Evaluation of an Algorithm for Detecting Visual Field Defects Due to Chiasmal and Postchiasmal Lesions: The Neurological Hemifield Test

Michael V. Boland,[1,2,3,4] *Allison N. McCoy,*[1,4,5] *Harry A. Quigley,*[1,2] *Neil R. Miller,*[5] *Prem S. Subramanian,*[5] *Pradeep Y. Ramulu,*[1,2] *Peter Murakami,*[6] *and Helen V. Danesh-Meyer*[7]

Purpose. To develop an automated neurologic hemifield test (NHT) to detect visual field loss caused by chiasmal or postchiasmal lesions.

Methods. Visual field locations from 24-2 pattern automated visual fields were grouped into two symmetric regions with 16 points on either side of the vertical meridian. A scoring system similar to the Glaucoma Hemifield Test (GHT) was used to calculate point scores using the pattern deviation values from the right and left regions. The cross-vertical difference in the sum of these values was the NHT score. The NHT was evaluated using visual fields from subjects with known neurologic disease, subjects with glaucoma, and glaucoma suspects (92 pairs of eyes each). The NHT score was calculated for each eye. Four masked reviewers scored all pairs of visual fields with regard to the likelihood of neurologic and glaucomatous optic neuropathy. Both NHT score and expert field ratings were compared with clinical diagnosis by receiver operating characteristic (ROC) analysis.

Results. The NHT effectively discriminated neurologic fields from those of glaucoma patients and glaucoma suspects (area under the ROC curve [AUC] = 0.90; 95% confidence interval [CI], 0.86–0.94). The NHT score correlated well with clinician grading (Pearson correlation estimates, 0.74–0.78). Even when field defects were subtle, the NHT had some ability to discriminate neurologic from nonneurologic fields (AUC 0.68; 95% CI, 0.56–0.79).

Conclusions. The NHT distinguished neurologic field defects from those of glaucoma and glaucoma suspects, rivaling the performance of subspecialist clinicians. Its implementation may help identify unsuspected neurologic disease. (*Invest Ophthalmol Vis Sci.* 2011;52:7959–7965) DOI:10.1167/iovs.11-7868

Visual field testing is central to the diagnosis and management of glaucoma and other diseases affecting the visual system. Automated perimetry is now used routinely to identify and monitor visual field defects quantitatively over time. To assist clinicians in the interpretation of automated perimetry, several algorithms have been developed that include statistical measures (e.g., mean deviation and pattern standard deviation), artificial neural network analysis,[1–3] and rule-based systems, including the Glaucoma Hemifield Test (GHT).[4,5] Although the diagnosis of neurologic disease is critical for patients and clinicians, the software provided for field analysis has tended to focus on glaucoma management. We are aware of only one commercial system designed to help identify neurologic disease from visual fields, and it is no longer available.[6]

The GHT exploits the tendency of glaucoma to damage the upper and lower fields differentially by comparing corresponding clusters of three to six test points above and below the horizontal midline. The method is effective in identifying glaucomatous visual field loss.[7] In contrast, neurologic diseases that affect the visual pathway at or posterior to the optic chiasm produce visual field defects that respect the vertical midline, due to the segregation at the chiasm of retinal ganglion cell axons arising from the nasal and temporal retina. Clinicians may identify these homonymous or heteronymous neurologic patterns by manual inspection of the field data, but until now there has been no automated analysis of the quantitative differences across the vertical midline. To address this need, we created a neurologic hemifield test (NHT) to improve the detection of chiasmal and postchiasmal field loss.

Methods

This research was reviewed and approved by the Institutional Review Board of the Johns Hopkins University School of Medicine and abided by the tenets of the Declaration of Helsinki.

The NHT score was constructed in a manner similar to the approach taken for the GHT. We calculated for each point in the 24-2 pattern test of the Humphrey Field Analyzer (HFA; Carl Zeiss Meditec, Inc, Dublin CA), a number that was inversely proportional to its pattern deviation probability (Table 1). The points were then grouped into two symmetrical regions of 16 points on either side of the vertical meridian (Fig. 1). This distribution of points maximized the number of test locations that could be included, while eliminating the region affected by the physiologic blind spot (Fig. 1, open circles) as well as the nasal periphery, where initial glaucoma damage is prevalent. Other point patterns were also tested, but the pattern used here had a larger

From the [1]Glaucoma Center of Excellence, the [2]Dana Center for Preventive Ophthalmology, and the [5]Neuro-ophthalmology Service, Wilmer Eye Institute, and the [3]Division of Health Sciences Informatics, Johns Hopkins University School of Medicine, Baltimore, Maryland; the [6]Department of Biostatistics, Johns Hopkins School of Public Health, Baltimore, Maryland; and the [7]Department of Ophthalmology, University of Auckland, Auckland, New Zealand.

[4]These authors contributed equally to the work presented here and should therefore be regarded as equivalent authors.

Supported in part by National Eye Institute Grant EY001765 (Core Facility Grant, Wilmer Institute); by unrestricted funds from the Leonard Wagner Trust, New York, NY, from Saranne and Livingston Kosberg, and from William T. Forrester; and by a grant to the Wilmer Eye Institute from Research to Prevent Blindness.

Submitted for publication May 12, 2011; revised June 21, 2011; accepted July 18, 2011.

Disclosure: **M.V. Boland**, None; **A.N. McCoy**, None; **H.A. Quigley**, None; **N.R. Miller**, None; **P.S. Subramanian**, None; **P.Y. Ramulu**, None; **P. Murakami**, None; **H.V. Danesh-Meyer**, None

Corresponding author: Michael V. Boland, Johns Hopkins University, 600 N. Wolfe Street, Wilmer 131, Baltimore, MD 21287; boland@jhu.edu.

Copyright 2011 The Association for Research in Vision and Ophthalmology, Inc.

Table 1. Derivation of Pointwise Scores for the Neurological Hemifield Test from Pattern Deviation Probability Values

Pattern Deviation Probability	NHT Point Score
>10%	0
≤10%,>2%	2
≤2%, >1%	5
≤1%	10

area under the receiver operating characteristic (ROC) curve for distinguishing normal from neurologic visual fields (data not shown). The visual fields used to develop the NHT pattern in Figure 1 were from subjects from the University of Auckland with or without neurologic disease, with the latter serving as controls. These visual fields were distinct from those described below used to evaluate the NHT.

The final NHT score was the absolute value of the difference in the sum of the point scores for right and left regions. This value ranged from 0 to 160 for the visual fields studied here. The greater the numerical value of the NHT, the greater the difference in pattern deviation probability values between nasal and temporal hemifields, and, presumably, the higher the likelihood of chiasmal or postchiasmal disease. A distinct NHT score was assigned to each eye of each subject.

The ability of the NHT to identify neurologic defects was evaluated using visual field tests for subjects with known neurologic disease, subjects with glaucoma, and glaucoma suspects, with 92 right–left pairs for each group. The neurologic fields were selected from the records of the Neuro-ophthalmology Division of the Wilmer Eye Institute and included eyes of patients with either a chiasmal lesion (e.g., pituitary syndrome) or a postchiasmal lesion causing homonymous field loss. Diagnoses in the latter group included cerebral infarction or hemorrhage, intracranial tumor, arteriovenous malformation, aneurysm, demyelinating disease, cortical atrophy, and trauma (Table 2). The diagnosis of neurologic disease was made by a neuro-ophthalmologist based on the clinical history, examination, visual fields, and neuroimaging studies for patients seen between 1999 and 2007.

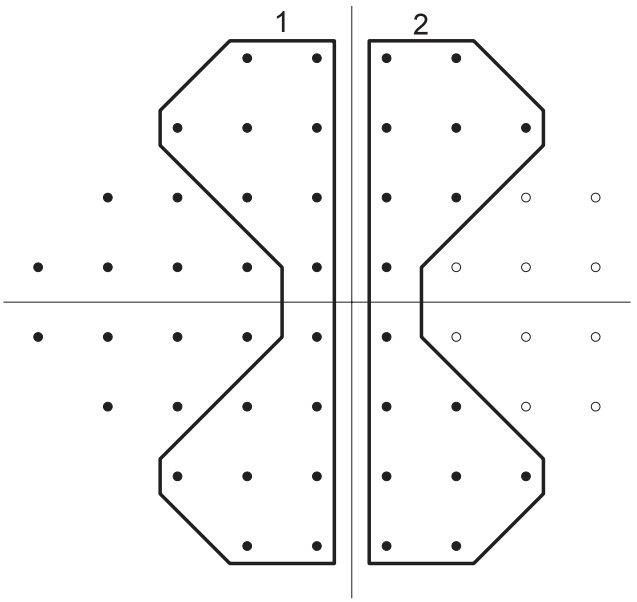

Figure 1. Pattern of points used in the NHT. The test points from a 24-2 pattern visual field are shown. A score is calculated for each of 16 points from each vertical hemifield (surrounded by *solid line*). These pointwise scores are summed for each region, and the final NHT value is the absolute value of the difference between nasal and temporal hemifields. *Open circles*: points that fall in or near the physiologic blind spot in a right eye.

Table 2. Demographics of Subjects, According to Diagnosis

	Neurological	Glaucoma	Glaucoma Suspect
Age, y	52.4 (16.3)	56.1 (15.4)	52.0 (16.6)
Mean deviation, dB	−9.3 (7.2)	−9.0 (7.4)	−1.5 (2.6)
Neurological diagnosis, *n*			
Nonpituitary tumor	31		
Infarction	28		
Vascular malformation	9		
Hemorrhage	7		
Pituitary tumor/lesion	6		
Demyelinating disease	4		
Trauma	3		
Aneurysm	2		
Alzheimer variant	2		

n = 92 in each group. Data are expressed as the mean (SD).

The glaucoma and glaucoma suspect visual fields were systematically selected from a database of approximately 120,000 visual fields of subjects examined at the Wilmer Eye Institute from 1999 through 2007. The diagnosis of glaucoma or glaucoma suspect was specified by clinical billing data. In a separate publication, we validated these diagnoses by chart review of clinical data. The diagnosis from the database matched that of an expert glaucoma subspecialist in 97% of cases.[8]

The fields of glaucoma suspects were used for comparison with the abnormal glaucoma and neurologic fields, because such fields are frequently encountered in an office setting in which clinicians need to distinguish neurologic disease from other diseases such as glaucoma. The visual fields from the glaucoma suspect group had low mean deviation values (Table 2), and 91% of the suspect eyes had a Glaucoma Staging System score of 0 or 1.[9]

To minimize differences between the three groups with respect to demographics and overall degree of field loss, we selected from the database of fields one pair of glaucomatous fields to match each neurologic field pair based on age. We further required both the right and left visual fields from the selected glaucoma subject to have a mean deviation within 30% of the same eye in the neurologic field pair. The visual fields of glaucoma suspects were matched with the neurologic cases on subject age alone. After the three groups of fields were selected, the pointwise pattern deviation values were extracted to calculate the NHT score as described above.

For expert clinician grading, the 276 final pairs of visual fields from the three diagnostic groups were printed in random order, but with both eyes of each subject presented sequentially. Two neuro-ophthalmologists (NRM and PSS) and two glaucoma specialists (HAQ and PYR) reviewed the visual fields independently. The reviewing clinicians were shown the entire visual field report for each eye except for patient identifiers and GHT results, which were obscured. For each pair of right and left visual fields, the reviewers assigned two independent scores: one for the likelihood of glaucoma, and one for the likelihood of neurologic disease. This was done on a scale of 1, unlikely; 2, possibly; 3, equivocal; 4, probably; and 5, definitely. The reviewers were unaware of the relative number of fields included from each category.

The performance of the NHT was compared with the ability of subspecialty-trained reviewers to identify neurologic disease using ROC analysis. Because the NHT score was calculated for each eye individually, we selected the higher NHT from the two eyes (NHT maximum) to compare with the clinician grading. This represents one way to combine the data from two eyes, as physicians would do subjectively. Other combinations of the NHT scores from two eyes of the same patient, including the minimum and average, were tested in our preliminary evaluations of the NHT and were similar to the maximum of the two eyes. Likewise the performance of the GHT was compared with the ability of clinicians to identify glaucomatous optic

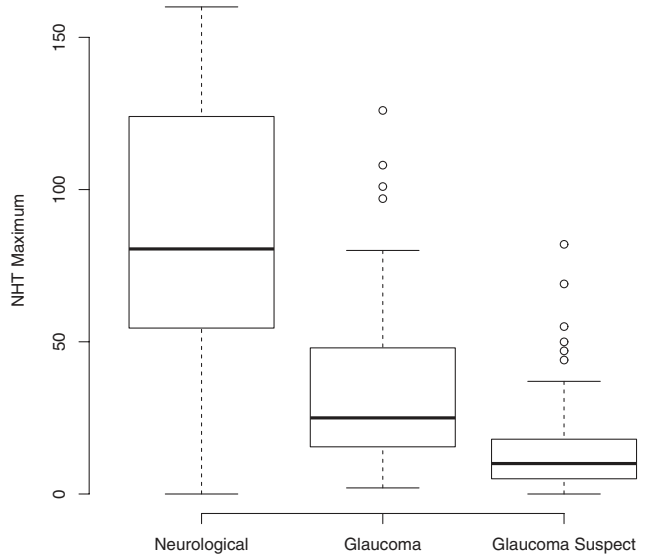

FIGURE 2. Comparison of NHT maximum score by diagnostic group. Box plot of NHT score for neurological, glaucoma, and glaucoma suspect subjects using the higher NHT score for the pair of fields from each patient (NHT Maximum). *Dark line*: the median; the box includes 75% of the values, and the outer flag is the range, with outliers indicated by *circles*.

neuropathy. GHT values "within normal limits," "borderline," and "outside normal limits" were considered to be an ordinal scale, and we selected the higher GHT for the two eyes to compare with the clinician grading. Any fields with GHT values of "general reduction of sensitivity" and "abnormally high sensitivity" were excluded from this analysis, and the GHT value of the fellow eye was chosen if it was within normal limits, borderline, or outside normal limits. The areas under ROC curves were compared using the method of DeLong et al.[10] All analyses were performed with R (ver. 2.12.1).[11–15]

RESULTS

The age, diagnosis, and visual field mean deviation for subjects are compared in Table 2. Mean age was similar across all three groups ($P = 0.16$, ANOVA). The visual field mean deviation was similar between the glaucoma and neurological groups ($P = 0.63$, Wilcoxon rank-sum test), although as might be expected, both of these groups had visual field loss greater than that of the glaucoma suspects ($P < 0.001$, ANOVA). Patients with neurologic disease had higher mean NHT maximum visual field scores than patients in the glaucoma or glaucoma suspect groups (Fig. 2). The mean NHT maximum for neurologic subjects (86) was significantly higher than that of nonneurologic subjects (24, $P < 0.001$, Wilcoxon rank-sum test). The NHT scores for right and left eyes of neurologic patients correlated highly (0.82; 95% confidence interval [CI], 0.78–0.86), as might be expected of homonymous and bitemporal defects.

Fields with GHT values of within normal limits or borderline had low NHT scores, all less than 50, whereas fields with a GHT score of outside normal limits spanned the full range of NHT scores (Fig. 3). Among fields with a GHT result of outside normal limits, those belonging to glaucoma patients fell at the

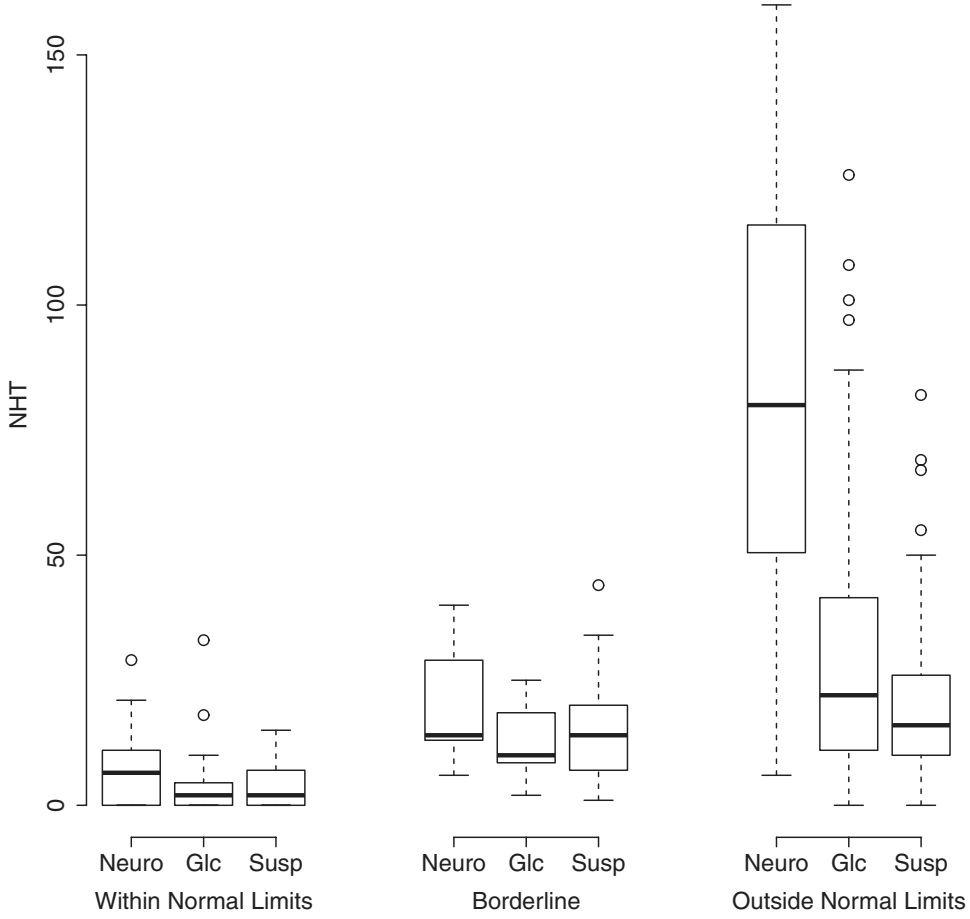

FIGURE 3. Distribution of NHT score by GHT category for each diagnosis group. All eyes, left and right, are grouped together.

low end of the NHT score range (Fig. 3). We note that the GHT is designed to classify a field as outside normal limits when the field mean deviation is high, even when the difference in summed point scores between upper and lower hemifields is low. Hence, the neurologic fields with higher mean deviation would be classified as outside normal limits by the GHT on this basis as well, despite their having no difference between upper and lower field sensitivity. The NHT, as implemented here, has no similar criterion to generate an abnormal NHT result due to diffuse vision loss.

The sensitivity and specificity of NHT scores were assessed for their ability to distinguish field loss corresponding to neurologic disease from that of glaucoma or glaucoma suspect patients combined. The area under the ROC curve (AUC) was 0.90 (95% CI, 0.86–0.94), when the NHT maximum value was used (Fig. 4). The result was similar whether we used the NHT score for right or left eyes, the minimum or the mean NHT score for the pair (range, 0.88–0.91). Using an NHT maximum score of 30 as a cutoff for classifying visual field defects as likely neurologic yielded a sensitivity of 87% (95% CI, 78%–93%) and specificity of 73% (95% CI, 66%–79%).

Physician reviewers graded each field pair on a five-level scale for the likelihood of neurologic disease. Although there were modest differences among the four reviewers, all generally were able to identify neurologic fields correctly, as judged by placing many of them into the probably or definitely categories (Fig. 5). The interrater reliability for the four reviewers was high, with ICCs of 0.86 (95% CI, 0.84–0.89) for neurologic fields and 0.78 (95% CI, 0.74–0.81) for glaucoma fields. The mean reviewer sensitivity (i.e., the proportion of true neurologic fields placed in the "probably" or "definitely" categories), was estimated at 73% (95% CI, 59%–84%). The mean reviewer specificity (i.e., the proportion of nonneurologic cases placed in the "unlikely, " "possibly, " or "equivocal" neurologic categories, was estimated at 97% (95% CI: 96%, 98%). For comparison, the specificity of the maximum NHT was 91% (95% CI, 86%–95%) for a sensitivity of 73% (95% CI 63%–82%). The area under the ROC curve for the four reviewers ranged from 0.89 to 0.92. The NHT and reviewers showed a similar ability to identify neurologic fields correctly, with Pearson correlation coefficients for the comparison between the NHT maximum and each reviewer ranging from 0.75 to 0.78, and no significant difference in the areas under the ROC curves (all $P > 0.05$; Fig. 6).

Using ROC analysis, the NHT and GHT algorithms were compared with each other for their ability to identify the conditions for which they were designed. The NHT discriminated neurologic fields from nonneurologic (glaucoma and glaucoma suspect) fields with an AUC of 0.9, whereas the GHT distinguished glaucomatous fields from nonglaucomatous (neurologic and glaucoma suspect) fields with an AUC of 0.63. The areas under these ROC curves were significantly different ($P < 0.0001$). For the easier task of distinguishing neurologic fields from those from glaucoma suspects, the NHT produced an AUC of 0.95. Similarly, the GHT distinguished glaucomatous fields from fields of glaucoma suspects with an AUC of 0.79 (significantly different, $P < 0.0001$). The NHT therefore performed significantly better than the GHT in distinguishing the condition for which it was designed from all other fields as well as from glaucoma suspect fields alone.

To evaluate the ability of the NHT to identify neurologic disease when the expert reviewers were most uncertain with regard to neurologic defects, we stratified the NHT analysis based on the clinician ratings of neurologic disease. In eyes with mean clinician gradings of 1 to 3 (i.e., fields with less certain neurologic characteristics), the NHT still had some ability to discriminate neurologic fields from the combined group of glaucoma and glaucoma suspect fields (AUC 0.68; 95% CI, 0.56–0.79). An example of a subtle neurologic field defect is shown in Figure 7. This particular pair of fields received an NHT score of 32 in each eye and was classified as probably or definitely neurologic (clinician grade 4 or 5)

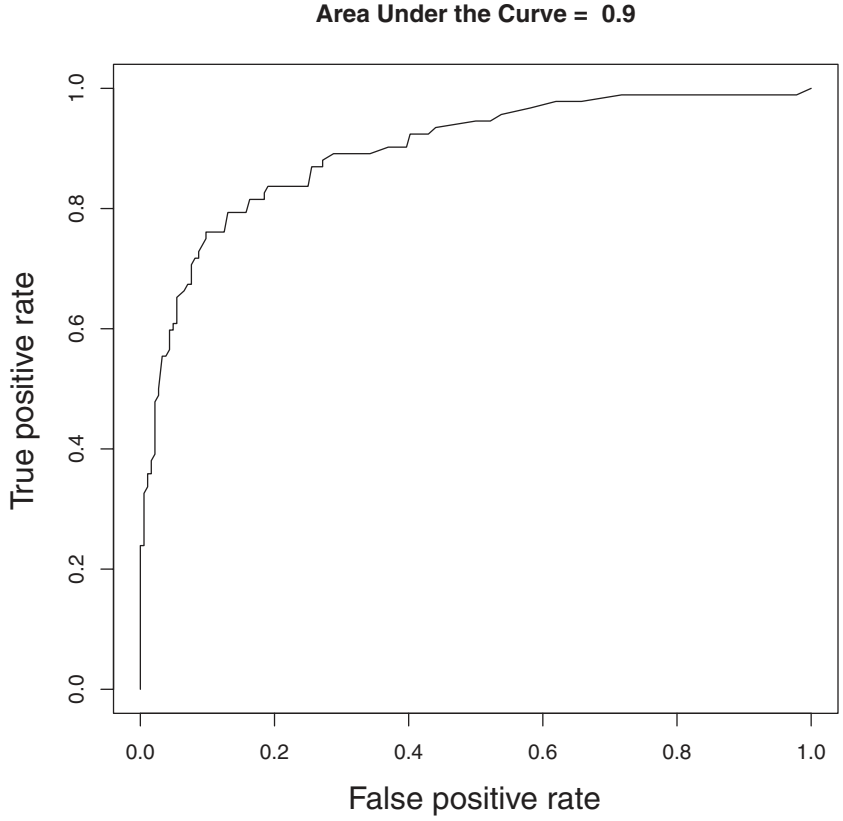

Figure 4. ROC curve for neurologic versus nonneurologic pairs of eyes using the maximum NHT value from each right/left pair of eyes.

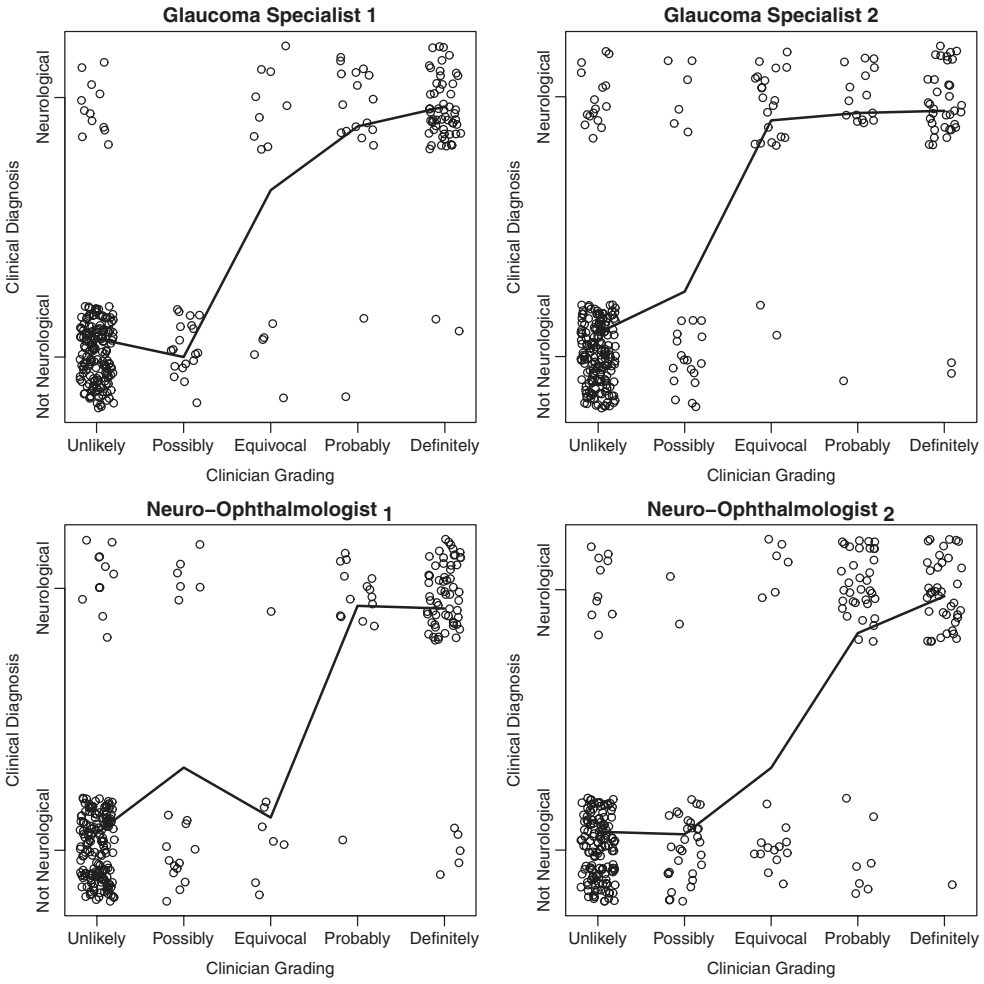

FIGURE 5. Clinician assessment of the likelihood of neurologic optic nerve disease. The diagnosis of "not neurological" includes visual fields from glaucoma and glaucoma suspect patients. *Line*: the proportion of neurologic diagnoses at each category of clinician grade.

by both neuro-ophthalmologists but by neither glaucoma specialist.

DISCUSSION

We have developed an algorithm, the NHT, to assist in the recognition of neurologic field defects in data from automated threshold perimetry. The NHT makes use of the pointwise pattern deviation probability data in each visual field test, comparing corresponding regions left and right of the vertical meridian. We have shown that this test can reliably detect visual field loss due to chiasmal or postchiasmal neurologic disease, and its performance compares favorably with that of subspecialty trained clinicians. Compared with the GHT, the NHT appears to be at least as sensitive and specific; in fact, the NHT performed significantly better than the GHT in distinguishing the disorder for which it is designed.

We were particularly encouraged that the NHT had a reasonable mixture of sensitivity and specificity for those neurologic cases that were considered subtle or equivocal by the reviewers. The field pair illustrated in Figure 7 provides an example of how early neurologic field defects could be identified by the NHT and neuro-ophthalmologists, but missed by glaucoma specialists and presumably general ophthalmologists as well. Although it would be easy to set criteria for the NHT that identified only obvious homonymous hemianopia, our purpose was to identify relatively subtle neurologic defects. The value of this approach will be greatest if it helps those clinicians who are not highly expert in evaluating neuro-ophthalmic disease to identify subtle chiasmal or postchiasmal field loss.

A sensitivity of 87% and specificity of 73% for an NHT score of 30 would not be highly useful in population screening for neurologic defects. With the low prevalence of true neurologic disease, the specificity would have to be set at a very high level to avoid having many more false positives than true positives. Given the results herein, one could still identify half of the neurologic defects at a specificity of nearly 99%. On the other hand, given the potential catastrophic consequences of missing a neurologic field defect and the ability to diagnose such defects with relatively noninvasive measures, it would also be reasonable to choose an NHT cutoff with high sensitivity at the expense of more false positives. We propose, however, that there is little use for complex automated perimetry in population screening. Rather, the NHT has greatest value in the clinical office where, most often, visual field testing is performed on glaucoma suspects. In this setting, unsuspected neurologic disease may be identified with such assistance, and it was for this reason that we compared visual field data from patients with neurologic disease with those of glaucoma and glaucoma suspect patients.

Although the NHT is promising for clinical use, our study has some limitations. Most important, the clinicians reviewed pairs of fields whereas the NHT, much like the GHT, only assessed one field at a time. To compare the NHT performance with the clinician grades, we selected the higher NHT for each pair of fields. We plan to develop a more sophisticated binocular analysis for a future version of the NHT. It will attempt to

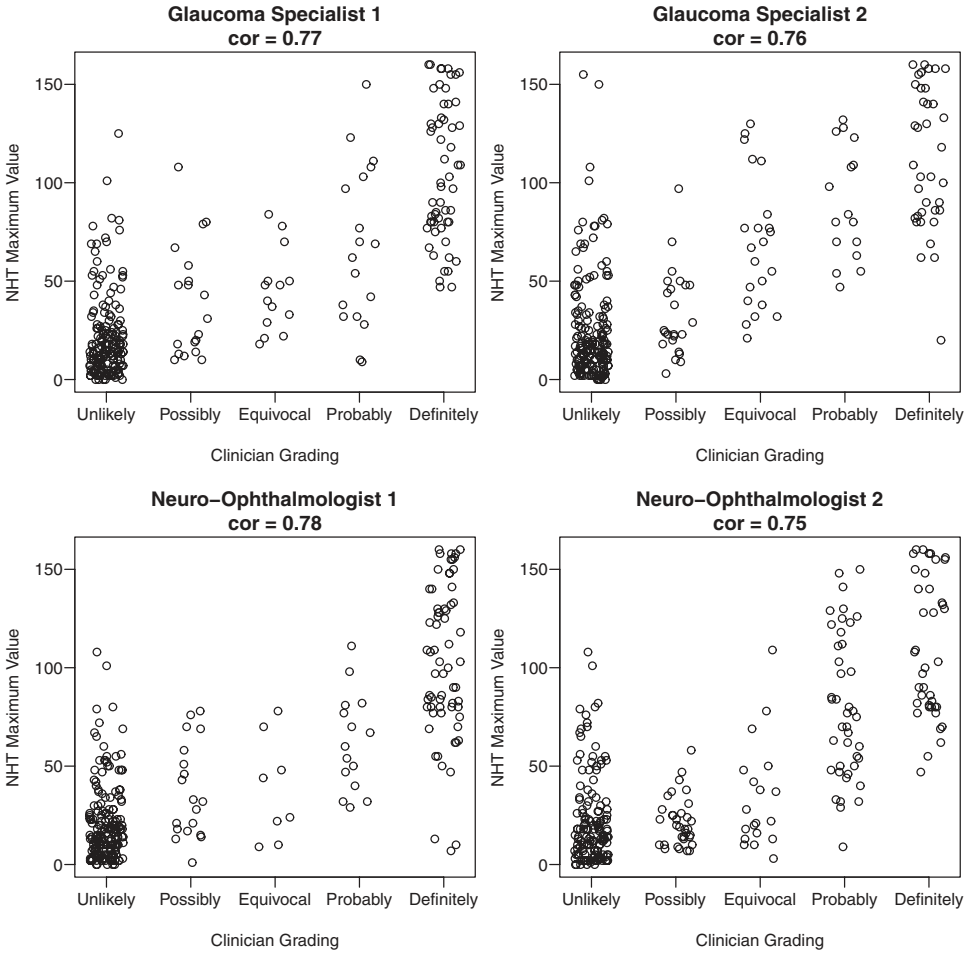

Figure 6. NHT maximum score versus clinician assessment of the likelihood of neurologic disease. For each of four reviewers, all pairs of fields were categorized from unlikely (1) to definitely (5) neurologic. Pearson correlation coefficients are shown.

integrate visual field data from both eyes and determine not only the probability of neurologic disease, but also if the pattern of field loss corresponds to a homonymous or heteronymous hemianopia.

Another important caveat is that the NHT is designed to detect chiasmal and postchiasmal visual field defects, but cannot distinguish defects due to nonglaucomatous optic neuropathies, which are often similar to those caused by glaucoma. Likewise, the GHT cannot distinguish many glaucomatous defects from those of ischemic optic neuropathy, compressive optic neuropathy, or branch artery occlusions. Additional clinical data must be used to assist in this differentiation, including visual acuity, color vision, ophthalmoscopic appearance of the optic disc and retina, and, of course, historical information. It was not our purpose to propose the NHT as an overall solution to the diagnosis of neurologic disease, but rather to aid identification of certain types of visual field defects.

In summary, the NHT can distinguish chiasmal and postchiasmal visual field defects from defects caused by glaucoma, and it rivals the performance of glaucoma and neuro-ophthalmology specialists in discriminating neurologic field

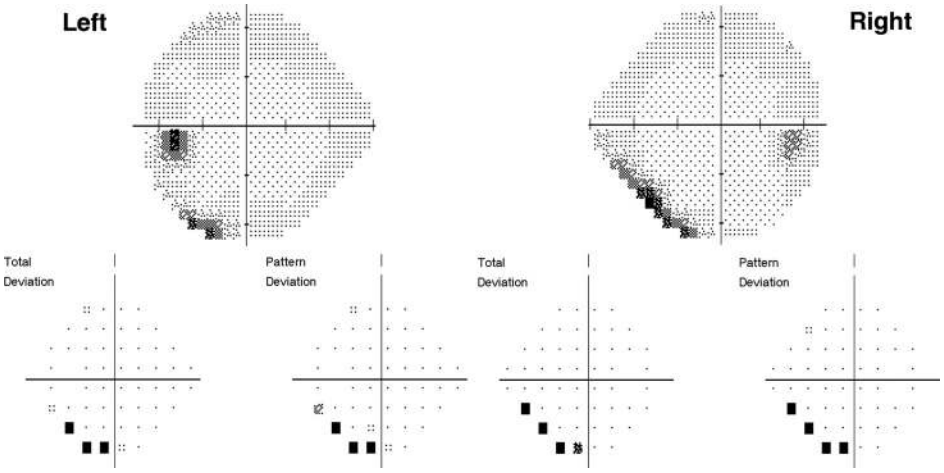

Figure 7. Example of an early neurologic field defect in the eyes of a single subject identified by the NHT and neuro-ophthalmology specialists.

defects. Despite its limitations, we believe this test may be a useful adjunct in the interpretation of automated perimetry. It can corroborate a clinician's suspicion for neurologic field defects, much as the GHT does for glaucomatous defects. Moreover, it can alert clinicians to the possibility of unsuspected neurologic disease.

References

1. Goldbaum MH, Sample PA, White H, et al. Interpretation of automated perimetry for glaucoma by neural network. *Invest Ophthalmol Vis Sci.* 1994;35:3362–3373.
2. Lietman T, Eng J, Katz J, Quigley HA. Neural networks for visual field analysis: how do they compare with other algorithms? *J Glaucoma.* 1999;8:77–80.
3. Bizios D, Heijl A, Bengtsson B. Trained artificial neural network for glaucoma diagnosis using visual field data: a comparison with conventional algorithms. *J Glaucoma.* 2007;16:20–28.
4. Asman P, Heijl A. Evaluation of methods for automated hemifield analysis in perimetry. *Arch Ophthalmol.* 1992;110:820–826.
5. Asman P, Heijl A. Glaucoma hemifield test. automated visual field evaluation. *Arch Ophthalmol.* 1992;110:812–819.
6. Nagata S, Kani K, Sugiyama A. A computer-assisted visual field diagnosis system using a neural network. Amsterdam: Kugler: *Perimetry Update 1990–1991.* 1991:291–295.
7. Katz J, Quigley HA, Sommer A. Repeatability of the glaucoma hemifield test in automated perimetry. *Invest Ophthalmol Vis Sci.* 1995;36:1658–1664.
8. Boland MV, Quigley HA. Evaluation of a combined index of optic nerve structure and function for glaucoma diagnosis. *BMC Ophthalmol.* 2011;11:6.
9. Mills RP, Budenz DL, Lee PP, et al. Categorizing the stage of glaucoma from pre-diagnosis to end-stage disease. *Am J Ophthalmol.* 2006;141:24–30.
10. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* 1988;44:837–845.
11. R Development Core Team. *R*: *A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2010.
12. Harrell FE Jr. Design: package design. 2009. Available at http://cran.r-project.org/web/packages/Design/index.html. Accessed July 18, 2011.
13. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: Visualizing the performance of scoring classifiers in R. *Bioinformatics*. 2005; 21:3940–3941.
14. Gamer M, Lemon J, Fellows I, Singh P. The irr package, Ver. 0.83.: Various coefficients of interrater reliability and agreement. 2010. Available at http://cran.r-project.org/web/packages/irr/index.html. Accessed July 18, 2011.
15. Lumley T. Survey: Survey analysis of complex survey samples. R package version 3.24-1. Seattle, WA: Department of Biostatistics, University of Washington; 2011.