Correlation of the Binocular Visual Field with Patient Assessment of Vision

Henry D. Jampel, David S. Friedman, Harry Quigley, and Rhonda Miller

PURPOSE. To determine which measures of the binocular visual field correlate best with the patient’s assessment of vision.

METHODS. Esterman binocular visual field testing and four other binocular visual field tests (designated peripheral 20 dB [p20], peripheral 22 dB [p22], central 24 dB [c24] and central 26 dB [c26]) were performed in 101 patients with glaucoma or suspected glaucoma. Scores from these five tests, as well as binocular visual field scores calculated from monocular testing (best-location summation and probability summation), were correlated with performance on the National Eye Institute’s Visual Function Questionnaire (VFQ)-25 and Short-Form (SF)-36 quality of life instruments, as well as with the linear rating scale utility test.

RESULTS. The mean percentage of correct responses was 87%, 69%, 59%, 78%, and 71% for the Esterman, p20, p22, c24, and c26 tests, respectively. The distribution of scores was much broader for the p20 and p22 tests than for the Esterman test. The mean decibels for the binocular visual fields calculated from the monocular visual fields were 21.5 ± 7.7 dB for the best-location algorithm and 25.1 ± 6.7 dB for the probability-summation algorithm. The binocular visual field score calculated with the best-location algorithm correlated better with the overall, general vision, distance activities, and peripheral vision domains of the VFQ-25 (partial correlation coefficients of 0.48, 0.48, 0.49, and 0.51, respectively) than did the probability-summation algorithm and all five binocular visual field tests. The best-location algorithm also had the strongest correlation with the linear rating scale utility test (partial correlation coefficient, 0.40).

CONCLUSIONS. In this sample of clinic-based patients with glaucoma or suspected glaucoma, a global score derived from a combination of two monocular fields correlated better with patient assessment of vision than did the Esterman and four novel binocular visual field tests. (Invest Ophthalmol Vis Sci. 2002;43:1059–1067)

Although it is critical to assess the function of each eye individually to determine the presence of ocular disease, our visual world is determined by the input from both eyes to the brain. Therefore, an appreciation of how a person’s vision affects quality of life (QOL) should incorporate binocular visual function, for both central and peripheral vision.

Glaucoma is a disease in which vision loss most commonly starts in the midperiphery, with central vision loss not occurring until late in the disease. An assessment of the binocular visual field in patients with glaucoma, particularly one that included the midperipheral points first lost in the disease, could be a valuable adjunct in understanding the visual limitations that glaucoma causes. Specifically, the degree of binocular visual field loss may correlate with vision-related QOL and mobility skills.

Two strategies can be used to assess the binocular visual field. First, the visual field can be determined with a binocular test, in which both eyes remain open. The concept of using a binocular visual field as a means of assessing visual disability was popularized by Esterman. The Esterman binocular visual field was designed first for manual perimetry, and was later adapted to automated perimeters, and is included as an algorithm on the Humphrey Field Analyzer II (HFA II; Humphrey Systems, San Leandro, CA).

Several investigators have used the Esterman binocular visual field test, as configured for the HFA II, in trying to assess visual disability in patients with glaucoma. All these studies, whether evaluating patients primarily with early glaucoma, or those with more advanced disease, and including a study that we have recently completed have found that most of the scores on the Esterman binocular visual field test are clustered in the 80% and above range. This lack of breadth in Esterman scores makes it difficult to find strong correlations between the Esterman score and other measures of visual function. For example, in our recently completed study, the mean Esterman score was 89.7 ± 13.4, with a maximum score of 100, and the partial correlation coefficient between the Esterman binocular visual field score and the overall score on the National Eye Institute Visual Function Questionnaire (VFQ)-25 was 0.32.

A second strategy for assessing a person’s binocular visual field is to combine the results obtained from testing the visual field of each eye separately. If this can be accomplished in patients with glaucoma who are already undergoing monocular visual field testing to detect and monitor disease, then no additional visual field tests would be necessary to make an assessment of the binocular vision. Recently, Nelson-Quigg et al. reported that using either of two binocular sensitivity-prediction models, known as best location and binocular summation, they could predict the results obtained on a binocular visual field test.

In this study, we sought to develop a method of assessing the binocular visual field in patients with glaucoma that would correlate better with patient-reported function and preference than does the Esterman test, as currently configured on the HFA II perimeter. First, we designed our own binocular visual field tests, using the Esterman visual field as a template, and tested them in a series of subjects with glaucoma or suspected glaucoma. Because the Esterman test uses a relatively bright, fixed stimulus intensity of 10 dB throughout the visual field, we adopted the idea of Harris and Jacobs that the stimulus intensity used in the Esterman test could be decreased to expand the useful range of scores. Furthermore, we tested the central and more peripheral visual field separately. Second, we applied the
best-location- and binocular-summation strategies published by Nelson-Quigg et al.\textsuperscript{13} to the monocular visual field tests of these subjects. Finally, we correlated these binocular visual field assessments with scores from the QOL and linear rating utility instruments.

**METHODS**

**Subjects**

From September 1999 through June 2000 we recruited subjects who were patients in a tertiary glaucoma practice. Records of all patients under observation for glaucoma or suspicion of glaucoma were reviewed. Inclusion criteria were visual acuity of at least 20/40 or better in one eye, age of 21 years or greater, and notation in the medical record of results of a Humphrey automated visual field test in at least one eye performed within the past 9 months. Exclusion criteria were diabetic retinopathy, macular degeneration, history of retinal reattachment surgery, intraocular surgery or laser treatment within the previous 2 months, scheduled intraocular surgery, and optic neuropathy other than glaucoma. Patients not fluent in English or judged mentally unable to complete the study were excluded. The study protocol, which adhered to the tenets of the Declaration of Helsinki and which was reviewed and approved by the Johns Hopkins Institutional Review Board, was explained to each patient. Each participant gave informed written consent.

**Patient Interview Materials**

All interviews were administered face to face by the same experienced interviewer. A comorbidity, medication, and demographics questionnaire—the 25-question VFQ-25\textsuperscript{14,15} and the short form (SF)-36 of the Medical Outcomes Study\textsuperscript{17} were administered in random order. Data on income were estimated from the subject’s Zip Code and United States census information. After the completion of these questionnaires, the patients were administered the linear rating utility scale test.

The linear rating scale was presented in the form of two subjective rating scales, or “feeling thermometers”\textsuperscript{18} (Fig. 1). The first thermom-

![Figure 1](http://iovs.arvojournals.org/pdfs/)

**Figure 1.** “Feeling thermometers” used in the linear rating scale utility test.

![Figure 2](http://iovs.arvojournals.org/pdfs/)

**Figure 2.** Binocular visual field tests. (A) Esterman binocular visual field test. (B) Custom central visual field test grid. (C) Custom peripheral visual field test grid.
eter was labeled ‘ideal vision’ at the top and ‘blind’ at the bottom. The subject placed a marker on the thermometer corresponding to his or her answer to the question, “On a scale of 0 to 100, where 0 represents blindness and 100 represents ideal vision, how would you rate your current vision?” This number is referred to in the Results section as the “linear rating score.” On the second feeling thermometer, a score of 100 represents perfect health and vision, and a score of 0 represents death. The interviewer asked the subject two questions: First, “On a scale where 0 now represents death and 100 represents ideal health and ideal vision, where would you rate your ‘overall health,’ assuming you had ideal vision?” and second, “On a scale where 0 now represents death and 100 represents ideal health and ideal vision, where would you rate being completely blind, assuming you had your same current health?” The subject placed markers on the thermometer corresponding to his or her answers to these two questions.

Because the answers to the first and second questions on the second thermometer are equivalent to the 100 and the 0 on the first thermometer, the rating of vision determined in the first thermometer on a blind-to-perfect vision scale can be expressed on the second thermometer on a life and death scale. This number is referred to in the Results section as the “adjusted linear rating score.”

Visual Field Testing

We examined results of both monocular and binocular visual field tests. Monocular visual field testing was not performed on the day of the study visit; however, all visual fields were performed within 9 months of the date of the interview. We recorded the mean deviation (MD) and pattern SD (PSD) from the printouts of the 24-2 Swedish interactive test algorithm (SITA) fast tests (Humphrey Systems).

During the study visit, we administered five binocular visual field tests on the field analyzer (Field Analyzer II; Humphrey Systems) to each patient with glaucoma or suspected glaucoma. The order of the tests was varied from subject to subject, so that equal numbers of subjects received each visual field test first. Subjects wearing glasses were instructed to wear them for the tests; if not, they were tested without correction. The tests administered were as follows.

**Esterman Binocular Visual Field Test.** The Esterman binocular visual field test on the field analyzer perimeter (Fig. 2, top) uses a grid of 120 test points to examine more than 130° of visual field. It was originally developed for manual perimeters and, similar to its monocular predecessor, gives more weight to the functionally more important parts of the visual field (i.e., central and inferior). The testing strategy plots the visual field exactly as the patient uses his or her own eyes, as a whole binocular unit, without occlusion.

**Custom Central 24-dB and 26-dB Tests.** We created the two central (Fig. 2, middle) and peripheral (Fig. 2, bottom) tests with the custom test option on the field analyzer. Briefly, from the system setup menu, we selected Additional Setup, then Custom Test, then Create Screening Test, and, finally, Full 90°. At this point, the screen directs the insertion of points to be tested and allows for a given decibel intensity to be selected.

These two central tests varied only by the intensity of the stimulus (24 or 26 dB) which was the same for the 60 points tested. The light sensitivity level for these tests and the peripheral tests were chosen based on estimates from monocular visual field testing of what appropriate thresholds would be at this degree of eccentricity and tested on the investigators themselves. The test points were in a rectangle extending from 20° below to 10° above the horizontal midline and 30° on each side of the vertical midline.
TABLE 1. Demographics

| Age (y) | 69.0 ± 12.0 (32–90)* |
| Race (%) | |
| White | 77 |
| Black | 20 |
| Other | 4 |
| Male | 45 |
| Taking glaucoma medications | 79 |
| Previous intraocular surgery | 47 |
| Family history of glaucoma | 46 |
| Family history of blindness from glaucoma | 9 |
| High school education or more | 77 |
| Employed full-time | 30 |
| Income (n = 93; US $) | 50,000 ± 17,000 (12,230–96,414)* |
| Hypertension | 37 |
| Diabetes | 12 |
| Heart problems | 22 |
| Breathing problems | 11 |
| Arthritis | 40 |

Data are number of subjects unless marked otherwise. n = 101.
* Mean ± standard deviation (range).

Custom Peripheral 20-dB and 22-dB Tests. These two tests (Fig. 2C, bottom) varied only by the intensity of the stimulus (20 or 22 dB) which was the same for the 42 points tested. The light sensitivity level was determined in a preliminary study in which the investigators determined what level was at their threshold in the periphery. The test points extended to 57° horizontally, 36° superiorly, and 43° inferiorly.

Data Analysis

Visual acuities were transformed from Snellen acuities to a log minimum angle of resolution (MAR) scale. Acuities of counting fingers, hand motions, light perception, and no light perception were assigned logMARS of 1.5, 2.0, 2.5, and 3, respectively. MD and PSD were obtained from the hard copy printout of the monocular visual field tests. The Advanced Glaucoma Intervention Study (AGIS) scores were calculated by entering the values for the deviation from the age-matched normal control at each point into software designed to calculate the AGIS score.19

The Enterman score was expressed both as the number of correct answers in a total of 120 and as a percentage of correct answers for comparison with the other tests. The scores on the customized tests were expressed as the percentage of correct answers, with a denominator of 56 for the central tests and 42 for the peripheral tests. Composite scores were calculated for the various combinations of the central and peripheral test (central 24 plus peripheral 20, and so forth).

Simulated binocular visual field scores were calculated based on published algorithms.13 Specifically, the absolute sensitivities in decibels from the printout of the subject’s SITA fast 24-2 visual field results were manually entered into a spreadsheet (Excel; Microsoft Corporation, Redmond, WA) and combined according to the best-location and probability-summation algorithms. The spreadsheet had been set up for 30-2 analysis, and the outer ring of points was therefore ignored. This method of analysis yields a simulated binocular visual field, but does not include any overall global analysis of the visual field (Fig. 3). Therefore, we calculated the mean for all points in the binocular visual field and called that the “binocular visual field score.” In addition, the binocular visual field was broken down into the following regions, and mean scores were calculated for each: superior hemifield, inferior hemifield, central 16 points, central 24 points, central inferior 8 points, central superior 8 points, and peripheral ring.

Subscale scores were calculated for the SF-36 and VFQ-25. For the SF-36, the number of items for each subscale ranges from 2 items for social functioning and bodily pain to 10 items for physical functioning.

For each subscale, the raw scores for each item are summed, and the percentage of the maximum score represented by the raw score becomes the subscale score (range, 0–100). For the VFQ-25, the number of items in each subscale ranges from 1 to 4. For subscales with one item, the score on the question (ranging from 0–100) becomes the subscale score. For subscales with more than one item, the scores on each question are averaged to obtain the subscale score. The scores of all 25 items are averaged to obtain an overall score.

All data were entered into a computer (Access software; Microsoft) and uploaded into statistical software (SAS, Cary, NC, and STATA, College Station, TX) for statistical analyses. To determine the strengths of linear associations between continuous variables for the study patients, partial correlation coefficients (PCCs) were calculated based on multiple linear-regression models. Regressions were adjusted for age, race, gender, and visual acuity, when visual acuity was not a primary variable. Regressions were run separately, adjusting for visual acuity in the better and worse eyes and the mean of the visual acuities of the two eyes. When visual acuity was the primary variable, regressions were adjusted only for age, race, and gender. Bootstrap methods were used to test the PCCs when the normality assumption was violated. To test the equality of two PCCs, bootstrap confidence intervals for the difference in two PCCs were used with a Bonferroni correction for multiple comparisons.

Four domain scores of the VFQ-25 (general health, general vision, color vision, and peripheral vision) were derived from the answer to only one question. Therefore, these domains were treated as ordinal variables and a multiple PCC was calculated when determining the strength of the linear association between these domains and other continuous variables.

Because higher values on some tests represented better vision (e.g., VFQ-25 scores) and on other tests (logMAR acuity) represented worse vision, some of the positive correlations have a negative sign in front of them. For clarity, all correlations have been transposed so that a plus denotes a positive correlation and a minus represents an inverse correlation.

Adjusted variable plots were made to exhibit the relationship between the adjusted QOL and utility scores and the adjusted visual field test results, and these plots were reviewed for outlying or influential observations. We performed two types of sensitivity analyses to determine to what degree removing these outlying or influential observations changes the results. First, we removed the subjects who scored in the lowest 5% or 10% on the simulated binocular visual field, and second, we removed subjects based on the calculated DFTIS from the regression models. The DFTIS are standardized differences in predicted values with and without the inclusion of each observation.20 Com-

TABLE 2. Clinical Characteristics of Eyes

<table>
<thead>
<tr>
<th>Clinical Characteristic</th>
<th>Measurement</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>logMAR, better eye</td>
<td>0.07 ± 0.10 (–0.12–0.3)</td>
<td>101</td>
</tr>
<tr>
<td>logMAR, worse eye</td>
<td>0.37 ± 0.65 (0–3)</td>
<td>101</td>
</tr>
<tr>
<td>IOP OD (mm Hg)</td>
<td>17.0 ± 4.6 (1–26)</td>
<td>99</td>
</tr>
<tr>
<td>IOP OS (mm Hg)</td>
<td>17.6 ± 4.7 (9–30)</td>
<td>101</td>
</tr>
<tr>
<td>Vertical c/d ratio, OD</td>
<td>0.77 ± 0.22 (0.05–0.99)</td>
<td>97</td>
</tr>
<tr>
<td>Vertical c/d ratio, OS</td>
<td>0.76 ± 0.24 (0.05–0.99)</td>
<td>100</td>
</tr>
<tr>
<td>MD, better eye</td>
<td>–5.3 ± 6.6 (–23.0–2.1)</td>
<td>99</td>
</tr>
<tr>
<td>MD, worse eye</td>
<td>–10.3 ± 8.8 (–32.3–1.7)</td>
<td>91</td>
</tr>
<tr>
<td>AGIS score, better eye</td>
<td>4.0 ± 5.4 (0–18)</td>
<td>97</td>
</tr>
<tr>
<td>AGIS score, worse eye</td>
<td>7.2 ± 6.2 (0–20)</td>
<td>88</td>
</tr>
<tr>
<td>PSD, better eye</td>
<td>4.2 ± 3.5 (0–13.8)</td>
<td>99</td>
</tr>
<tr>
<td>PSD, worse eye</td>
<td>7.3 ± 4.5 (1.2–16.1)</td>
<td>91</td>
</tr>
</tbody>
</table>

Data are expressed as the mean ± SD, with the range in parentheses. Not all visual field and optic disc measurements were available in all eyes, because of media opacity, which obscured the cup-to-disc ratio, and lack of performance of Program 24-2 (Humphrey, San Leandro, CA) testing in eyes with extremely limited vision.
Commonly used criteria suggest that data values with a DFITS greater than $2 \times \sqrt{n}$, where $n$ is the square root (number of variables in the model/sample size), warrant further investigation. After removal of these subjects, we ran the multiple logistic regressions again. We also studied the adjusted-variable plots to see whether there might be any relationships between the specific degrees of visual field loss and the QOL and utility scores that might not be detected in the multiple logistic regression of the entire subject population.

**RESULTS**

**Demographics**

The mean age of the subjects was 69 years and the study group was 76% white and 45% male (Table 1). Seventy-nine percent were taking eye drops to lower intraocular pressure (IOP) and roughly half had undergone intraocular surgery (either cataract...
or glaucoma surgery) in at least one eye. Forty-six percent reported a family member with glaucoma. The study group was educated, predominantly retired, and moderately affluent. Concurrent systemic chronic diseases were common in our population, with 40% reporting arthritis.

**Clinical Characteristics**

By study inclusion criteria, the Snellen acuity in the better eye had to be 20/40 or better (logMAR ≤ 0.3); acuity ranged from 20/20 to no light perception in the fellow eye. The cup-to-disc (c/d) ratios (0.77 ± 0.22 OD, 0.76 ± 0.24 OS) indicate that the mean amount of glaucoma damage was moderate. Monocular visual field indices demonstrated, on average, mild visual field loss in the better eye and moderate visual field loss in the worse eye (Table 2).

**Binocular Visual Field Testing**

The mean percentage of correct responses on the Esterman test was 87.4, with a distribution markedly skewed toward higher scores (Fig. 4). The percentage of correct responses was lower on the custom tests, ranging from a mean of 78 on the c24 test down to a mean of 59 for the p22 test, and these tests demonstrated a much broader distribution of scores than the Esterman. Combining a subject’s response on a central and a peripheral test did not further broaden the distribution.

The Esterman binocular visual field test took an average of 4.9 ± 0.7 minutes to complete, whereas the c24, c26, p20, and p22 tests took 2.8 ± 0.8, 3.1 ± 0.8, 2.5 ± 0.5, and 2.6 ± 0.6 minutes, respectively.

**Quality of Life Instruments**

The overall score on the VFQ-25 was 81.7 ± 16.2 with a range of 16.5 to 99.2 (Fig. 5). Subscale scores ranged from a low of 68.1 ± 26.5 for general health to 92.8 ± 14.7 for color vision. These scores were similar to those reported in patients with glaucoma who were studied as part of the development of the VFQ-25.14 On the SF-36, subscale scores ranged from 70.2 ± 26.5 for vitality to 88.4 ± 19.3 for social functioning.

**Linear Rating Scale**

On the linear rating scale, subjects rated their vision as 75 ± 18, on a scale where 100 is ideal vision and 0 is total blindness (Table 3). General health was rated as 89 ± 11 and the state of total blindness as 34 ± 28 on the second scale, on which 100 represents ideal health and ideal vision and 0 represents death. Adjusting the linear rating of vision for the patient’s rating of general health and rating of total blindness had no effect: The adjusted linear rating of vision was also 75 ± 18.

**Age, Race, Gender, and Acuity-Adjusted Correlation of Binocular Visual Field Assessments with Utility Instruments**

The results of the regression analyses did not substantially differ when adjusted for visual acuity in the better eye, worse eye, or the mean acuity of both eyes (Table 4). The data presented are the result of the analyses, adjusting for the mean acuity of both eyes.

The correlation between the Esterman and the custom-designed binocular visual field tests (p20, p22, c24, and c26) with the linear rating of vision (Table 4, top row) were poor to fair and similar to each other (all between 0.24 and 0.28). These correlations were very similar for the custom binocular visual field tests and the adjusted linear rating of vision (Table 4, bottom row). Combining the central and peripheral scores (p20 + c24, p20 + c26, p22 + c24, and p22 + c26) did not alter the correlations (data not shown). The monocular measures of vision (MD, logMAR acuity, CPSD, and AGIS scores for the better and worse eyes) had PCCs ranging from 0.20 to 0.32 on the two linear rating scales. (The MD in the better eye was representative of these measures and is included in Table 4).

The correlation of the probability-summation simulated binocular visual field with the linear rating of vision was no different on the visual field tests (0.27 and 0.26 for linear rating and adjusted linear rating, respectively). Even though the results did not reach statistical significance, the correlation of the best-location simulated binocular visual field was higher than the other measures (0.40 and 0.35 for the linear rating and adjusted linear rating, respectively; Fig. 6).

**Age, Race, Gender, and Acuity-Adjusted Correlation of Binocular Visual Field Assessments with VFQ**

The correlations between the binocular visual field assessments and the overall VFQ scores ranged from 0.40 for the p22 up to 0.48 for the best location (Fig. 7), with an Esterman correlation of 0.44. The individual subscales with which the binocular visual field assessments had the strongest correlation were general vision (range, 0.30–0.48), distance activities (range, 0.36–0.49), and peripheral vision (range, 0.40–0.51). Correlations were poorest with general health (range, 0.19–0.39) and ocular pain (range, 0.05–0.25). The best LOCATION summation had the highest correlation of the binocular visual field assessments with the overall VFQ-25 score (0.48) as well as on 6 of the 12 subscales, including general vision (0.48), distance activities (0.49), and peripheral vision (0.51), although the observed differences in correlations did not reach statistical significance (Table 5).

The monocular vision measures that in general correlated best with the VFQ-25 and its subscales were the MD and the AGIS scores, in particular the MD in the worse eye (Table 5).

### Table 3. Linear Rating of Vision

<table>
<thead>
<tr>
<th>Rating</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear rating of vision</td>
<td>75 ± 18 (25–100)</td>
</tr>
<tr>
<td>Rating of general health</td>
<td>89 ± 11 (60–100)</td>
</tr>
<tr>
<td>Rating of total blindness</td>
<td>54 ± 28 (0–100)</td>
</tr>
<tr>
<td>Adjusted linear rating</td>
<td>75 ± 18 (32–100)</td>
</tr>
</tbody>
</table>

Data are expressed as the mean score ± SD, with the range in parentheses.

### Table 4. Correlation of Binocular Visual Field Assessments with Utility Instruments

<table>
<thead>
<tr>
<th></th>
<th>Esterman</th>
<th>P20</th>
<th>P22</th>
<th>C24</th>
<th>C26</th>
<th>PS</th>
<th>BL</th>
<th>MD  (better eye)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear rating</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.28</td>
<td>0.24</td>
<td>0.27</td>
<td>0.40</td>
<td>0.28</td>
</tr>
<tr>
<td>Adjusted linear rating</td>
<td>0.24</td>
<td>0.26</td>
<td>0.27</td>
<td>0.23</td>
<td>0.27</td>
<td>0.26</td>
<td>0.35</td>
<td>0.28</td>
</tr>
</tbody>
</table>

Data are adjusted for logMAR mean. n = 98. Bold values are significant at P ≤ 0.001. PS, probability summation; BL, best location; MD, mean deviation.

---

Downloaded From: http://iovs.arvojournals.org/pdfaccess.ashx?url=/data/journals/iovs/933593/ on 06/24/2017
For instance, the PCC for the MD in the worse eye was comparable to and not statistically different from the best binocular measure (best-location summation) for the overall VFQ-25 score and was higher on the vision-specific role difficulties subscale than any other measure.

**Correlation of Regions of the Best-Location and Probability Summations with QOL and Linear Rating Scale Scores**

The mean sensitivity of each region (superior hemifield, inferior hemifield, central 16 points, central 24 points, central inferior 8 points, central superior 8 points, and peripheral ring) of the visual fields derived from the best-location– and probability-summation algorithms (Fig. 2) were correlated with the QOL and the linear rating scale scores. None of the regions correlated with the QOL and utility tests results as well as the best-location algorithm (data not shown).

**Correlation of Binocular Visual Field Assessments with SF-36**

The correlations between the binocular visual field assessments were low on all tests and all subscales of the SF-36 range, $-0.049$ [p20 and pain] to 0.212 [probability summation and role-physical; data not shown].

**Sensitivity Analyses**

Neither removal of subjects who scored in the lowest 5% or 10% on the simulated binocular visual field nor removal of subjects with a DFITS index greater than $2 \times$ the square root (number of variables in the model/sample size) altered the correlations obtained (data not shown). Inspection of the adjusted variable plots for the multiple logistic regressions of visual function and QOL and utility test scores did not reveal any strong correlations that were not present in the PCCs calculated from the multiple logistic regressions.

**DISCUSSION**

We proposed to devise a binocular visual field test that correlates better with patient-reported assessment of vision than does the Esterman binocular visual field test. Our reason for performing the study was the relatively poor correlation of the Esterman test with patient-reported assessment of vision reported in the literature and confirmed in a previous study of ours. However, in the patients with glaucoma or suspected glaucoma whom we studied, we were unable to demonstrate a correlation of any binocular visual field test that was stronger than the correlation of monocular measures of visual function with patient assessment of vision. In fact, the best-location algorithm for combining sensitivity values from monocular threshold visual field tests$^{13}$ gave stronger correlations than either the monocular or the binocular tests, although the MD in the worse eye visual field was almost as good (results did not reach statistical significance at $\alpha = 0.05$).

We and others$^9$ have speculated that a problem with the Esterman test is its inability to provide a broad enough range of values over the entire spectrum of visual field loss. We found that we were able to obtain a much greater range of responses to our custom binocular visual field tests, but the correlation of our custom test results with the VFQ-25 and utility instrument scores was still no better than with Esterman. Correlation of all test results with the SF-36 were only weak—an expected result, because the SF-36 is not targeted toward visual function.

Our custom tests were nonthreshold, with the threshold sensitivity held constant throughout the visual field. Either a
unknown how well our variation of the Nelson-Quigg algo-
We did not perform binocular threshold testing and thus do

demonstrated a high correlation between the binocular sensitivities
which generally takes 10 to 13 minutes per eye. They demon-
over a statistical combination of the two monocular threshold
tests.
The algorithm of Nelson-Quigg et al.13 combines the results
of full threshold, Program 30-2 testing (Humphrey Systems),
which generally takes 10 to 15 minutes per eye. They demon-
strain a high correlation between the binocular sensitivities
predicted by their models and actual binocular sensitivities.
Our patients, in contrast, had undergone central Program 24-2
testing, using the STA fast algorithm (Humphrey Systems),
a test that takes approximately 3 to 4 minutes per eye. Never-
theless, we used the same algorithms, minus the most per-
ipheral points, to generate a simulated binocular visual field test.
We did not perform binocular threshold testing and thus do
not know how well our variation of the Nelson-Quigg algo-
would approximate a binocular visual field. Neverthe-
less, the predicted binocular visual field correlated better with
the QOL and utility instrument scores than did results of any of
the binocular testing that we performed.
Our study was designed to minimize fatigue that might
influence the patient’s performance on the visual field tests.
Because the protocol required five binocular visual field tests in
one session, we felt that the performance of monocular visual
field testing on the same day would be too likely to tire the
patient. For that reason we used the results of previous visual
fields for the monocular analyses. This study design contains a
bias against correlations between the QOL and utility test and
monocular visual field parameters and the combination of
monocular visual field parameters, because the performance of
the monocular visual fields and the interviews was separated in
time. Nevertheless, we found the strongest correlations using
data derived from the monocular tests.
The algorithms of Nelson-Quigg et al.13 combine the thresh-
old sensitivities of central points and do not include the points
as eccentric as 5° presented in our peripheral tests or the
even more eccentric points tested in Esterman. In terms of
correlating with QOL and utility instruments, the testing of
these more peripheral points does not improve the correlation,
and so consideration of them may be unimportant. However,
although sensitivity in this far peripheral field may not corre-
late with a patient’s responses on testing, the ability to see in
the far periphery may be functionally important (e.g., for see-
ing automobiles rapidly approaching from the side).
Contrary to intuition, binocular visual field performance
may not always correlate well with patients’ assessments of
their visual function. At least in this select group of subjects
with glaucoma, monocular measures, particularly pertaining to
the visual field, correlated as well as binocular measures with
patients’ assessments of vision. Furthermore, visual function in
the worse eye correlated at least as highly with patients’ as-
essment of vision as did visual function in the better eye.
The literature is mixed as to whether patients’ assessment of vision
is more closely tied to visual function in the better or worse
eye. Although several studies suggest that correlation is higher in
the better-seeing eye,21,22 our data corroborate other reports
suggesting the importance of worse-eye function. For instance,
Bass et al.23 reported that cataract patients’ preference values
for their preoperative vision correlated more strongly with
the visual acuity in the worse eye than in the better eye. Turano et
al.15 studied the correlation between clinical measures of vision
in patients with glaucoma and their walking speed through an
obstacle course and found that the MD of the visual field in the
worse eye had the strongest correlation. In a recent report by
Owsley et al.24 for the Impact of Cataracts on Mobility Project,
the risk of automobile crash was more highly correlated with
worse-eye than better-eye function. Finally, in their description
of the development of the VFQ-25, Mangione et al.14 report
little difference between the correlations of the VFQ-25 scores
with visual acuity and visual field in the better and worse eyes.
Clearly, we need a better understanding of the aspects of visual
impairment that affect a patient’s perception of visual function.
The findings in this study suggest that combining the results
of monocular threshold visual field testing into an aggregate
binocular visual field score provides at least as good a represen-
tation of a patient’s assessment of visual function in glau-
coma as does central visual acuity testing or currently available
binocular visual field testing. However, some monocular visual
field parameters performed almost as well. Because the sample
studied was a specialized one of patients attending a glaucoma
clinic in an academic center, our results must be tested and
confirmed in other patient groups. Other databases that may be
readily available for analysis include those of the ongoing
multicenter trials of therapy for glaucoma,3,25 because visual
fields and QOL assessments are routinely performed on an
ongoing basis. Perhaps any battery of objective tests that does
not incorporate multiple facets of vision, such as acuity, con-
trast sensitivity, light and dark adaptation, motion detection,
visual field, and depth perception, for example, fall short of
capturing the totality of our patients’ visual experience. Fur-
ther work is warranted, both to improve our understanding of
the visual requirements necessary for everyday living and to

TABLE 5. Correlation of Binocular Visual Field Assessments with VFQ

<table>
<thead>
<tr>
<th></th>
<th>Esterman</th>
<th>P20</th>
<th>P22</th>
<th>C24</th>
<th>C26</th>
<th>PS</th>
<th>BL</th>
<th>MD (worse eye)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VFQ overall</td>
<td>0.44</td>
<td>0.46</td>
<td>0.40</td>
<td>0.40</td>
<td>0.44</td>
<td>0.42</td>
<td>0.48</td>
<td>0.49</td>
</tr>
<tr>
<td>General health</td>
<td>0.23</td>
<td>0.19</td>
<td>0.25</td>
<td>0.26</td>
<td>0.21</td>
<td>0.39</td>
<td>0.37</td>
<td>0.31</td>
</tr>
<tr>
<td>General vision</td>
<td>0.30</td>
<td>0.38</td>
<td>0.36</td>
<td>0.36</td>
<td>0.41</td>
<td>0.36</td>
<td>0.48</td>
<td>0.37</td>
</tr>
<tr>
<td>Ocular pain</td>
<td>0.25</td>
<td>0.12</td>
<td>0.03</td>
<td>0.17</td>
<td>0.17</td>
<td>0.19</td>
<td>0.23</td>
<td>0.25</td>
</tr>
<tr>
<td>Near activities</td>
<td>0.26</td>
<td>0.23</td>
<td>0.20</td>
<td>0.19</td>
<td>0.19</td>
<td>0.24</td>
<td>0.34</td>
<td>0.29</td>
</tr>
<tr>
<td>Distance activities</td>
<td>0.39</td>
<td>0.45</td>
<td>0.43</td>
<td>0.36</td>
<td>0.41</td>
<td>0.43</td>
<td>0.49</td>
<td>0.45</td>
</tr>
<tr>
<td>Vision-specific social functioning</td>
<td>0.42</td>
<td>0.35</td>
<td>0.27</td>
<td>0.26</td>
<td>0.30</td>
<td>0.38</td>
<td>0.37</td>
<td>0.41</td>
</tr>
<tr>
<td>Vision-specific mental health</td>
<td>0.26</td>
<td>0.31</td>
<td>0.28</td>
<td>0.31</td>
<td>0.35</td>
<td>0.24</td>
<td>0.35</td>
<td>0.37</td>
</tr>
<tr>
<td>Vision-specific role difficulties</td>
<td>0.37</td>
<td>0.35</td>
<td>0.34</td>
<td>0.36</td>
<td>0.51</td>
<td>0.27</td>
<td>0.35</td>
<td>0.37</td>
</tr>
<tr>
<td>Vision-specific dependency</td>
<td>0.32</td>
<td>0.37</td>
<td>0.30</td>
<td>0.26</td>
<td>0.34</td>
<td>0.31</td>
<td>0.29</td>
<td>0.24</td>
</tr>
<tr>
<td>Driving</td>
<td>0.24</td>
<td>0.32</td>
<td>0.35</td>
<td>0.36</td>
<td>0.37</td>
<td>0.29</td>
<td>0.38</td>
<td>0.35</td>
</tr>
<tr>
<td>Color Vision</td>
<td>0.35</td>
<td>0.26</td>
<td>0.26</td>
<td>0.29</td>
<td>0.23</td>
<td>0.26</td>
<td>0.24</td>
<td>0.31</td>
</tr>
<tr>
<td>Peripheral Vision</td>
<td>0.45</td>
<td>0.46</td>
<td>0.41</td>
<td>0.43</td>
<td>0.47</td>
<td>0.40</td>
<td>0.51</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Data are adjusted for logMAR mean. Bold values significant at P ≤ 0.001. PS, probability summation; BL, best location; MD, mean deviation.
elucidate the role that patient-reported assessment of vision should play in the management of glaucoma.

Acknowledgments

The software (Excel, Microsoft Corp.) for these calculations was generously provided by Chris Johnson, PhD. Expert statistical support was provided by Elizabeth Johnson, MS.

References