Automated Suprathreshold Screening for Glaucoma: The Baltimore Eye Survey

Joanne Katz,* James M. Tielsch,* Harry A. Quigley,* Jonathan Javitt,† Kathe Witt,* and Alfred Sommer%

Purpose. To evaluate automated suprathreshold perimetric screening for glaucoma in a population-based survey of ocular disorders in east Baltimore, Maryland.

Methods. A population-based sample of persons ≥ 40 years of age residing in 16 clusters was selected for an ocular screening examination that included automated suprathreshold testing with the Full Field 120 program of the Humphrey Field Analyzer. Subjects who failed the test underwent manual testing to confirm the defect. Subjects were referred for definitive examination by an ophthalmologist if they had an abnormal field, visual acuity worse than 20/30, intraocular pressure > 21 mm Hg, optic disc damage, a history of glaucoma, or shallow angles. The sensitivity and specificity of the automated visual field testing for identifying glaucoma was estimated and compared with other methods to screen for glaucoma.

Results. Of 5,341 subjects ≥ 40 years of age who underwent a screening eye examination at neighborhood centers, 4,735 (89%) completed the automated field test. The median test time was 7.25 minutes per eye. Screening test results were abnormal in one or both eyes in 1,234 (26%) of the subjects. Kinetic perimetry was performed on 95% of these subjects, and defects were confirmed for 448 (36%) of them. Hence, 9.5% of the 4,735 subjects who completed the automated test were referred for definitive examination because the defect on automated perimetry was confirmed on manual testing. For a specificity of 90%, the sensitivity of the screening visual field test to detect glaucoma was 52% for 17 or more relative or absolute defects, higher than that of intraocular pressure at 39% for a cut-off of 20.5 mm Hg, vertical cup-to-disc ratio at 45% for a cut-off of 0.53, narrowest remaining rim width at 42% for a cut-off of 0.16, and was comparable to a combination of these and other nonfield parameters.

Conclusion. Suprathreshold testing performed better than nonperimetry-based screening tests for glaucoma. However, a number of logistical weaknesses of this visual field screening method were identified. Invest Ophthalmol Vis Sci. 1993;34:3271–3277.

A variety of techniques have been advocated to screen populations for glaucoma.1–8 These include tonometry to measure intraocular pressure, direct ophthalmoscopy to observe the features of the optic nerve head, and questionnaires to ascertain family and personal histories of glaucoma and other risk factors such as diabetes. These methods have drawbacks that were highlighted in a recent population-based survey of ocular disorders in a multiracial urban setting in east Baltimore.9 Tonometry, using the traditional cut-off of greater than 21 mm Hg for intraocular pressure, yielded a sensitivity of 47% and a specificity of 92% for identifying persons with glaucomatous optic nerve damage. Assessment of optic disc characteristics from stereoscopic color photographs did not produce better results.10 The addition of age, race, family history of glaucoma, and history of diabetes to optic disc findings and tonometry slightly improved sensitivity and specificity. Most of the improvement was attributable to the addition of optic disc characteristics. However, photographic optic disc findings were available for only 53% of those with glaucoma and 78% of those...
without glaucoma because of difficulties in obtaining readable photographs among the elderly.\textsuperscript{9}

Within the past decade, automated threshold testing of the visual field has become widely used as a definitive test for glaucoma. Suprathreshold screening programs take less time to complete than threshold testing. Our purpose was to determine the performance characteristics of automated suprathreshold screening for glaucoma in a representative sample of blacks and whites who underwent a battery of examinations in a population-based setting. We compare this technique to other traditional screening methods.

METHODS

A population-based prevalence survey of ocular disorders (the Baltimore Eye Survey) was conducted in east Baltimore, Maryland, from January 1985 to November 1988. All noninstitutionalized persons 40 years of age or older permanently residing in 16 neighborhood clusters were eligible to participate. The sampling design and study methods have been described previously.\textsuperscript{10,11} Briefly, 7,104 black and white persons 40 years of age and older residing in 16 geographic clusters representative of the east Baltimore population were identified as eligible for the survey. A total of 6,850 (96.4\%) persons completed a short enrollment interview. One hundred forty-five subjects were ineligible because they were medically incapable of participating or they died after the enrollment interview but before their scheduled appointment for an ocular screening examination at a neighborhood center. A screening examination was completed for 5,341 persons (79.2\% of eligible subjects), of whom 33 were nonwhite and nonblack. They were excluded from this analysis because of their small numbers, leaving a study population of 5,308 persons. The screening examination included refraction, visual acuity, applanation tonometry, stereoscopic fundus photography of the optic disc and macula, and visual field testing. Anthropometric measurements, blood pressure, pulse rate, and a detailed medical and personal history were also taken. Informed consent was obtained from all participating subjects in accordance with the tenets of the Helsinki Declaration, and ethical approval for the study was obtained from the Joint Committee on Clinical Investigations of the Johns Hopkins Medical School.

The Full Field 120 protocol of the Humphrey Field Analyzer was used as the initial visual field screening test. If both eyes were normal as described above, no further visual field testing was done. If either eye was abnormal, kinetic manual perimetry using a Goldmann perimeter was performed.

The Full Field 120 protocol of the Humphrey Field Analyzer first establishes the foveal threshold for the individual being tested. This threshold value is then used to construct an expected hill of vision assuming a 3 dB decrease in threshold sensitivity for every 10° of eccentricity from fixation. One hundred twenty points are then tested within a 60° field (Fig. 1). A size III (0.5°) stimulus is used for all presentations. Each of the 120 points is presented at a luminance of 6 dB brighter than the expected threshold at that location. If this stimulus is seen, the location is marked with an open circle. If not, a stimulus of 10,000 apostilbs is presented. If this stimulus is seen, a relative defect is marked with an X. If the 10,000 apostilb stimulus is not seen, an absolute defect is denoted by a black square (Fig. 1). The central 30° field is tested with correction for refractive error and addition near correction appropriate for age. We judged the test abnormal if there were 17 or more relative or absolute defects out of the 120 points tested. This cut-off was chosen because it resulted in a sensitivity of 96\% and a specificity of 79\% in a previously tested clinic-based population.\textsuperscript{12} We also considered the test abnormal if eight relative or absolute defects were present in any one quadrant. This criterion is not included in the automated calculation of the perimeter and was ascertained by inspection of the printout. We added this second criterion because a very high sensitivity for visual field screening was considered more critical than specificity in this research setting. However, different criteria might be used in designing cost-effective screening programs where high specificity at the expense of some sensitivity is more appropriate.

The Full Field 120 suprathreshold screening test of the Humphrey Field Analyzer was used as the initial visual field screening test. If both eyes were normal as described above, no further visual field testing was done. If either eye was abnormal, kinetic manual perimetry using a Goldmann perimeter was performed.
on that eye. Our standard protocol for Goldmann perimetry has been published previously and included testing of at least four isopters kinetically, three of them within the central 30° field. This is a modification of the Armaly Drance protocol. Static testing was done central to each isopter with the same stimulus used for that isopter. Goldmann visual field results were considered abnormal if there was a paracentral or full arcuate scotoma at least 0.4 log units deep, a nasal step at least 10° wide and present to at least two isopters, a substantial arc-shaped blind spot extension, temporal or central islands, or hemispheric loss. If the Goldmann field was abnormal, the subject was referred to the Wilmer Institute for a definitive eye examination by a glaucoma subspecialist. If Goldmann visual field results were normal, no further visual field testing was done and the subject was referred for a definitive examination only if he or she met any of the following criteria: best corrected visual acuity < 20/30 in either eye; intraocular pressure > 21 mm Hg in either eye; horizontal or vertical cup-to-disc ratio ≥ 0.7 or a narrowest neuroretinal rim width < 0.15 in either eye based on photographic readings; a history of treatment for glaucoma; shallow or narrow anterior chamber angles on penlight examination.

A detailed description of the definition of glaucoma has been provided previously. Briefly, visual fields and optic nerve characteristics obtained and evaluated during definitive examinations were used by a glaucoma specialist to determine whether the subject had glaucoma and was in need of clinical management. The entire clinical record of each person who was clinically determined to have glaucoma was then reviewed by an independent glaucoma specialist, who graded the severity of visual field defects in both the automated and Goldmann fields and determined whether the damage was from glaucoma. This grading system has been described previously. Visual fields were compared with optic nerve findings and nerve fiber layer photographs (where available) for consistency of severity and location of damage. Glaucomatous damage was based on visual fields and on optic nerve and nerve fiber layer appearance, and it was categorized as definite, probable, or uncertain/unknown. Congruency between automated and manual field defects provided stronger evidence of glaucomatous damage, as did congruency between field and disc findings. Angles were graded as open or narrow/closed based on gonioscopy and slit-lamp examination. Glaucoma was also categorized as primary or secondary. Intraocular pressure was not used as a diagnostic criterion. All persons who met the criteria for definite or probable glaucoma of any type were defined as cases for this analysis. One hundred ninety-six persons met this definition. There were 161 cases of primary narrow-angle glaucoma, 42 of secondary open-angle glaucoma, 25 of primary open-angle glaucoma, and 12 of secondary narrow-angle glaucoma.

The specificity of the automated screening visual field test was defined as the proportion of persons with normal results on visual field testing from among all those nonglaucomatous subjects able to complete the automated field test. This assumed that all cases of glaucoma in the population were detected. Although one-third of the 5,308 persons received a definitive ophthalmic examination, the rate of disease among subjects with normal visual fields, visual acuity ≥ 20/30, intraocular pressures ≤ 21 mm Hg, normal optic discs, and no history of glaucoma was likely low. Such low rates would have limited the effects of any potential misclassification for this analysis. The "sensitivity" of automated field testing was defined as the proportion of abnormal visual field test results among persons with any type of glaucoma who were able to perform automated testing. Because all cases of glaucoma were not used in the denominator for this analysis (only persons who had glaucoma and could perform automated testing were included), our sensitivity is not true sensitivity. However, it seemed only fair to evaluate automated perimetry's performance among those who could actually take this test. To evaluate changes in sensitivity and specificity of automated suprathreshold testing with varying cut-offs, we varied the total number of relative or absolute defects and did not include the criterion of 8 or more defects. The sensitivity and specificity of the total Full Field 120 defects was compared with the sensitivity and specificity of other screening methods alone and in varying combinations using different cut-offs of the predicted distribution of being a case obtained from logistic regression models.

RESULTS

A total of 5,308 white and black subjects completed the ocular screening examination. There were 4,735 subjects able to perform the Full Field 120 test in at least one eye (Table 1). In 233 subjects, only manual Goldmann perimetry was performed. Goodlite or confrontation testing was performed on 190 subjects. The Goodlite test is similar to tangent screen testing but
uses a clear plastic bowl. This was done only for homebound subjects unable to attend the neighborhood screening site and a small number who were unable to perform Goldmann perimetry. No visual field testing was done for 150 (2.8%) subjects, half of whom were homebound and the remainder of whom were blind or had severe central vision loss and accompanying poor fixation. The median time taken to complete the Full Field 120 test was 7 minutes, 16 seconds per eye (range, 1 minute 15 seconds to 17 minutes 13 seconds). Right eyes were tested first. There was no statistically significant difference between the mean times taken to complete the test for right and left eyes.

Twenty-six percent of Full Field 120 tests performed at the screening center were abnormal based on the number of total defects or quadrant-based criteria (Table 2). Of those persons who had manual Goldmann testing because they were unable to perform the automated test, 67.8% had abnormal visual fields. The rate of abnormality among those tested by Goodliffe or confrontation strategies was similar to that of automated testing. Eighty-three percent of subjects with an abnormal Full Field 120 test had 17 or more relative or absolute defects. Sixteen percent were abnormal only because of eight or more defects in any one quadrant. In two cases, the basis for judging the field abnormal was unknown because the printout could not be located and was not saved on the perimeter’s computer.

Of the 1,234 subjects who tested abnormal in at least one eye on the Full Field 120 test, 448 (36.3%) also had abnormal results on manual Goldmann perimetry (Table 3). This corresponds to 9.5% of all subjects able to complete the Full Field 120 test. These subjects were referred for definitive examination based on visual acuity, intraocular pressure, or optic disc appearance. On further testing, some of these subjects were found to have an abnormal manual or automated field. We also calculated the sensitivity and specificity for different cut-offs of the total relative or absolute defects. Sensitivity and specificity could not be calculated for cut-offs below 17 defects because no confirmatory testing or referral for definitive examination was done for subjects with tests below this cut-off. The use of ≥17 defects without the cluster of 8 criterion resulted in a decline in sensitivity from 83.6% to 74.7%, and an increase in specificity from 74.9% to 78.9% (Table 4). If the Full Field 120 test had been the only way in which subjects were screened, 74.7% of subjects with glaucoma would have been detected at a cut-off of ≥17 defects. The other 25% would have been referred for definitive examination either be-

### Table 2. Number and Percent of Visual Field Abnormalities on Initial Testing

<table>
<thead>
<tr>
<th>Field Test</th>
<th>N</th>
<th>No. Abnormal</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full field 120</td>
<td>4735</td>
<td>1234</td>
<td>26.1</td>
</tr>
<tr>
<td>Manual Goldmann</td>
<td>223</td>
<td>158</td>
<td>67.8</td>
</tr>
<tr>
<td>Goodliffe/confrontation</td>
<td>190</td>
<td>40</td>
<td>21.1</td>
</tr>
</tbody>
</table>

### Table 3. Results of Confirmatory Manual Goldman Perimetry for Those With Abnormal Full Field 120 Tests

<table>
<thead>
<tr>
<th>Outcome</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>722</td>
<td>58.5</td>
</tr>
<tr>
<td>Abnormal</td>
<td>448</td>
<td>36.3</td>
</tr>
<tr>
<td>Not done</td>
<td>64</td>
<td>5.2</td>
</tr>
<tr>
<td>Total</td>
<td>1234</td>
<td>100.0</td>
</tr>
</tbody>
</table>

74.9%, respectively, using the combined criteria of ≥17 defects or ≥8 defects, or both, per quadrant (Table 4).

### Table 4. Sensitivity and Specificity of Suprathreshold Visual Field Testing for all Types of Glaucoma

<table>
<thead>
<tr>
<th>Cutoff</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥17 and/or cluster of 8</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>≥17</td>
<td>3439</td>
<td>74.9</td>
</tr>
<tr>
<td>≥20</td>
<td>3695</td>
<td>80.6</td>
</tr>
<tr>
<td>≥25</td>
<td>4005</td>
<td>87.3</td>
</tr>
<tr>
<td>≥30</td>
<td>4141</td>
<td>90.5</td>
</tr>
<tr>
<td>≥35</td>
<td>4229</td>
<td>92.4</td>
</tr>
<tr>
<td>≥40</td>
<td>4316</td>
<td>94.1</td>
</tr>
<tr>
<td>Total</td>
<td>4587*</td>
<td>146†</td>
</tr>
</tbody>
</table>

* Five subjects with glaucoma were unable to complete the full field 120 test.† Two subjects with abnormal full field 120 tests did not have test results saved on the Humphrey Analyzer and printouts were not available in charts; 523 subjects without glaucoma were unable to complete the full field 120 test.
cause of a cluster of 8 defects (an additional 13 cases or 9% of all cases), visual acuity < 20/30, optic disc abnormalities, elevated intraocular pressure, or a history of glaucoma treatment.

Twenty-four subjects identified as having glaucoma did not meet our definition of abnormality on the Full Field 120 test. Seventeen of these had a normal Full Field 120 in both eyes. Seven had a normal Full Field 120 in one eye but were unable to do the automated test in the fellow eye because of poor visual acuity, difficulties with fixation, or both. Five of these seven had an abnormal manual field test and two had visual acuities so poor they were unable to do a Goldmann test. These seven subjects were identified as having glaucoma because of grossly asymmetric cupping or defects on Goldmann perimetry. Nasal steps at least 10° wide and present to at least 2 isopter, arcuate, or paracentral scotomas at least 0.4 log units deep and 5° wide, and central or temporal islands constituted such defects. The 17 persons in which Full Field 120 test results were normal in both eyes were identified as cases because of clear evidence of optic nerve damage, asymmetric cupping, or nerve fiber layer defects consistent with disc damage using all available photographic and clinical evidence.

At the cut-off of ≥17 defects, 21.1% of those without glaucoma would have been referred for further testing. A specificity of 90% was achieved if a cut-off of ≥30 defects was used (Table 4 and Fig. 2). At this cut-off, only 53.4% of glaucoma cases would have been correctly identified as such. An analysis using only cases of primary open-angle glaucoma rather than all types of glaucoma yielded similar results, with a sensitivity of 54.8% and a specificity of 90% (cut-off ≥30 defects).

The sensitivity and specificity of the Full Field 120 test was compared to those of other more traditional screening methods in the same east Baltimore population. The visual field screening test using total defects as the sole criterion had higher sensitivity at each level of specificity than intraocular pressure, cup-to-disc ratio, and narrowest rim width (Table 5 and Fig. 3). The Full Field 120 had a sensitivity of 52%, whereas intraocular pressure, cup-to-disc ratio, and narrowest rim width had sensitivities of 39%, 45%, and 42%, respectively, at a specificity of 90%.

A combination of age, race, intraocular pressure, family history of glaucoma, and history of diabetes produced sensitivities roughly comparable to the Full Field 120 at the same levels of specificity. If cup-to-disc and narrowest rim width parameters were added to these other screening variables, the sensitivity was slightly higher at each level of specificity than the Full Field 120 test alone. Although this combination of parameters had higher sensitivity for the same specificity than the Full Field 120, the full combination of parameters was only available for 53% of glaucoma cases and 78% of those without glaucoma. Ninety-five percent of the unavailable data was attributed to the inability to take photographs of adequate quality to assess the optic disc parameters. Seventy-five percent of persons with glaucoma and 90% of those without it completed the Full Field 120 test, a higher proportion than those able to complete photography. We also assessed the sensitivity and specificity of age, race, family history of glaucoma, history of diabetes, and total defects found on the Full Field 120. We examined this combination of screening parameters because they represented methods that would require the skills of a technician.
rather than the skills of an ophthalmologist or an optometrist. For each level of specificity, the sensitivity of this combination was higher than that of age, race, intraocular pressure, family history of glaucoma, and history of diabetes. The addition of the total number of defects on the Full Field 120 test to all other screening parameters produced the highest sensitivity for each level of specificity. The combination of age, race, family history of glaucoma, history of diabetes, intraocular pressure, cup-to-disc ratio, rim width, and total defects on Full Field 120 test.

DISCUSSION

The Full Field 120 test was successfully completed by 89.8% of all subjects. Subjects visited in their homes may have been able to complete this test, but transportation of the equipment into the home was impractical. Thirty-six (51%) subjects who came to the screening center but were unable to complete the test in either eye were blind or visually impaired. In contrast, intraocular pressure measurements could be obtained on 98.8% of all subjects, but the sensitivity of intraocular pressure in detecting glaucoma was 39%, compared to 52% for suprathreshold testing at a specificity of 90% for both screening methods. The cup-to-disc ratio and narrowest rim width generally had slightly poorer sensitivities at the same specificity than intraocular pressure or the Full Field 120 test. For the same specificity, a combination of methods that required only a technician had better sensitivity than the Full Field 120 alone. The addition of intraocular pressure and optic disc characteristics improved on the sensitivity for each level of specificity. However, optic disc photographs of adequate quality could not be obtained on a substantial fraction of subjects, especially those who were older and those who had glaucoma.

Photographic assessment as a screening method is also difficult because pupils have to be dilated for stereo photographs to be obtained, and the results are not available as soon as the photographs have been taken. Hence, subjects must be contacted days after screening if photographs show evidence of potential glaucomatous optic nerve damage. A nonmydriatic camera that avoids the problem of dilation has been used, but stereo photographs cannot be obtained and the reliability of the nonmydriatic camera has not been documented. As with visual field testing, photographers and photographic readers must be trained and quality control has to be tightly and constantly maintained. Photographic assessment of other optic disc characteristics have been shown to have better sensitivity and specificity than cup-to-disc ratio in a clinic-based study in which a glaucoma specialist evaluated the photographs. However, the ability to train and standardize ophthalmologist or technicians to evaluate these more subjective characteristics has not been demonstrated. Screening for optic disc abnormalities using direct ophthalmoscopy by trained technicians is not a viable option because even glaucoma specialists have been unable to produce good agreement by this method.

Fifty percent of subjects were able to complete the Full Field 120 test in 7 minutes, 16 seconds per eye. This is approximately half the time taken to complete the C-30-2 threshold test among subjects without glaucoma. The Full Field 120 test displayed better performance characteristics than a variety of other screening methods and could be completed by about 90% of subjects in a population naive to visual field testing. In spite of this advantage, only half of all glaucoma cases would have been detected if a specificity of 90% was desired. In a glaucoma survey undertaken in St. Lucia, approximately 69% of subjects were able to complete the Full Field 120 test. However, among their more selective population able to complete the test, the sensitivity and specificity of the Full Field 120 was 86% and 84%, respectively.

It is possible that manual Goldmann perimetry was not as sensitive as automated perimetry in detecting glaucoma. In some cases, glaucoma that was de-
tected on the Full Field 120 may have been missed because the Goldmann failed to confirm the defect. However, approximately two-thirds of subjects who tested abnormal on the Full Field 120 tested normal on Goldmann perimetry. This was probably because of the low specificity of automated perimetry, whose cutoff of 17 or more defects was designed to ensure that subjects with true defects were likely to be identified for follow-up testing. Furthermore, because the criterion was nonspecific for glaucoma, visual field defects that did not result from glaucoma would have been detected using the automated criteria but rejected by the Goldmann criteria that were more specific for glaucoma.

Because of the continuing poor results shown by traditional screening parameters or risk factor assessment, some form of visual field testing may be necessary to screen effectively for glaucoma. Although the Full Field 120 took a median of 7 minutes per eye to complete, visual field screening using this test would take at least half an hour for each subject if both eyes were to be tested and the time to seat the subject and to explain and demonstrate the test were added. Several new visual field testing strategies have become available since this study was completed. The C-24-2 test of the Humphrey Field Analyzer provides threshold values but takes less time to complete than the C-30-2 threshold test. The new Fast-Pac software of the Humphrey Field Analyzer uses a thresholding technique that results in faster threshold testing at the expense of some variability in the measurements. In addition, suprathreshold screening strategies like the Full Field 120 that test fewer locations might reduce the time taken for each eye. Other more complex algorithms that use clusters of defective points in specified locations may improve sensitivity and specificity. The completion time and sensitivity and specificity of these tests need to be evaluated in a population-based setting before they can be recommended for clinical practice or glaucoma screening.

Key Words

glaucoma, screening, survey, visual field, perimetry

Acknowledgments

The authors thank John D. Gottsch, Kuldev Singh, Brenda Casey, James Sutton, Valerie Jenkins, Nancy Morrow, Teresa Mckissack, Donna Gilbert, James Martone, Sandi Ezrine, and the staff of Survey Research Associates, Inc., for contributing to this project.

References


