Using Optical Imaging Summary Data to Detect Glaucoma

Cesar Sanchez-Galeana, MD, Christopher Bowd, PhD, Eytan Z. Blumenthal, MD, Parag A. Gokhale, MD, Linda M. Zangwill, PhD, Robert N. Weinreb, MD

**Purpose:** To evaluate the sensitivity and specificity for discriminating between early to moderate glaucomatous and normal eyes using summary data reports from the Heidelberg Retina Tomograph (HRT), the GDx Nerve Fiber Analyzer (GDx), and the Optical Coherence Tomograph (OCT).

**Design:** Comparative cross-sectional study.

**Participants:** One eye each of 50 normal subjects and 39 glaucoma patients with early to moderate visual field damage (mean deviation, $-5.04 \pm 3.32$ dB; range, $-0.85$ to $-13.2$ dB).

**Methods:** Three experienced graders masked to patient identity and diagnosis evaluated each summary data report from the HRT, GDx, and OCT independently.

**Main Outcome Measures:** Each summary report was classified as either normal or glaucomatous. Sensitivity and specificity are reported for each grader, and agreement between graders is reported.

**Results:** For the HRT, sensitivity and specificity ranged from 64% to 75% and 68% to 80%, respectively. Agreement ($\kappa \pm$ standard error [SE]) between observers one and two, two and three, and one and three was 0.73 $\pm$ 0.07, 0.77 $\pm$ 0.07, and 0.67 $\pm$ 0.08, respectively. For the GDx, sensitivity and specificity ranged from 72% to 82% and 56% to 82%, respectively. Agreement ($\kappa \pm$ SE) between observers one and two, two and three, and one and three was 0.66 $\pm$ 0.08, 0.66 $\pm$ 0.08, and 0.50 $\pm$ 0.09, respectively. For the OCT, sensitivity and specificity ranged from 76% to 79% and 68% to 81%, respectively. Agreement ($\kappa \pm$ SE) between observers one and two, two and three, and one and three was 0.73 $\pm$ 0.07, 0.58 $\pm$ 0.08, and 0.51 $\pm$ 0.09, respectively.

**Conclusions:** When used alone, HRT, GDx, and OCT summary data reports can differentiate between normal and glaucomatous eyes with mild to moderate visual field loss. However, none of the instruments provided sensitivity and specificity that justify summary data reports being used as a screening tool for early to moderate glaucoma. *Ophthalmology* 2001;108:1812–1818 © 2001 by the American Academy of Ophthalmology.

Several instruments have been introduced to quantitatively assess optic disc topography and/or the retinal nerve fiber layer (RNFL). Each of them relies on different properties of the retina to provide large amounts of descriptive information, which can be subsequently condensed into graphical summary data reports, including numeric indices for clinical interpretation. Compared with optic disc and RNFL stereophotographs, these instruments provide data that are objective, quantitative, highly reproducible, and do not require laborious processing.

Although there has been considerable research on the diagnostic sensitivity and specificity using these instruments, most of these studies have focused on objective quantitative parameters. The aim of this study was to compare the sensitivity and specificity for differentiating between normal and glaucomatous eyes based solely on the qualitative assessment of summary data reports from three commercially available instruments.

**Materials and Methods**

One randomly selected eye from each of 50 healthy and 39 glaucoma patients was included. All individuals underwent a complete eye examination, including slit-lamp biomicroscopy, gonioscopy, dilated funduscopy, stereophotography, and static automated perimetry.

Normal eyes were defined as those with intraocular pressure (IOP) $\leq 21$ mmHg with no prior history of elevated IOP, a best-corrected visual acuity of at least 20/40, and healthy optic disc appearance (asymmetry of vertical cup/disc ratio $<0.2$, no evidence of rim thinning, notching, excavation, hemorrhage, or RNFL defects). A glaucoma hemifield test and corrected pattern standard deviation (Humphrey Field Analyzer program 24-2, Humphrey Instruments, San Leandro, CA) within normal limits were required. Patients with a history of diabetes, ocular inflammation, trauma, or past intraocular surgery were excluded.

Glaucomatous eyes had at least two consecutive abnormal standard automated perimetry results defined as glaucoma hemifield test or corrected pattern standard deviation outside normal limits. Glaucomatous eyes had early to moderate visual field damage with mean ($\pm$ standard deviation [SD]) deviation of $-5.0 \pm 3.3$ dB (range, $-0.8$ to $-13.2$ dB).

Mean ($\pm$ SD) age was lower in the normal group ($57.8 \pm 12.0$ years; range, 42.4–81.6) than in the glaucoma group ($68.0 \pm 12.7$ years; range, 43.5–91.7) ($t$ test, $P < 0.001$). There were more females than males in the normal group (4 males and 36 females).

Originally received; December 6, 2000. Accepted: May 15, 2001. Manuscript no. 200851.

From the Glaucoma Center, University of California San Diego, La Jolla, California.

Reprint requests to Robert N. Weinreb, MD, Glaucoma Center, University of California San Diego, 9500 Gilman Drive, La Jolla, CA, 92093–9046.
Figure 1. Heidelberg Retina Tomograph initial examination report showing reflectance image, color topography image, topography cross-sections, and height variation of the retinal surface along the contour line. Stereometric parameters also are shown. Patient was diagnosed as glaucomatous by all observers.

Stereometric Analysis ONH:
- Disk Area: 1.881 mm²
- Cup Area: 1.022 mm²
- Cup/Disk Area Ratio: 0.543
- Rim Area: 0.859 mm²
- Cup Volume: 0.196 mm³
- Rim Volume: 0.102 mm³
- Mean Cup Depth: 0.218 mm
- Maximum Cup Depth: 0.459 mm

Stereometric Analysis ONH:
- Cup Shape Measure: -0.055
- Height Variation Contour: 0.167 mm
- Mean RNFL Thickness: 0.101 mm
- RNFL Cross Section Area: 0.490 mm²
- Classification: Glaucoma
- Reference Height (Std.): 0.279 mm

Software: IR1-V2.01


Figure 2. Heidelberg Retina Tomograph ranked segment analysis report showing cumulative frequency distributions of 32 measured segments of retinal nerve fiber layer (RNFL) thickness, rim/disc ratio, rim/disc area ratio, and rim volume for the same patient as shown in Figure 1. (*) Classification based on statistics; diagnosis is physician’s responsibility.
and in the glaucoma group (15 males and 24 females) (both chi-square test, $P < 0.001$).

All patients underwent ocular imaging with the Heidelberg Retina Tomograph (HRT; Heidelberg Engineering, Heidelberg, Germany), GDx Nerve Fiber Analyzer (Laser Diagnostic Technologies, San Diego, CA), and Optical Coherence Tomograph (OCT; Humphrey-Zeiss, Dublin, CA). For each subject, all ocular imaging and visual field tests were completed within 6 months. These instruments have been described in detail elsewhere. A brief overview of each instrument and the information available, known as summary data in the standard report, is described below.

Confocal Scanning Laser Ophthalmoscopy

The HRT provides topographic measures of the optic disc and peripapillary retina derived from 32 optical sections at consecutive focal depth planes. Each image consists of 256 × 256 pixels, with each pixel corresponding to retinal height at its location.

For this study, three 15° field of view scans centered on the optic disc judged to be of acceptable quality were obtained for each test eye. A mean topography image of these three scans was created with HRT software version 2.01 and used in all analyses. Experienced technicians outlined the margin of the optic disc while viewing stereo photographs.

To assess the images, HRT initial examination report version 1R1-V2.01 was reviewed by each observer (Fig 1). This report contains the color topography image, the reflectance image, one horizontal and one vertical cross-section of the topography, and the height variation of the retinal surface along the contour line. The report also contains the following stereometric parameters: disc area, cup area, cup/disc area ratio, rim area, rim volume, mean cup depth, maximum cup depth, cup shape measure, height variation contour, mean RNFL thickness, RNFL cross-section area, and the HRT classification. The HRT classification uses a discriminant function to classify an eye as glaucomatous or normal. In addition, the ranked segment analysis report was reviewed. This report contains the cumulative frequency distributions of 32 measured segments of RNFL, normalized graph, rim/disc ratio, rim/disc area ratio, and rim volume (Fig 2).

Scanning Laser Polarimetry

The scanning laser polarimeter (GDx Nerve Fiber Analyzer) uses scanning laser technology coupled with an integrated polarization modulator to provide a retardation map of the optic disc and peripapillary retina. This instrument measures retardation of light that has double-passed the birefringent fibers of the RNFL. Each resulting image consists of RNFL thickness measurements (in microns) how a patient’s RNFL measurements compare with normal for each of four quadrants. The nerve fiber analysis compares the patient’s thickness values with a database of age-matched and race-matched normal individuals and evaluates a series of ratios, averages, and other parameters. The RNFL parameters of the report include average thickness, volume, symmetry (superior quadrant thickness/inferior quadrant thickness), superior ratio (superior quadrant thickness/temporal quadrant thickness), inferior ratio (inferior quadrant thickness/temporal quadrant thickness), superior/nasal ratio, maximum modulation (thickest quadrant-thinnest quadrant/thinnest quadrant), superior maximum (average of thickest 1500 pixels in superior quadrant), inferior maximum, ellipse modulation, ellipse average, superior average, inferior average, normal integral. The GDx number is the result of analysis by a neural network that analyzes more than 200 parameters from a GDx image and assigns a number between 0 and 100, where 0 is normal and 100 is glaucoma.

Optical Coherence Tomography

The OCT 2000 uses low-coherence interferometry to assess peripapillary RNFL thickness. This instrument quantifies RNFL thickness by measuring the difference in temporal delay of back-scattered light from the RNFL and a reference mirror. RNFL is differentiated from other retinal layers with an algorithm that detects the anterior edge of retinal pigment epithelium and determines the photoreceptor layer position (software version ASX 1). Each resulting image consists of RNFL thickness measurements (in microns) at 100 points along a 360° circular (ring) path (one thickness value per 3.6°) around the optic disc.

For this study, three circular scans of 3.4 mm diameter centered on the optic disc judged to be of acceptable quality were obtained for each test eye. This approximate scan diameter was found to be optimal for RNFL analysis in a prototype instrument. Mean RNFL thicknesses for quadrant and clock-hour measurements were calculated from the three images obtained.

To assess the data, the OCT RNFL thickness report was reviewed (Fig 4). The report contains different color-coded layers of the retina including the RNFL and a black-and-white fundus image of the location of the scan. In addition, RNFL thickness is reported in each of four quadrants (superior, inferior, nasal, and temporal), in clock hours, and as average RNFL thickness.

Data Analysis

Three experienced graders, two ophthalmologists (PAG and EZB) and one vision scientist experienced with these imaging techniques (CB) independently evaluated each report masked to patient identity and diagnosis. A forced choice diagnosis of healthy versus glaucoma was made after reviewing each report based on a subjective clinical impression of all included information. The diagnosis was not based on any specific normal/abnormal criteria.

Graders used a standard form to document their impression of each report and to determine which information was used in the classification (healthy versus glaucoma). These forms contained questions tailored to each instrument report. The HRT form included the following items: (1) reflectance image (normal, focal defect, diffuse defect), (2) double-hump graph (normal, focal defect, diffuse defect), (3) location of defect (superior, inferior, nasal, temporal), (4) stereometric analysis (normal, glaucomatous, undetermined), and (5) ranked segment analysis (normal, glaucomatous, undetermined). The GDx analysis documented (1) the polarization map (normal, focal defect, diffuse defect), (2) double hump graph (normal, focal defect, diffuse defect), (3) location of defect (superior, inferior, nasal, temporal), and (4) nerve fiber analysis (normal, glaucoma, undetermined). The OCT form documented (1) the color map (normal, focal defect, diffuse defect), (2) double hump graph (normal, focal defect, diffuse defect), and (3) location of defect (superior, inferior, nasal, temporal). Each form concluded
Figure 3. GDx Nerve Fiber Analyzer extended analysis report showing reflectance image, polarization map, double-hump graph, deviation from normal table, and nerve fiber analysis parameter information for the same patient as shown in Figure 1.

Figure 4. Optical Coherence Tomograph retinal nerve fiber layer (RNFL) thickness report showing color-coded retinal thickness map, monochromatic fundus image, and RNFL thickness values by quadrant, clock-hour, and on average for the same patient as shown in Figure 1.
with a forced-choice final diagnosis (normal, glaucoma). Graders were instructed to base their subjective decision on all summary data conveyed on the form and not any strict diagnostic criteria, such as GDx number > 30 or HRT classification “glaucoma.” This diagnosis was then compared with the diagnosis based on standard visual field results, and percentage of correct diagnoses (glaucoma, normal) was compared. The sensitivity and specificity for each instrument was also calculated. Furthermore, sensitivity and specificity was calculated for each observer. Level of agreement between each observer was calculated using the $\kappa$ statistic analysis and the JMP (SAS Institute, Inc, Cary, NC) statistical software package.

### Results

The sensitivity and specificity for detecting glaucoma by each grader for each instrument is presented in Table 1. Agreement among graders is presented in Table 2. For differentiating between normal and glaucomatous eyes by qualitative assessment of the HRT initial examination reports, sensitivity was 64% and specificity was 68% for grader one, 75% and 80% for grader two, and 71% and 80% for grader three, respectively. Agreement ($\kappa \pm$ standard error [SE]) between observers one and two, two and three, and one and three was $0.73 \pm 0.07$, $0.77 \pm 0.07$, and $0.67 \pm 0.08$, respectively. These values are considered to indicate moderate to substantial agreement. Twenty-one glaucoma patients and 31 normal subjects were diagnosed correctly by all three observers. Eight glaucoma patients and six normal subjects were erroneously diagnosed as normal by all three observers. Using the HRT classification discriminant function (glaucoma versus normal), the sensitivity was 52% and the specificity was 87%. Agreement ($\kappa \pm$ SE) of the HRT classification with observers one, two, and three was moderate to substantial $0.58 \pm 0.08$, $0.64 \pm 0.08$, and $0.69 \pm 0.08$, respectively.

For qualitative assessment of GDx extended analysis reports, sensitivity was 74% and specificity was 71% for grader one, 72% and 82% for grader two, and 82% and 56% for grader three, respectively. Agreement ($\kappa \pm$ SE) between observers one and two, two and three, and one and three was $0.66 \pm 0.08$, $0.66 \pm 0.08$, and $0.50 \pm 0.09$, respectively. These values are considered to be moderate to substantial agreement. Twenty-four glaucomatous patients and 22 normal subjects were diagnosed correctly by all three observers. Four glaucoma patients and seven normal subjects were erroneously diagnosed by all three observers.

Using the GDx number neural network algorithm cutoff value of 35, the sensitivity was 64% and the specificity was 82%. Agreement ($\kappa \pm$ SE) between the GDx number and observers one, two, and three was $0.45 \pm 0.09$, $0.51 \pm 0.09$, and $0.39 \pm 0.08$, respectively. These values are considered to be fair to moderate agreement.

For qualitative assessment of OCT reports, sensitivity was 76% and specificity was 81% for grader one, 79% and 68% for grader two, and 76% and 68% for grader three, respectively. Agreement ($\kappa \pm$ SE) between observers one and two, two and three, and one and three was $0.73 \pm 0.07$, $0.58 \pm 0.08$, and $0.51 \pm 0.09$, respectively. These values are considered to be fair to moderate agreement. Twenty-five glaucomatous patients and 24 normal subjects were diagnosed correctly by the three observers. Five glaucoma patients and four normal subjects were erroneously diagnosed by all three observers.

We also investigated the sensitivity and specificity for consensus assessment of diagnostic group. When agreement between two of the three graders was required to diagnose eyes as either glaucoma or normal, sensitivity was 67.6%, 97.9%, and 71.4% and specificity was 79%, 71.4% and 80.4% for HRT, GDx, and OCT, respectively.

In addition, we evaluated which information from the instrument reports was used when three graders correctly and incorrectly diagnosed the subjects. In all cases of correctly diagnosed glaucoma, the double-hump graph of the GDx and OCT was graded as abnormal. In addition, in 85% of eyes the HRT stereometric analysis was graded as abnormal; in 83% of eyes the GDx nerve fiber analysis was graded as abnormal. When glaucoma patients were erroneously diagnosed as normal, graders recorded some evidence of focal defects, whereas HRT stereometric analysis, and ranked segment analysis, and GDx nerve fiber analysis were normal.

In the correctly diagnosed normal eyes, little evidence of focal defects was found, and HRT stereometric analysis and ranked segment analysis, and GDx nerve fiber analysis were normal. When normal subjects were incorrectly diagnosed with glaucoma, graders found evidence of glaucoma from the HRT stereometric

### Table 1. Sensitivity and Specificity for Assessment of HRT, GDx and OCT Summary Data Reports by Three Observers

<table>
<thead>
<tr>
<th></th>
<th>Observer 1</th>
<th>Observer 2</th>
<th>Observer 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivity (%)</td>
<td>Specificity (%)</td>
<td>Sensitivity (%)</td>
</tr>
<tr>
<td>HRT</td>
<td>64</td>
<td>68</td>
<td>75</td>
</tr>
<tr>
<td>GDx</td>
<td>74</td>
<td>71</td>
<td>72</td>
</tr>
<tr>
<td>OCT</td>
<td>76</td>
<td>81</td>
<td>79</td>
</tr>
</tbody>
</table>

GDx = GDx Nerve Fiber Analyzer; HRT = Heidelberg Retina Tomograph; OCT = Optical Coherence Tomograph.

### Table 2. Agreement Between Observers for Classifying Eyes as Normal or Glaucomatous (kappa ± standard error)

<table>
<thead>
<tr>
<th></th>
<th>Observer 1 and 2</th>
<th>Observer 2 and 3</th>
<th>Observer 1 and 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>HRT</td>
<td>$0.73 \pm 0.07$</td>
<td>$0.77 \pm 0.07$</td>
<td>$0.67 \pm 0.08$</td>
</tr>
<tr>
<td>GDx</td>
<td>$0.66 \pm 0.08$</td>
<td>$0.66 \pm 0.08$</td>
<td>$0.50 \pm 0.09$</td>
</tr>
<tr>
<td>OCT</td>
<td>$0.73 \pm 0.07$</td>
<td>$0.58 \pm 0.08$</td>
<td>$0.51 \pm 0.09$</td>
</tr>
</tbody>
</table>

GDx = GDx Nerve Fiber Analyzer; HRT = Heidelberg Retina Tomograph; OCT = Optical Coherence Tomograph.
The severity of glaucoma as measured by mean deviation was related to diagnostic accuracy. For patients correctly diagnosed as glaucomatous by all three graders, mean deviation ranged from −5.42 to −6.72 dB. For glaucoma patients incorrectly diagnosed as normal, mean deviation ranged from −2.64 to −2.77 dB. For subjects correctly diagnosed as normal, mean deviation ranged from −0.12 to −0.14 dB. Finally, mean deviation ranged from 0 to −1.10 dB in normal subjects who were incorrectly diagnosed as glaucoma.

Discussion

This study demonstrates moderate to good sensitivity and specificity for differentiating glaucomatous eyes from normal eyes, with summary data from HRT, GDx and OCT standard reports as the only source of clinical information. Moderate to substantial agreement was found among graders for classifying glaucoma and normal eyes. The classification included in the reports for HRT (discriminant analysis) and GDx (neural network number) had moderate to substantial agreement with the graders.

Mikelberg et al.17 originally reported a sensitivity of 87% and specificity of 84% for detecting glaucoma with the current HRT classification discriminant function, which is better than the 52% sensitivity and 87% specificity reported in this study. In other populations, sensitivities and specificities of the HRT classification ranging from 42% to 93% and 84% to 96%, respectively, have been reported.11–13

Tjon-Fo-Sang and Lemij14 detected high sensitivity (96%) and specificity (93%) with an earlier version of the GDx, the Nerve Fiber Analyzer I (NFA I). Their study selected patients with early to severe visual field defects; mean deviation was −10.33 dB (range, −31.5–0.76 dB). In contrast, we included only patients with early to moderate visual field defects (mean deviation, −5.04 ± 3.32; range, −0.85 to −13.2). A substantial number of subjects with severe visual field defects may have resulted in the higher sensitivity and specificity of the latter study. Using the GDx compatible NFA II, Weinreb et al.15 detected sensitivity of 82% and specificity of 62% for a GDx neural network “number” of 17 or greater in 84 healthy individuals and 83 patients with early to moderate glaucomatous visual field loss. In this study, sensitivity and specificity, based on a subjective clinical evaluation of GDx summary data reports, ranged from 72% to 82% and 76% to 82%, respectively, in a similar group of patients.

Trible et al.16 using a GDx “number” greater than 35 to define glaucoma, found a specificity of 89% and corresponding sensitivities of 57%, 71%, and 81% to detect early, moderate, and severe glaucoma, respectively. In this study, a GDx number 35 cut-off value resulted in a sensitivity of 64% and a specificity of 82%. Hence, our results support Trible’s finding using the GDx number.

Using a prototype instrument, Pieroth et al.17 found OCT sensitivity of 65% and specificity of 81% for detecting focal RNFL defects. Our study shows similar OCT results, with sensitivity and specificity ranging from 76% to 79% and 68% to 81%, respectively.

RNFL photography, as a qualitative parameter to detect glaucomatous optic neuropathy, can have a high specificity (100%) but relatively low overall sensitivity (15%), as Jonas et al.18,19 demonstrated with red-free RNFL photography in a group of 398 open-angle glaucoma patients at different stages (early, moderate, end stage) and 234 normal individuals. Sommer et al.20 analyzed serial RNFL photographs of 14 eyes that eventually developed visual field abnormalities and 76 control eyes. Each of the 14 eyes with subsequent visual field abnormalities, but only 9% of control eyes, had prior RNFL defects.

In this study, the glaucoma patients were significantly older than the normal subjects. Hence, it is possible that we may have overestimated the difference between the normal and glaucoma groups. This could have biased sensitivities and specificities, because there are fewer nerve fibers with increased age.21–23 However, the decrease in nerve fibers, particularly with only one decade of age difference, is small (approximately 5000/year).22,23 Moreover, subjects were selected for inclusion so that the youngest glaucomatous patient (43.5 years) and the youngest normal individual (42.4 years) were approximately the same age, and there was no reason for graders to suspect an age difference between the diagnostic groups. Therefore, it is not likely that the age difference between the glaucoma patients and the normal subjects would have impacted the results.

Information obtained with HRT, GDx, and OCT reports has the unique advantage of combining qualitative data such as fundus appearance (HRT) with graphical visual information (OCT and GDx RNFL thickness maps, and HRT reflectance and topography image), as well as quantitative data (stereometric disc parameters and RNFL thickness measures). The use of these reports in clinical practice can aid the clinician in determining the status of the optic disc and RNFL. With further research to identify parameters to improve the sensitivity and specificity of these instruments, the summary data reports could better assist in the formulation of clinical decisions for management of glaucoma patients.

HRT, GDx, and OCT summary data reports when used alone provided a means for differentiating between normal eyes and glaucomatous eyes with mild to moderate visual field loss. The present results are similar to those of studies using HRT, GDx and/or OCT quantitative analysis.7,12,15,24,25 Although the qualitative evaluation of summary data reports did not provide sensitivities and specificities that justify implementing them as primary population screening tools for early to moderate glaucoma, such modalities may be useful in conjunction with current diagnostic techniques.

References