilities near 0 or 1 and these values are incorrect, this will be reflected in the other measures such as AIC, AUC, sensitivity, and specificity. Also, for two models that produce similar predicted probabilities, the PIL lets us know which model is predicting the probabilities with less uncertainty. We would then prefer the model with the smallest PIL, since both predictions are essentially equal. For these reasons, we suggest using PIL as a tool in conjunction with the other measures.

Because detecting glaucoma at an early stage is critical to delay the progression of structural and functional damage, a good diagnostic test should be highly sensitive and specific not only for moderate to severe, but for early disease as well. Logically, a combined structure-function approach would be more indicated for the detection of early glaucoma, for which establishing a reference standard remains difficult, since in some people, optic disc damage precedes VF, whereas in others it is the other way around.<sup>1</sup> While the search for an ideal method combining structural and functional tests for glaucoma population screening continues, our results suggest that the combination proposed here may be particularly useful for detecting moderate to severe glaucoma. Interestingly, the results also show that the model proposed herein is also suitable for detection of all stages of glaucoma. As noted in Table 1, the optimal specificities are located around 90% for each multivariable model and each type of glaucoma (all: 89.5%, early: 89.5%, moderate: 88.2%, severe: 92.1%). The associated sensitivities are also shown to be high in comparison with the other models as 3 out of the 4 are at least 94.6% and two of those are 100%. This suggests that at the optimal performance level of the multivariable models, the specificity and sensitivities are reasonably high. This represents an improvement over the optimal performance of the single variable models which often have optimal specificities much lower than 90% or in the case when specificity is high, very low sensitivities. Thus, the multivariable models balance the need for high sensitivity and specificity and represent an improvement over the single variable models in this regard.

This study has some limitations. First, it is a hospital-based study that was performed in a case-control manner, with the diagnosis of glaucoma based on typical glaucomatous ONH changes and supporting characteristic VF defects whereas normal subjects were required to



have normal IOP, normal looking ONH, and normal VF. Also, normal subjects were younger than subjects with glaucoma, and would be less likely to have lens opacities that might create non-glaucomatous FDT defects. Therefore, the results herein presented cannot directly be extrapolated to the general population; however, they provide valuable information of what to expect when combining the results of FDT and GDx-VCC TSNIT. These two instruments are portable, have short testing time, are easy for the patient to understand, are easy to administer and interpret by a non-physician operator, and do not require pupil dilation. These characteristics, which are not specific to these two devices, make their concurrent use one of the alternatives among possible combinations of glaucoma diagnostic devices. Second, we did not use the GDx-VCC typical scan score (TSS) as a quantitative measure of discriminating between scans with normal and abnormal retardation pattern.<sup>29, 30</sup> The TSS ranges from 0 (atypical retardation) to 100 (very typical retardation) and is derived from a support vector machine. Since glaucoma patients with abnormal retardation pattern are likely to be wrongly classified as false negative, abnormal retardation in glaucoma patients decreases the accuracy of the NFI as a result of the decrease of typical scan score (TSS) accuracy. The fact that the NFI had good AUCs in our study may indirectly indicate that TSSs were also high, particularly in the typical to very typical range. Third, although GDx provides more than 10 parameters, the NFI has been consistently shown to yield the best glaucoma diagnostic performance. The TSNIT average has also been shown to perform better than most of the other parameters, such as temporal and nasal averages, superior and inferior ratios, maximum modulation, superior and inferior maximum, ellipse modulation, temporal to nasal and inferior to nasal ratios, total, superior and inferior integrals. For these reasons, we chose to include only the NFI and TSNIT average in the assessment. Whether including the other GDx parameters would have yielded better performances than reported in the manuscript is doubtful. Interestingly, our model with these two parameters generated encouraging results. Fourth, we also acknowledge that the GDx-VCC has been outdated by the GDx-ECC. The GDx-VCC may exhibit an atypical retardation pattern that sometimes makes interpretation of the results difficult, as a result of poor signal-to-noise ratio. In contrast, the GDx-ECC is an improvement relative to the VCC version, with improved signal-to-noise ratio and introduction of a large birefringence bias that shifts the measurement of total retardation into a higher value region, ultimately resulting in better polarimetric image analysis, sensitivity, and specificity. However, it is important to bear in mind that for the advent of the GDx-ECC did not take away the good diagnostic value of the GDxVCC.

In conclusion, combining FDT C-20-5 protocol and GDx-VCC improves glaucoma detection in comparison with GDx or FDT used alone. The combination of GDx-TSNIT, NAP-FDT, and their interaction provides the best glaucoma discriminating model.

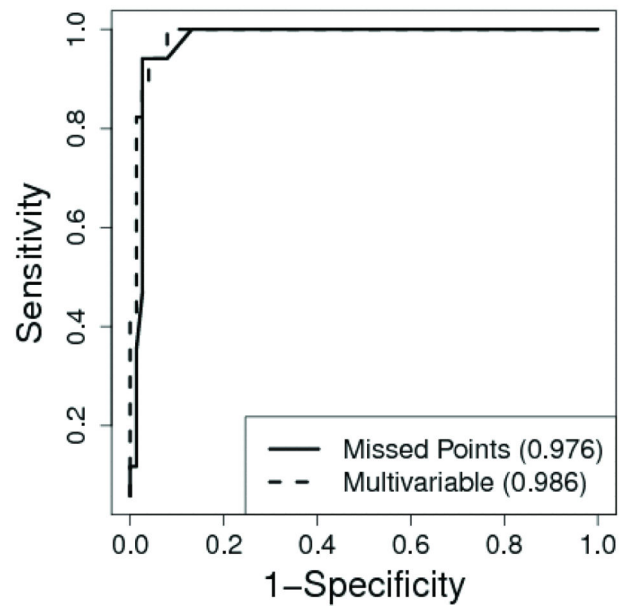
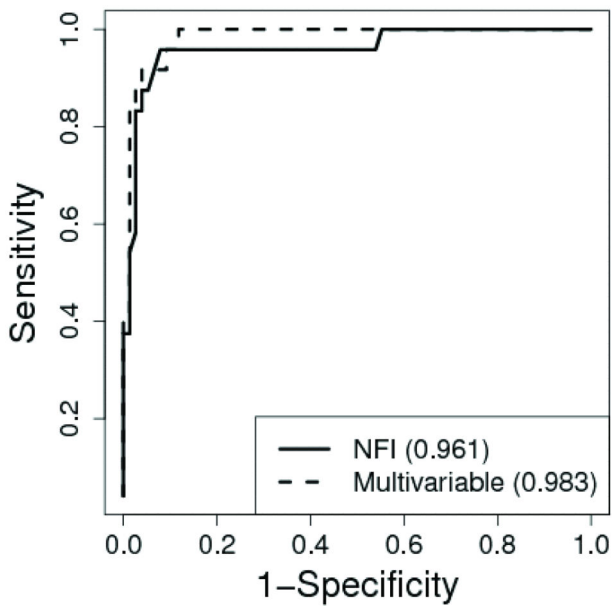
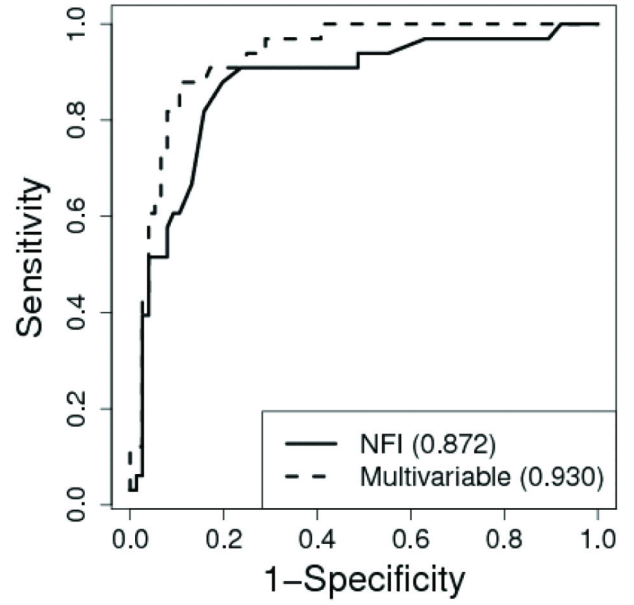
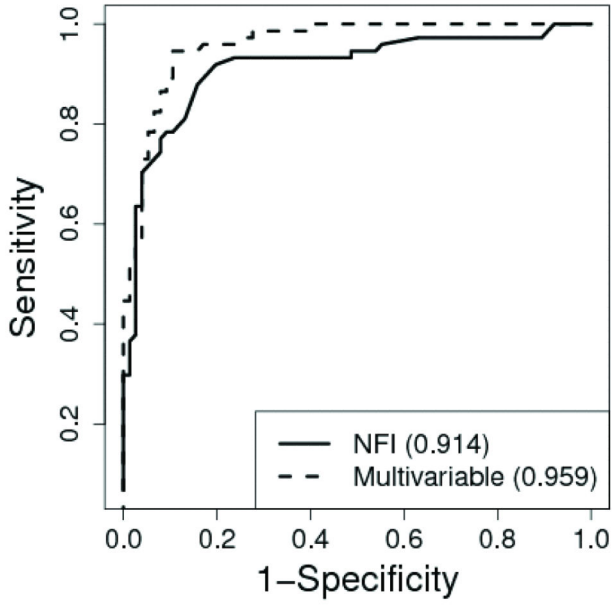
## Acknowledgments

Financial Support: Supported by an unrestricted grant from Research to Prevent Blindness, Inc., New York, NY, an unrestricted grant from Carl Zeiss Meditec, Inc., Dublin, CA, and the Heed Fellowship.

## REFERENCES

1. Kass MA, Heuer DK, Higginbotham EJ, et al. The Ocular Hypertension Treatment Study: a randomized trial determines that topical ocular hypotensive medication delays or prevents the onset of primary open-angle glaucoma. *Arch Ophthalmol.* 2002; 120:701–713. [PubMed: 12049574]
2. Leske MC, Hyman L, Hussein M, Heijl A, Bengtsson B. Comparison of glaucomatous progression between untreated patients with normal-tension glaucoma and patients with therapeutically reduced intraocular pressures. The effectiveness of intraocular pressure reduction in the treatment of normal-tension glaucoma. *Am J Ophthalmol.* 1999; 127:625–626. [PubMed: 10334369]
3. Quigley HA, West SK, Rodriguez J, Munoz B, Klein R, Snyder R. The prevalence of glaucoma in a population-based study of Hispanic subjects: Proyecto VER. *Arch Ophthalmol.* 2001; 119:1819–1826. [PubMed: 11735794]
4. Tielsch JM, Katz J, Singh K, et al. A population-based evaluation of glaucoma screening: the Baltimore Eye Survey. *Am J Epidemiol.* 1991; 134:1102–1110. [PubMed: 1746520]
5. McKendrick AM. Recent developments in perimetry: test stimuli and procedures. *Clin Exp Optom.* 2005; 88:73–80. [PubMed: 15807638]
6. Townsend KA, Wollstein G, Schuman JS. Imaging of the retinal nerve fibre layer for glaucoma. *Br J Ophthalmol.* 2009; 93:139–143.
7. Shah NN, Bowd C, Medeiros FA, et al. Combining structural and functional testing for detection of glaucoma. *Ophthalmology.* 2006; 113:1593–1602. [PubMed: 16949444]
8. Horn FK, Nguyen NX, Mardin CY, Junemann AG. Combined use of frequency doubling perimetry and polarimetric measurements of retinal nerve fiber layer in glaucoma detection. *Am J Ophthalmol.* 2003; 135:160–168. [PubMed: 12566019]
9. Heeg GP, Stoutenbeek R, Jansonius NM. Strategies for improving the diagnostic specificity of the frequency doubling perimeter. *Acta Ophthalmol Scand.* 2005; 83:53–56. [PubMed: 15715557]
10. Reus NJ, Colen TP, Lemij HG. Visualization of localized retinal nerve fiber layer defects with the GDx with individualized and with fixed compensation of anterior segment birefringence. *Ophthalmology.* 2003; 110:1512–1516. [PubMed: 12917165]
11. Weinreb RN, Bowd C, Zangwill LM. Glaucoma detection using scanning laser polarimetry with variable corneal polarization compensation. *Arch Ophthalmol.* 2002; 120:218–224. [PubMed: 11831931]
12. Tannenbaum DP, Hoffman D, Lemij HG, Garway-Heath DF, Greenfield DS, Caprioli J. Variable corneal compensation improves discrimination between normal and glaucomatous eyes with the scanning laser polarimeter. *Ophthalmology.* 2004; 111:259–264. [PubMed: 15019373]
13. Burnstein Y, Elish NJ, Magbalon M, Higginbotham EJ. Comparison of frequency doubling perimetry with Humphrey visual field analysis in a glaucoma practice. *Am J Ophthalmol.* 2000; 129:328–333. [PubMed: 10704548]
14. Landers J, Sharma A, Goldberg I, Graham S. Topography of the frequency doubling perimetry visual field compared with that of short wavelength and achromatic automated perimetry visual fields. *Br J Ophthalmol.* 2006; 90:70–74. [PubMed: 16361671]
15. Medeiros FA, Sample PA, Weinreb RN. Frequency doubling technology perimetry abnormalities as predictors of glaucomatous visual field loss. *Am J Ophthalmol.* 2004; 137:863–871. [PubMed: 15126151]
16. Khong JJ, Dimitrov PN, Rait J, McCarty CA. Can the specificity of the FDT for glaucoma be improved by confirming abnormal results? *J Glaucoma.* 2001; 10:199–202. [PubMed: 11442182]
17. Mansberger SL, Johnson CA, Cioffi GA, et al. Predictive value of frequency doubling technology perimetry for detecting glaucoma in a developing country. *J Glaucoma.* 2005; 14:128–134. [PubMed: 15741814]
18. Hodapp, E.; Parrish, R.R.; Anderson, D.R. *Clinical decisions in glaucoma.* CV Mosby; St. Louis: 1993.
19. Anderson AJ, Johnson CA. Frequency-doubling technology perimetry and optical defocus. *Invest Ophthalmol Vis Sci.* 2003; 44:4147–4152. [PubMed: 12939339]
20. Anderson AJ, Johnson CA. Frequency-doubling technology perimetry. *Ophthalmol Clin North Am.* 2003; 16:213–225. [PubMed: 12809159]

21. Hernandez RA, Burr JM, Vale LD. Economic evaluation of screening for open-angle glaucoma. *Int J Technol Assess Health Care*. 2008; 24:203–211. [PubMed: 18400124]
22. Burr JM, Mowatt G, Hernandez R. The clinical effectiveness and cost-effectiveness of screening for open angle glaucoma: a systematic review and economic evaluation. *Health Technol Assess*. 2007; 11:1–190.
23. Weinreb, RN.; Healey, PR.; Topouzis, F. *Glaucoma screening*. Kugler Publications; Amsterdam: 2008.
24. Bowd C, Hao J, Tavares IM, et al. Bayesian machine learning classifiers for combining structural and functional measurements to classify healthy and glaucomatous eyes. *Invest Ophthalmol Vis Sci*. 2008; 49:945–953. [PubMed: 18326717]
25. Medeiros FA, Lisboa R, Weinreb RN, Girkin CA, Liebmann JM, Zangwill LM. A combined index of structure and function for staging glaucomatous damage. *Arch Ophthalmol*. 2012; 130:1107–1116. [PubMed: 23130365]
26. Robin TA, Muller A, Rait J, Keeffe JE, Taylor HR, Mukesh BN. Performance of community-based glaucoma screening using Frequency Doubling Technology and Heidelberg Retinal Tomography. *Ophthalmic Epidemiol*. 2005; 12:167–178. [PubMed: 16036475]
27. Toth M, Kothy P, Hollo G. Accuracy of scanning laser polarimetry, scanning laser tomography, and their combination in a glaucoma screening trial. *J Glaucoma*. 2008; 17:639–646. [PubMed: 19092459]
28. Akaike, H. Information theory and an extension of the maximum likelihood principle. In: Petrov, BH.; Czaki, F., editors. *2nd International Symposium on Information Theory*; Budapest. Akademiai Kiadó; 1973. p. 267-281.
29. Bagga H, Greenfield DS, Feuer WJ. Quantitative assessment of atypical birefringence images using scanning laser polarimetry with variable corneal compensation. *Am J Ophthalmol*. 2005; 139:437–446. [PubMed: 15767051]
30. Schrems WA, Laemmer R, Hoesl LM, et al. Influence of atypical retardation pattern on the peripapillary retinal nerve fibre distribution assessed by scanning laser polarimetry and optical coherence tomography. *Br J Ophthalmol*. 2011; 95:1437–1441. [PubMed: 21242583]



Receiver operating characteristic (ROC) curves of FDT and GDx-VCC used individually and in combination for detection glaucoma regardless of severity (top left), early glaucoma (top right), moderate glaucoma (bottom left), and severe glaucoma (bottom right).

**Table 1**

Glaucoma Status Prediction with Univariable and Multivariable Models of GDx and FDT

Glaucoma group	Model	AIC	AUC	Sensitivity	Specificity	CC*	PIL*
All	NFI**	120.88	0.914	87.8	84.2	86.5	0.180
	TSNIT	126.58	0.902	79.7	93.4	82.4	0.189
	NAP-FDT	129.12	0.881	83.8	86.8	78.4	0.165
	FDT	129.00	0.853	83.8	86.8	78.4	0.163
	CDR	102.90	0.931	93.2	81.6	85.1	0.137
	Multivariable	88.28	0.959	94.6	89.5	83.8	0.147
Early	NFI**	102.26	0.872	87.9	80.3	80.4	0.178
	TSNIT	101.40	0.854	72.7	86.8	76.8	0.199
	NAP-FDT	109.98	0.796	69.7	86.8	71.4	0.162
	FDT	103.67	0.783	69.7	86.8	71.4	0.155
	CDR	79.32	0.908	87.9	81.6	80.4	0.186
	Multivariable	76.06	0.930	87.9	89.5	82.1	0.240
Moderate	NFI**	44.16	0.961	95.8	92.1	95.2	0.105
	TSNIT	49.49	0.946	91.7	94.7	92.9	0.124
	NAP-FDT	55.56	0.929	91.7	92.1	81.0	0.104
	FDT	61.80	0.893	91.7	86.8	71.4	0.155
	CDR	50.99	0.936	91.7	90.8	92.9	0.112
	Multivariable	39.37	0.983	100	88.2	83.3	0.125
Severe	NFI	39.14	0.931	88.2	94.7	100	0.106
	TSNIT	44.12	0.932	94.1	93.4	93.2	0.134
	NAP-FDT**	36.48	0.976	94.1	97.4	88.6	0.079
	FDT	38.07	0.934	100	86.8	77.3	0.110
	CDR	37.13	0.968	100	81.6	84.1	0.122
	Multivariable	30.62	0.986	100	92.1	88.6	0.121

AIC, Akaike's information criterion; AUC, area under the curve; CC, proportion subjects correctly classified; PIL, median prediction interval length; NFI, nerve fiber indicator; TSNIT, average RNFL; NAP-FDT, number of abnormal points on FDT; FDT, normal or abnormal; CDR, cup-to-disc ratio.

\* Calculated using the validation set;

\*\* Best single variable model

**Table 2**

Estimated Sensitivities at Fixed Specificity Levels with Confidence Intervals

<b>Glaucoma group</b>	<b>Model</b>	<b>Sensitivity (90% Spec.)</b>	<b>95% CI</b>	<b>Sensitivity (95% Spec.)</b>	<b>95% CI</b>
All	NFI	78.4	(60.4, 89.6)	71.6	(49.7, 86.6)
	TSNIT	79.7	(62.0, 90.4)	71.6	(49.7, 86.6)
	NAP-FDT	82.4	(65.4, 92.1)	60.8	(38.8, 79.2)
	FDT	--	--	--	--
	CDR	83.8	(67.1, 92.9)	66.2	(44.0, 83.0)
	Multivariable	89.2	(74.4, 95.9)	73.0	(51.1, 87.4)
Early	NFI	60.6	(37.7, 79.6)	51.5	(27.7, 74.7)
	TSNIT	63.6	(40.5, 81.8)	54.5	(30.1, 76.9)
	NAP-FDT	60.6	(37.7, 79.6)	33.3	(14.7, 59.1)
	FDT	--	--	--	--
	CDR	75.8	(52.7, 89.8)	48.5	(25.3, 72.3)
	Multivariable	81.8	(59.6, 93.2)	60.6	(35.3, 81.3)
Moderate	NFI	95.8	(72.3, 99.5)	87.5	(60.9, 96.9)
	TSNIT	91.7	(68.1, 98.3)	87.5	(60.9, 96.9)
	NAP-FDT	91.7	(68.1, 98.3)	79.2	(51.2, 93.2)
	FDT	--	--	--	--
	CDR	91.7	(68.1, 98.3)	83.3	(55.9, 95.2)
	Multivariable	95.8	(72.3, 99.5)	91.7	(66.0, 98.4)
Severe	NFI	88.2	(59.1, 97.5)	88.2	(57.1, 97.7)
	TSNIT	94.1	(64.4, 99.3)	82.4	(50.8, 95.5)
	NAP-FDT	100	--	94.1	(62.6, 99.4)
	FDT	--	--	--	--
	CDR	88.2	(59.1, 97.5)	76.5	(44.9, 92.8)
	Multivariable	100	--	94.1	(62.6, 99.4)

AUC, area under the curve; NFI, nerve fiber indicator; TSNIT, average RNFL; NAP-FDT, number of abnormal points on FDT; FDT, normal or abnormal; CDR, cup-to-disc ratio.

**Table 3**

Glaucoma Multiple Logistic Regression Results Using the Modeling Set

Glaucoma group	Parameter	Estimate	P-Value
All	Intercept ( $\beta_0$ )	18.29	<0.0001
	GDx-TSNIT ( $\beta_1$ )	-0.41	<0.0001
	NAP-FDT ( $\beta_2$ )	-1.18	0.0012
	GDx-TSNIT*NAP-FDT ( $\beta_3$ )	0.03	0.0001
Early	Intercept ( $\beta_0$ )	18.67	0.0001
	GDx-TSNIT ( $\beta_1$ )	-0.42	<0.0001
	NAP-FDT ( $\beta_2$ )	-1.80	0.0006
	GDx-TSNIT*NAP-FDT ( $\beta_3$ )	0.05	0.0002
Moderate	Intercept ( $\beta_0$ )	16.49	0.0045
	GDx-TSNIT ( $\beta_1$ )	-0.42	0.0015
	NAP-FDT ( $\beta_2$ )	-1.00	0.0627
	GDx-TSNIT*NAP-FDT ( $\beta_3$ )	0.03	0.0233
Severe	Intercept ( $\beta_0$ )	14.85	0.0448
	GDx-TSNIT ( $\beta_1$ )	-0.40	0.0174
	NAP-FDT ( $\beta_2$ )	-0.89	0.1200
	GDx-TSNIT*NAP-FDT ( $\beta_3$ )	0.03	0.0360

TSNIT, average RNFL; NAP-FDT, number of abnormal points on FDT