GLAUCOMA PATIENTS' ASSESSMENT OF THEIR VISUAL FUNCTION AND QUALITY OF LIFE*

BY *Henry D. Jampel, MD, MHS*

ABSTRACT

Purpose: To determine how glaucoma patients with various degrees of vision loss rate their vision, and to determine if the Esterman binocular visual field test and other visual function tests correlate with those ratings.

Methods: Two hundred thirty-seven glaucoma patients evaluated their vision using 2 utility tests, the linear rating scale and the time trade-off test, and 2 quality-of-life instruments, the National Eye Institute Visual Function Questionnaire (VFQ) and the Short Form 36 (SF-36). Their results were compared to clinical tests of their vision and to persons with normal vision (n=12) and blind persons (n=12).

Results: On a scale of 0 (blind) to 100 (ideal), subjects with normal vision rated their vision higher (90 ± 8.0) than did glaucoma subjects and suspects (75.7 ± 17.6) and "blind' subjects (15.6 ± 15.3), *P* =.001. Mean scores for the Esterman test were 89.7 ± 13.4 for the glaucoma group. The Esterman test correlated moderately with the overall VFQ score (partial correlation coefficient [PCC] = 0.32, *P* = .001), but only weakly with the linear rating scale (PCC = 0.17, *P* = .02) and the time trade-off test (PCC = -0.16, *P* = .06). Correlation between the linear rating scale and the overall VFQ score was good (PCC = 0.56, *P* = .0001) and was moderate with several domains of the SF-36 (eg, social function PCC = 0.32, *P* = .0001).

Conclusions: Utility values that glaucoma patients assign to their vision do not correlate well with Esterman test results. A challenge for the future will be designing clinical tests that better correlate with patient perceptions.

Tr Am Ophth Soc 2001;99:301-317

INTRODUCTION AND LITERATURE REVIEW

The glaucomas are a heterogenous group of diseases that have in common a characteristic form of damage to the optic nerve. The damage generally results in typical changes in optic disc morphology and visual field. The primary focus in the care of glaucoma patients has been the prevention of ongoing damage to the optic nerve and consequent visual field loss.

Visual field defects in glaucoma tend to affect the midperipheral visual field first and only later in the disease involve central vision and then fixation. This pattern of visual field loss in glaucoma has led to the impression that the glaucoma patient is asymptomatic until late in the disease. Only when visual field loss impinges upon or involves central vision does the patient become aware of a functional defect.

Objective end points in the management of patients with glaucoma are important and include the level of intraocular pressure, appearance of the optic nerve, and status of the visual field. In addition, over the past several years an increased awareness of the effect of glaucoma on the patient's quality of life (QOL) has developed. This parallels an increased interest throughout ophthalmology in the impact of disease and therapy on QOL. The impact of cataract,[1-4] macular degeneration,[5-9] diabetic retinopathy,[10,11] refractive error,[12-14] corneal disease,[15] and choroidal melanoma[16,17] on QOL have all been evaluated.

QOL in patients with ocular disease can be measured by using either vision-directed instruments or generic instruments designed for examining overall health. Other investigators have described the characteristics and the strengths and weaknesses of these instruments,[8,18,19] and so they will not be described at length herein. Vision-directed instruments described in this thesis include the National Eye Institute Visual Function Questionnaire (VFQ),[20,21] the VF-14[22-24] the Activities of Daily Vision Scale (ADVS),[25,26] and the Glaucoma Symptom Scale;[27] the generic instruments mentioned are all versions of the Medical Outcomes Study Short Form (eg, SF-36).[28-30]

Several investigators have examined QOL in glaucoma patients. Sherwood and coinvestigators[31] examined QOL in glaucoma patients by using a generic instrument. They found that patients with glaucoma had lower scores on all domains of the SF-20 than control subjects, but they did not adjust for general medical comorbidity, which could have influenced results. Similarly, Wilson and colleagues[32] administered the SF-36 instrument to glaucoma

*From the Department of Ophthalmology, Johns Hopkins University School of Medicine, Baltimore, Maryland. Supported by the Glaucoma Research Foundation, San Francisco, California.

patients, people suspected of having glaucoma (glaucoma suspects), and controls (no glaucoma) and found that the scores were lowest for the glaucoma patients, intermediate for the glaucoma suspects, and highest for the controls. However, they did not report the severity of damage among the glaucoma patients, again limiting the conclusions that they could reach.

Other investigators have looked at vision-specific measures of QOL in glaucoma patients. Parrish and coworkers[33] found that there was a moderate correlation between binocular visual field impairment and scores on the VFQ and the VF-14. Mills[19] has reported only weak correlations between QOL and the ocular characteristics at enrollment of subjects in the Collaborative Initial Glaucoma Treatment Study (CIGTS).

Finally, Gutierrez and associates[34] found that greater visual field defects in the better eye of glaucoma patients were associated with poorer scores on the VFQ and VF-14 QOL instruments. Of particular note was the suggestion that there was a linear relationship between visual field loss and QOL responses and that changes in QOL were present even with small amounts of visual field loss. These findings are particularly provocative because they suggest that even early visual field loss may affect patients' QOL and that visual field loss need not threaten fixation nor involve an entire hemifield before affecting the functional well-being of patients. These findings call into question the clinical aphorism that glaucoma is a "sneak thief" of vision and that it has no effect until damage is profound.

Preliminary attempts have been made to design QOL instruments that are specific to the symptoms of glaucoma[27] and the effect of glaucoma therapy on QOL.[35] QOL assessment in glaucoma is becoming increasingly important, as attested to by the incorporation of QOL instruments into the study design of both the Ocular Hypertension Treatment Study (OHTS)[36] and the CIGTS clinical trials.[19]

The QOL instruments that are in common use suffer from an important limitation. While they identify declines in visual function among patients, they fail to elicit from patients how their visual function influences their lives. Two patients with the same score on a QOL instrument may differ substantially in how dissatisfied or distressed they feel about their QOL. For instance, 2 patients may give the same response to a question concerning how difficult it is to perceive colors, and hence their score on this question would be the same. However, for the patient who is a painter, the importance of having a color vision problem may be much greater than for other patients.

To address this aspect of the QOL effects of a disease, it is necessary to measure the preference values that patients assign to their health status,[37] where preference values are defined as "the levels of subjective satisfaction, distress, or desirability that people associate with a particular health state."[38] Several tools, generally referred to as utility tests, have been developed to measure the preference values that patients assign to their health status.[39] This utility approach has several strengths for assessing health-related QOL, including producing a single score on a 0-to-1 scale; incorporating information on risk attitudes, time preference, and trade-offs among different situations; and being able to be combined with pecuniary measure of costs and benefits.[40] On the other hand, there are drawbacks to the utility approach, including lack of precision, the need for labor-intensive interviews, and the requirement of native language ability in the language in which the materials are presented.[40]

Some of the most widely used tools are the standard gamble,[7,41] the time trade-off,[7] and the linear rating scale.[38] The *standard gamble* is the classic method of measuring preferences. The subject is asked what risk he or she would take in order to reach a certain health state. For example, a completely paralyzed patient would be offered a theoretical treatment, which would either cure the paralysis or immediately kill the patient. The subject would determine how much of a risk of death he or she would tolerate for the chance of a cure. The amount of chance taken is used as a measure of how undesirable the patient perceives the present disease state: the greater the risk tolerated, the worse the disease state. The *time trade-off* technique is an alternative to the standard gamble that is simpler to administer.[42] Subjects are asked how many years of their remaining life in their current state of health they would be willing to give up in order to have perfect health for the rest of their life. The *linear rating scale* originated in psychometrics.[39] Subjects are shown a line on a page where 0 at one end represents an undesirable state (death or blindness) and 100 at the other end represents perfect health (or vision). The subject then places his or her assessment of health at the appropriate point along the line. The linear rating scale is the most efficient of the 3 methods to administer but has the disadvantage of not providing direct cardinal utility measures.[43]

The use of utility measures to determine patient preferences in ophthalmology is relatively new, but several interesting studies have been reported. Torrance[44] refers to the state of "being blind, deaf, or dumb" as being given a utility value of 0.39 on a scale where 0 represents death and 1.00 represents perfect health. Bass and colleagues[45] used a linear rating scale to determine how patients awaiting cataract surgery felt about their vision. They found that the patients' preference values for their vision were related to problems in specific aspects of daily life (such as feelings of depression and problems interacting with

people) but generally not to traditional clinical measures of visual acuity. Although the cataract patients' preference values regarding their vision were significantly correlated with a visual function index, the correlation coefficient was only about 0.5, indicating that the assessment of visual functioning used did not fully predict how a patient felt about his or her visual impairment.

Brown, Brown, and colleagues have assessed utility problems in patients with age-related macular degeneration[7] and diabetic retinopathy.[11] In patients with age-related macular degeneration, they demonstrated that patients with visual acuity in the better eye of 20/20 to 20/25 would be willing on average to trade 11% of their remaining life to obtain perfect vision in both eyes, whereas those with visual acuity of counting fingers or light perception in their better eye would be willing to give up 60% of their remaining life. Similarly, for diabetic retinopathy the investigators used both time trade-off and standard gamble techniques to demonstrate that vision loss is associated with a decrease in patients' ratings of their vision.[11] In his American Ophthalmological Society thesis, Brown[8] reported the relationship between visual acuity and utility instruments in a large group of his patients with predominantly retinal disease. In a study to determine whether there are differences in QOL between patients with choroidal melanoma treated with enucleation or radiation therapy, Cruickshanks and associates[17] used the time trade-off measure and found that there was no difference between the 2 treatment groups.

To my knowledge, there have been no studies of glaucoma patient preference values for the visual states associated with glaucoma damage. Given physicians' impression that vision loss in glaucoma patients has little impact early in the disease, and the finding by Gutierrez and associates[34] that patient responses to QOL instruments may become abnormal early in the course of the disease, further investigation into patient perception of the disease process seemed warranted. For this reason, I undertook an investigation into glaucoma patients' perception of the impact of glaucoma on their vision and life.

SUBJECTS AND METHODS

SUBJECTS

Subjects were recruited from 3 glaucoma practices during the period of October 1998 to August 1999. One practice was university-based (site A); one was a university-affiliated, community hospital–based, practice (site B); and the third was a suburban private practice (site C).

Before the beginning of each clinical session, records of all patients who had undergone automated static perimetry and were being followed up for glaucoma or suspicion of glaucoma were reviewed. The eligibility criteria were visual acuity of at least 20/40 or better in 1 eye, age of 21 years or older, and the presence in the medical record of a reliable automated visual field in at least 1 eye within the past 9 months (patients with poor vision in 1 eye or with a perfectly normal fellow eye may not have undergone visual field testing in both eyes). Patients were excluded from consideration if they had diabetic retinopathy, macular detachment, or a history of retinal reattachment surgery, intraocular surgery, or laser treatment within the previous 2 months; were scheduled for intraocular surgery; or were thought to have an optic neuropathy other than glaucoma. Patients whose pupils were pharmacologically dilated were not considered for participation on that day because of the unknown effect of pupil size on the Esterman binocular visual field test. Likewise, patients who were scheduled to undergo automated visual field testing were also not considered that day owing to concerns of fatigue from taking multiple visual field tests in one session. Patients not fluent in English or judged not mentally able to complete the study were excluded.

The charts of all potential patients were flagged. Many potential patients were not approached about participation in the study for logistical reasons, which included lateness of the hour, inability of the patient to stay for participation, lack of availability of the interviewer, and lack of availability of a perimeter. The remaining patients were approached about participation in the study by either the study coordinator or the patient's physician. The study protocol, which had been reviewed and approved by the Institution Review Board governing each center, was explained to each patient. The age, race, sex, and visual acuity of patients who declined participation in the study were recorded. Each participant gave informed written consent.

In addition to those subjects who were glaucoma patients or patients followed up for suspicion of glaucoma, we enrolled 2 additional groups of subjects. The *blind* subjects were patients followed up by the physicians participating in the study. They all had visual acuity recorded as no better than counting fingers in their better eye, and none were able to walk without help because of their visual impairment. All had a history of glaucoma; many had other ocular diseases. The *normal* subjects were patients who came to site A annually for an eye checkup, had no known ocular disease except for refractive error, and had normal acuity and a normal eye examination, including a normal optic disc examination. The *blind* and *normal* subjects were recruited to obtain an estimate of the floor and ceiling for answers to the linear rating scale and time trade-off tests. Therefore, no effort was made to match them with the glaucoma patients and glaucoma suspects in terms of age, race, sex, or any other demographic characteristic. With the exception of 1

"blind" subject, the 12 "blind" subjects and the 12 "normal" subjects were recruited at site A.

INTERVIEW PROCESS

All interview materials were administered face-to-face by the same experienced interviewer. Periodically, with the patient's permission, the interviews were audiotaped and reviewed by the principal investigator for quality assurance.

All subjects (glaucoma, blind, and normal) were given the following questionnaires, which were administered in random order:

- A *comorbidity*, *medication*, and *demographics questionnaire* in which subjects were asked about whether or not they had diabetes, hypertension, heart disease, breathing difficulties, or arthritis, and what medications they were taking for their general health. Subjects were also asked about the highest level of education obtained, family history of glaucoma or blindness, and current job status (employed, unemployed, or retired).
- The 25-question version of the *National Eye Institute VFQ*. This questionnaire was chosen as an instrument to assess how subjects fare with their day-to-day visual tasks. The psychometric properties of the VFQ have been well defined,[21] and its utility as a vision-targeted, health-related QOL survey has been demonstrated.[20] A longer (51-item) version of the VFQ has been used previously in published studies of the effect of glaucoma on QOL.[33,34] The 25-question version has been shown to correlate well with the longer version.[46]
- The *Short Form 36 (SF-36)* of the Medical Outcomes Study. This is a survey that was designed as a generic measure of health status. Multiple studies in many disciplines of medicine have used the SF-36.[18] We chose it so that the scores of our subjects might potentially be compared to those of patients with other diseases.

We initially had planned to use the visual ophthalmic symptoms (FUNC-4) portion of the Glaucoma Symptom Scale designed by Lee and associates[27] as well. However, we decided not to use it because of the substantial number of patients that we studied who had little vision in 1 eye. When piloted, these patients answered "no" to such questions as, "Do you see halos around lights?" with their "non-seeing" eye. Therefore, both patients with perfect vision and those with no vision would have provided the same answer to this question.

After completion of these questionnaires, patients were administered 2 utility tests. Visual props using large print, such as a "feeling thermometer" for the rating scale, were used to present the tasks to decrease the "cognitive burden" placed upon subjects.[43]

- The *linear rating scale* was presented in the form of 2 "feeling thermometers," which were large cardboard props with a "0" at the bottom and "100" at the top (Fig 1). The first thermometer was labeled "ideal vision" at the top and "blind" at the bottom. Subjects were asked 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?" Subjects then placed a marker on the thermometer corresponding to their assessment of vision (in this example, 75). If the subjects asked, they were instructed to rate the vision while wearing glasses. This number is referred to in the "Results" section as the *linear rating*.

Subjects were then asked to turn their attention to the second "feeling thermometer." The interviewer told the subjects that a score of 100 on this thermometer represents perfect health and vision and a score of 0 represents death. The interviewer asked the subject 2 questions: (1) "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?" (2) "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 answers to the first and second questions on the second thermometer (80 and 20, respectively, in this example) are equivalent to the 100 and the 0 on the first

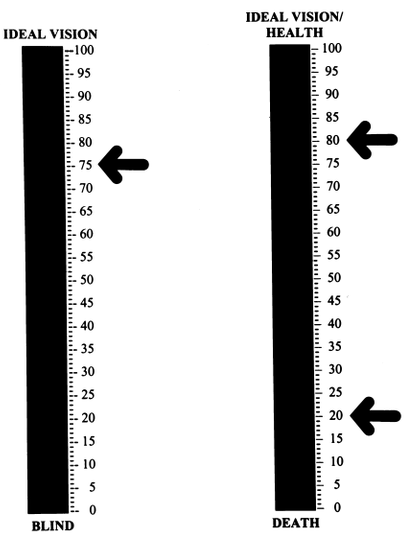

FIGURE 1

Two "feeling thermometer" props used in administration of linear rating scale. On thermometer at left, the subject has rated his vision as 75; on thermometer at right, the subject rated his general health as 80 and blindness as 20.

thermometer. Therefore 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*. In the example illustrated in Fig 1, the adjusted linear rating is 75% of the distance between 20 and 80, yielding a score of 65.

- In the *time trade-off test*, the interviewer first calculated from mortality tables the life expectancy of the individual subject on the basis of age, sex, and race. The subject was then presented with a choice of 2 lives. In the first life, the subject would live for the time equivalent to his or her life expectancy, with his or her current vision. In the second life, the subject would be given ideal vision, but the remaining life would be shorter (Fig 2). Through a series of bracketed questions, the percentage of remaining life that the subject would sacrifice in order to have ideal vision during remaining life was determined.[43] In this example, a patient with a life expectancy of 40 years (life B) would be willing to give up 16 years of remaining life for ideal vision (life A).

One to 6 months later, 13 patients completed the linear rating scale, and 14 patients the time trade-off, a second time, to determine the reproducibility of the tests.

BINOCULAR VISUAL FIELD TESTING

After completion of the interview, we tested each glaucoma patient and glaucoma suspect with the binocular Esterman visual field testing on the Humphrey Field Analyzer II perimeter. The Esterman binocular visual field test was originally developed for manual perimeters, and like its monocular predecessor, it gives more weight to the functionally more important parts of the visual field (ie, central and inferior).[47] The testing strategy plots the visual field exactly as the patient uses his or her eyes, as a whole binocular unit, without occlusion.[48]

The Esterman binocular visual field test has been adapted to automated perimeters.[48-50] On the Humphrey Field Analyzer II, the test uses a grid of 120 test points to examine more than 130 degrees of visual field (Fig 3). Each location is tested once with a size III white stimulus with an intensity of 10 dB. Missed points are retested, and 2 negative responses are recorded as a defect. Stability of fixation is monitored indirectly by observation.[50]

Patients were asked to wear their current refractive correction for distance, if they had their glasses with them.

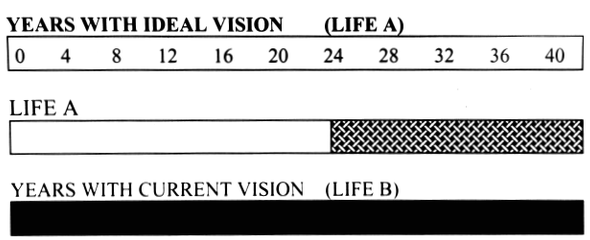

FIGURE 2
Prop used to administer the time trade-off test. In this example, a patient with a life expectancy of 40 years (life B) would be willing to give up 16 years of his remaining life for ideal vision (life A).

CLINICAL RECORD REVIEW

The clinical records of all glaucoma patients and glaucoma suspects were reviewed and their ocular medications, ocular comorbidity (eg, cataract, posterior capsular opacification, diabetic retinopathy), and past ocular surgery were recorded. The Advanced Glaucoma Intervention Study (AGIS) visual field score for each eye was calculated from the subject's most recent automated threshold visual field tests.[51]

DATA ANALYSIS

Visual acuities were transformed from Snellen acuities to logMAR scale. Acuities of counting fingers, hand motions, light perception, and no light perception were assigned logMAR values of 1.5, 2.0, 2.5, and 3, respectively.

Mean deviation and corrected pattern standard deviation (CPSD) were obtained from the hard-copy printout of the visual fields. The AGIS scores were calculated by entering the values for the deviation from age-matched normal at each point into software designed to calculate the AGIS score. The determination of which eye was the "better" eye was based on its mean deviation. The Esterman score was calculated by dividing the number of correct responses by the total number of stimuli (120) and

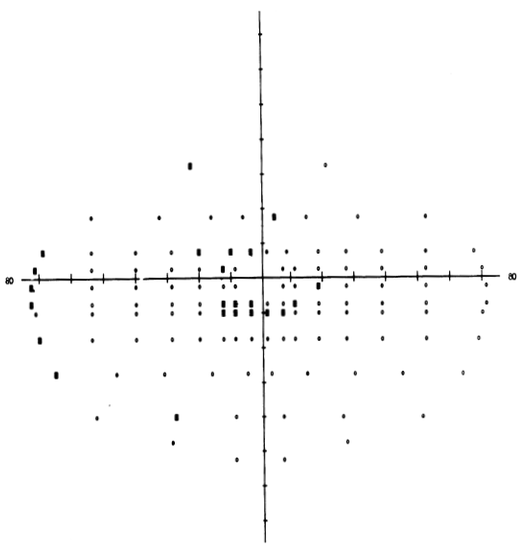

FIGURE 3
Printout of test points of automated Esterman binocular visual field test.

multiplying by 100.

Subscale scores were calculated for the SF-36 and VFQ. The time trade-off score was calculated by dividing the number of years given up by the patient by the life expectancy and multiplying by 100.

All data were entered into Microsoft Access and then uploaded into the statistical packages SAS and STATA for statistical analyses. Descriptive data were tabulated for the study patients as well as the groups of "blind" and "normal" subjects. Analysis of variance (ANOVA) procedures were used to test for differences between continuous variables measured for the study patients across the 3 centers. When assumptions for the ANOVA were violated, the Kruskal-Wallis procedure was used, as in the case of the utility tests and VFQ and SF-36 scores.[52] Chi-square tests were used to test for associations between categorical measures for the study patients across the 3 centers. Similar testing procedures were used to compare the participating and nonparticipating study patients and to compare the 3 groups of subjects, namely, the glaucoma patients and glaucoma suspects, the "blind" subjects, and the normal patients.

To determine the strengths of linear associations between continuous variables for the study patients, partial correlation coefficients (PCCs) were calculated on the basis of multiple linear regression models. Regressions were adjusted for age, race, sex, and the logMAR acuities for both the better and worse eyes, when visual acuity was not a primary variable. When visual acuity was the primary variable, regressions were adjusted only for age, race, and sex. Data quality checks were made using adjusted variable plots generated from the multiple linear regressions. Bootstrap methods were used to test the PCCs when the normality assumption was violated.

To consider the possibility that some associations might not be linear but instead demonstrate a threshold effect, we examined scatterplots of clinical function tests (visual acuity and monocular and binocular visual field scores) plotted against the QOL and utility scores.

Four of the domain scores of the VFQ (general health, general vision, color vision, and peripheral vision) were derived from the answer to only one question, so that the only possible scores were 0, 25, 50, 75, and 100. 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.

In interpreting the values of correlation, it must be appreciated that for some of the visual function tests and utility tests, higher scores signify better vision, and for others, higher scores signify worse vision. Table I lists which variables are positively correlated with better function and which with worse function. This is important in determining whether a correlation is truly positive or negative. For instance, the higher the Esterman binocular visual field score, the better the function, and the lower the time trade-off score, the better the patient's assessment of function. Therefore, the Esterman binocular visual field test and the time trade-off score are correlated when the PCC is negative.

Correlations were considered good if the PCC was between 0.4 and 0.6, fair if between 0.2 and 0.39, and poor if less than 0.2.

RESULTS

DEMOGRAPHICS

Two hundred and thirty-seven patients followed up for glaucoma or suspicion of glaucoma who met eligibility criteria and agreed to participate were enrolled in the study. One hundred and three subjects were seen at the site A, 49 were seen at site B, and 85 were seen at site C.

The mean age of the glaucoma patients and glaucoma suspects was 71 years, 21% were African American, and there was a slight preponderance of women (Table II). One third had a positive family history of glaucoma. Most had at least a high school education and were now retired. Many of them had chronic health conditions, such as arthritis and hypertension. The "normal" subjects were a younger and healthier group, and the "blind" subjects were intermediate in age, had a stronger family history of glaucoma, and had had more glaucoma surgery than the study subjects (glaucoma patients and glaucoma suspects).

The subjects in the 3 centers did not differ in terms of age, race, sex, or use of glaucoma medications, but the subjects from site C were more likely to have undergone surgery for glaucoma (Table III).

Forty-five patients (16%) were invited to participate in the study but declined. The age, race, and sex did not differ between the study participants and the decliners (Table IV). The rate of participation did not differ among the 3 study centers. The visual acuity in the better eye was worse in the nonparticipants than in the participants.

CLINICAL CHARACTERISTICS

The mean logMAR visual acuity was 0.09 ± 0.10 (Snellen equivalent of 20/25) in the study subjects' better eye and 0.48 ± 0.65 (Snellen equivalent of 20/60) in the worse eye (Table V). The mean intraocular pressures were 16.5 ± 4.9 in the right eye and 17.4 + 7.1 mm Hg in the left eye. Not surprisingly, these values were intermediate between the "normal" and the "blind" subjects. The visual field scores in the better eye of –5.3, 4.0, and 4.2, for mean deviation, AGIS score, and CPSD represent on average mild visual field loss; the corresponding scores in the worse eye of –10.6, 7.8, and 6.4 represent on average moderate visual

TABLE I. DIRECTION OF CORRELATIONS*

VARIABLE	LINEAR RATING + SCALES	TIME TRADE-OFF -	VFQ SCORES+	SF-36 SCORES +
Esterman binocular visual field scores +	+	-	+	+
LogMAR scores -	-	+	-	-
Mean deviation +	+	-	+	+
AGIS scores -	-	+	-	-
CPSD -	-	+	-	-
Linear rating scales +			+	+
Time trade-off -			-	-

AGIS, Advanced Glaucoma Intervention Study; CPSD, corrected pattern standard deviation; SF, Short Form; VFQ, Visual Function Questionnaire.
*If a test has a + by its name, then the more positive the test result, the better the visual function or assessment of the visual function. If a test has a - its name, then the more positive the test result, the worse the visual function or assessment of the visual function.

The symbols within the grid are the sign expected when there is a positive correlation between the variable listed down the left side and the variable across the top. For instance, a positive relationship between a logMAR score and the time trade-off test would have a plus sign in front of it, whereas a positive relationship between an AGIS score and a VFQ score would have a negative sign in front of it.

TABLE II. DEMOGRAPHICS

	STUDY SUBJECTS (GLAUCOMA AND SUSPECT) N = 237	"NORMAL" N = 12	"BLIND" (≤ CF ACUITY OU) N = 12
Age	70.6 ± 11.8 [22-92]*	49.7 ± 8.5 [31-63]	63.3 ± 19.5 [30-92]
Caucasian	187 (79%)	12 (100%)	3 (25%)
Male	101 (43%)	10 (83%)	5 (42%)
No. of glaucoma medications	1.3 ± 1.2 [0-4]	0.0	1.3 ± 1.5 [0-4]
No. of glaucoma surgeries	0.8 ± 1.5 [0-10]	0.0	1.9 ± 3.1 [0-10]
Family history of glaucoma	81 (34%)	3 (25%)	7 (58%)
Family history of blindness	30 (13%)	1 (8%)	4 (30%)
Family history of blindness from glaucoma	17 (7%)	0 (0%)	2 (17%)
No. with ≥ high school education	211 (89%)	12 (100%)	8 (67%)
Full-time employed	39 (16%)	8 (67%)	1 (8%)
Hypertension	98 (41%)	2 (17%)	6 (50%)
Diabetes	28 (12%)	0 (0%)	4 (33%)
Heart problems	53 (22%)	0 (0%)	4 (33%)
Breathing problems	43 (18%)	1 (8%)	3 (25%)
Arthritis	136 (57%)	1 (8%)	5 (42%)

CF, counts fingers.
*Mean ± standard deviation [range].

TABLE III. CHARACTERISTICS OF STUDY CENTERS

	SITE A	SITE B	SITE C
No. of Patients	125	49	63
Age	69.0 ± 13.2*	72.7 ± 9.5*	72.2 ± 10.1*
Caucasian	100 (80%)	34 (69%)	53 (84%)
Male	58 (46%)	17 (35%)	26 (41%)
Patients taking glaucoma medication	78 (62%)	33 (67%)	44 (70%)
Patients with previous glaucoma surgery†	33 (26%)	17 (35%)	32 (51%)

* Mean ± standard deviation.
† $P < .001$ for previous glaucoma surgery; $P > .05$ for all other comparisons.

field loss. The mean Esterman binocular visual field score (maximum of 100) was 89.7 (range, 15.8-100).

AGGREGATE PERFORMANCE ON QUALITY OF LIFE AND UTILITY INSTRUMENTS

On the linear rating scale, on which subjects rated their vision on a scale of 0 (blind) to 100 (ideal), the glaucoma patients had scores between the "normal" and the "blind" subjects with means of 90 ± 8.0, 75.7 ± 17.6, and 15.6 ± 15.3 for the "normal", glaucoma, and "blind" subjects, respectively (Table VI). The differences were statistically significant ($P = .0001$). When the same preference was

TABLE IV: COMPARISON OF PARTICIPANTS AND NONPARTICIPANTS

	PARTICIPANTS N = 237	NONPARTICIPANTS N = 45
Age	70.6	71.8
Race		
White	187 (78.9%)	20 (71.4%)*
Black	47 (19.8%)	8 (17.8%)
Hispanic	1 (0.4%)	0
Other	2(0.8%)	0
Sex		
Male	101(42.6%)	18 (40%)
Female	136 (57.4%)	27 (60%)
Site		
A	125 (52.7%)	19 (42.2%)
B	49 (20.7%)	12 (26.7%)
C	63 (26.6%)	14 (31.1%)
Visual acuity		
logMAR better eye†	0.09	0.17
logMAR worse eye	0.48	0.54

* Of 28 known (17 with missing data).
† P = .009 for logMAR better eye; P > .05 for all other comparisons.

transferred to a death (0) to ideal vision (100) scale (adjusted linear rating), the same statistically significant relationship held, with means of 89.6 ± 8.2, 71.4 ± 19.3, and 53.7 ± 24.2, respectively (P= .0003). The ratings of their general health and of the theoretical state of total blindness were lower in the glaucoma subjects than in the other 2 groups, but the difference was not statistically significant (P = .06 and .11, respectively).

On the time trade-off test, none of the 12 "normals" was willing to trade life for improved vision, whereas 6 of 12 "blind" patients were willing to trade some time for ideal vision. On average, the "blind" patients would give up one third of their remaining life for ideal vision. Glaucoma patients were closer to "normals" than to "blind" in this regard, with only 45 of 228 (20%) willing to trade any life for ideal vision.

The scores of all 3 groups were similar on the SF-36 general health perception subset (P = .62, ANOVA), although the glaucoma subjects were again intermediate. The summary scores on the VFQ differed among the 3 groups (P= .0001, ANOVA). As anticipated, they were highest in the "normals" at 92.8 ± 6.4, intermediate for the

TABLE V: CLINICAL CHARACTERISTICS

	STUDY SUBJECTS (GLAUCOMA AND SUSPECT) N = 237*	NORMAL N =12*	BLIND N = 12*	NO. OF SUBJECTS
logMAR better eye	.09 ± .10† [-0.12-0.4]	.01 ± .05 [-0.12-0.1]	1.7 ± 0.5 [0.7-2]	
logMAR worse eye	.48 ± .65 [0-3]	.05 ± .09 [0.00-0.3]	2.5 ± 0.6 [1.5-3]	
IOP OD (mm Hg)	16.5 ± 4.9 [4-40]	15.5 ± 3.5 [10-21]	28.0 ± 11.1 [12-44]	
IOP OS (mm Hg)	17.4 ± 7.1 [3-70]	15.2 ± 2.9 [12-21]	23.3 ± 17.7 [4-58]	
MD, better eye	-5.3 ± 6.4 [-27.9, 3.78]			207
MD, worse eye	-10.6 ± 8.3 [-32.1, 0.73]			197
AGIS score, better eye	4.0 ± 4.9 [0, 20]			211
AGIS score, worse	7.8 ± 6.1 [0, 20]			213
CPSD, better	4.2 ± 3.5 [0, 12.8]			207
CPSD, worse	6.4 ± 4.3 [0, 16.3]			197
Esterman binocular visual field	89.7 ± 13.4 [15.8, 100]			200

AGIS, Advanced Glaucoma Intervention Study; CPSD, corrected pattern standard deviation; IOP, intraocular pressure; MD, median deviation.
*Mean ± standard deviation [range].
†Applies to first 4 rows only.

TABLE VI: AGGREGATE PERFORMANCE ON QUALITY OF LIFE AND UTILITY INSTRUMENTS

	NORMAL N = 12	STUDY SUBJECTS (GLAUCOMA AND SUSPECT)	BLIND N = 12	*P* VALUE
Linear rating of vision	90 ± 8.0	75.7 ± 17.6 (n=234)	15.6 ± 15.3	.0001
Rating of general health	93.3 ± 5.0	81.7 ± 16.8 (n=233)	89.0 ± 13.5	.062
Rating of total blindness	54.6 ± 37.6	37.6 ± 27.7 (n=225)	43.8 ± 28.5	.11
Adjusted linear rating	89.6 ± 8.2	71.4 ± 19.3 (n=223)	53.7 ± 24.2	.0003
Time trade-off (% life remaining)	0.0 ± 0.0	6.1 ± 16.7 (n=228)	33.1 ± 39.2	.048
SF-36 general health perception	69.8 ± 11.9	65.9 ± 19.5 (n=237)	62.2 ± 16.7	.59
Overall score VFQ	92.8 ± 6.4	78.3 ± 14.9 (n=237)	36.6 ± 12.9	.0001

*Mean + standard deviation.

glaucoma patients at 78.3 ± 14.9, and lowest for the "blind" patients, at 36.6 ± 12.9.

REPRODUCABILITY OF THE LINEAR RATING SCALE AND TIME TRADE-OFF TESTS

The intraclass correlation for the responses to the linear rating scale question "How do you rate your vision?" was 0.88 (95% confidence interval, 0.77-1.00) (Fig 4), whereas

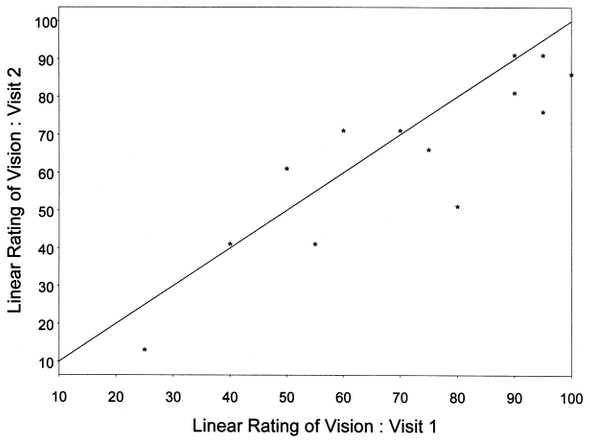

FIGURE 4

Reproducibility of linear rating scale. Thirteen patients performed linear rating scale twice 1 to 6 months apart. Intraclass correlation was 0.88 (95% CI, 0.77-1.00).

the intraclass correlation for the adjusted linear rating score, in which answers are transposed to a death to perfect health and vision scale was 0.78 (95% confidence interval, 0.56-0.99) (data not shown). Therefore, the linear rating scale demonstrated good reproducibility in this small sample.

On the time trade-off test, 9 of 14 subjects were unwilling to trade any time at either visit. Of the other 5 subjects, 3 were willing to give up 10% to 30% of their remaining life on the first visit, but none on the second, whereas one subject gave up 10% on the first visit and 90% on the second visit. Because so many of the answers were 0%, the data concerning reproducibility of the time trade-off test are difficult to interpret.

RELATIONSHIP OF CLINICAL TESTS OF VISUAL FUNCTION TO QOL AND UTILITY TESTS

An important hypothesis tested in this study was that the patient's binocular visual field, as assessed using the Esterman binocular visual field test, would be correlated with patient responses on the utility tests (linear rating and time trade-off). However, the correlation of the Esterman binocular visual field test with the linear rating scale (PCC, 0.17), the adjusted linear rating scale (PCC, 0.17), and the time trade-off test (PCC, –0.14) was poor (Table VII). The correlation of the Esterman binocular visual field test with all domains of the SF-36 was poor (best PCC of 0.18 for the physical function domain), whereas its correlation with the overall score on the VFQ-25 was fair (PCC, 0.32). Correlation was highest (PCC, 0.38) with the vision social function domain of the VFQ-25. Figure 5 contains scatterplots of the Esterman score versus the linear rating score and VFQ-25 summary score and demonstrates clustering of the Esterman scores in the range of 80 to 100.

We next examined whether the other clinical vision measures correlated better with the linear rating and time trade-off tests than did the Esterman. The measure that correlated best with the linear rating scale was the logMAR visual acuity in the worse eye (PCC, –0.30), and the measures correlating best with the time trade-off were the CPSD in the better eye (PCC = 0.24) and logMAR acuity in the worse eye (PCC = 0.22).

TABLE VII: RELATIONSHIP OF QUALITY OF LIFE AND UTILITY TESTS TO CLINICAL TESTS OF VISUAL FUNCTION

	LINEAR RATING	LINEAR RATING (ADJUSTED)	TIME TRADE-OFF	VFQ OVERALL	VFQ (MOST CORRELATED DOMAIN)	SF-36 (MOST CORRELATED DOMAIN)
Esterman binocular	0.17 (.07)*	0.17 (.13)	-0.14 (.31)	0.32 (.001)	0.38 (0.003) VSF	0.18 (.04) PF
LogMAR better eye	-0.15 (.03)	-0.10 (.16)	0.01 (.87)	-0.18 (.004)	-0.27 (.001) DRIVE	-0.11 (.11) VITAL
MD better eye	0.20 (.008)	0.08 (.34)	-0.17 (.02)	0.32 (.001)	0.44 (.001) VSF	0.20 (.007) SF
AGIS score better eye	-0.15 (.03)	-0.07 (.46)	0.10 (.18)	-0.22 (.008)	-0.36 (.001) VSF	-0.17 (.02) SF
CPSD better eye	-0.21 (.003)	-0.05 (.51)	0.24 (.005)	-0.12 (.13)	-0.27 (.001) PV	-0.10 (.20) SF
LogMAR worse eye	-0.30 (.001)	-0.12 (.13)	0.22 (.04)	-0.32 (.001)	-0.32 (.001) VD	-0.10 (.14) GH
MD worse eye	0.13 (.07)	0.14 (.08)	-0.16 (.013)	0.21 (.003)	0.38 (.001) PV	0.04 (.54) SF
AGIS score worse eye	-0.10 (.15)	-0.08 (.30)	0.06 (.36)	-0.22 (.001)	-0.41 (.001) PV	-0.11 (.10) SF
CPSD worse eye	-0.15 (.04)	-0.11 (.15)	0.13 (.01)	-0.03 (.63)	-0.18 (.006) PV	0.16 (.03) PHY

AGIS, Advanced Glaucoma Intervention Study; CPSD, corrected pattern standard deviation; DA, distance activity; DRIVE, driving; GH, general health; MD, mean deviation; PF, physical function; PHY, role physical; PV, peripheral vision; SF, social function; VD, dependence; VITAL, vitality; VSF, vision social function.

*Partial correlation coefficient, with *P* value in parentheses.

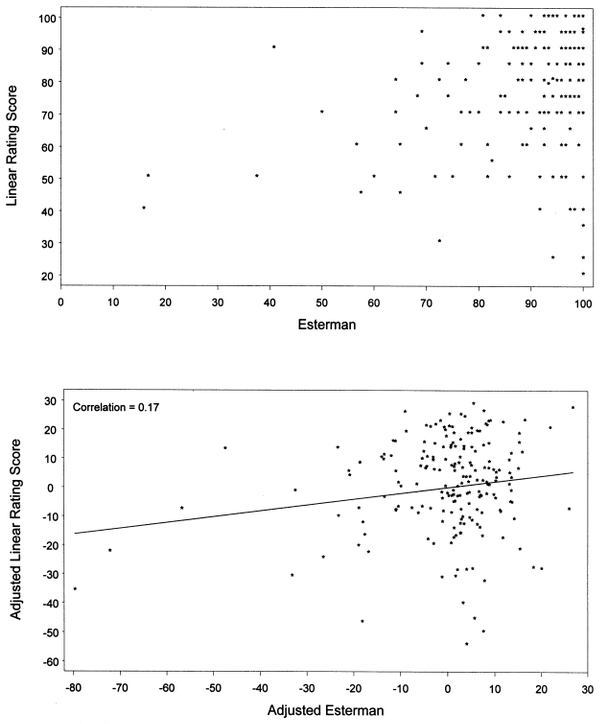

FIGURE 5A

Correlation of Esterman score with scores on linear rating test and VFQ-25 summary score. Top, Scatterplot of Esterman score versus score on linear rating test. Note clustering of Esterman scores between 80 and 100. Bottom, Adjusted variable plot of Esterman score versus linear rating score.

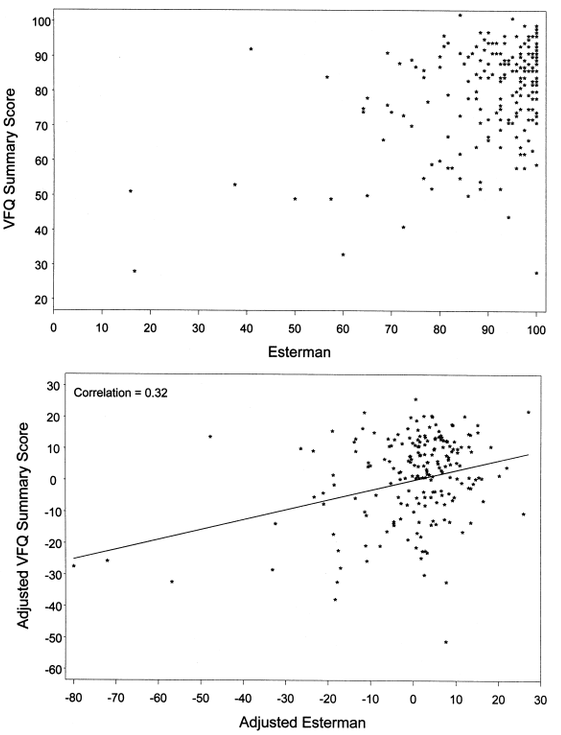

FIGURE 5B

Correlation of Esterman score with scores on linear rating test and VFQ summary score. Top, Scatterplot of Esterman score versus VFQ summary score. Bottom, Adjusted variable plot of Esterman score versus VFQ summary score.

Of all the utility and QOL instruments, the overall VFQ score appeared to have the best overall correlation with the clinical parameters of visual function (Esterman test, visual acuity, and visual field scores). Although not formally statistically analyzed, the ranking of these tests in terms of correlation with the clinical parameters is VFQ overall > linear rating > time trade-off > best SF-36 domain > linear rating adjusted. It should be re-emphasized that none of the correlations were very strong. Within the VFQ, the social function and peripheral vision domains had the strongest correlations with the clinical parameters. The PCC for the social function domain was 0.38 with the Esterman and 0.46 with the mean deviation in the better eye. It generally had a stronger correlation than the overall VFQ score (Fig 6). The single-question peripheral vision domain question also correlated fairly well with several clinical parameters.

The vision tests that appeared to correlate best with the battery of utility and QOL tests were the logMAR in the worse eye (low of –0.12 to high of –0.32), followed by the Esterman binocular visual field test (low of –0.14 to high of 0.38), and the mean deviation in the better eye (low of 0.08 to high of 0.44). The linear rating and VFQ overall scores correlated as highly with the logMAR in the worse eye as with any other measure of visual function. Overall, the clinical tests of the better eye and of the worse eye correlated about equally with the QOL and utility tests.

CORRELATION OF UTILITY TESTS AND THE VFQ-25

The correlation of the linear rating score, (PCC, 0.56) and the adjusted linear rating score (PCC, 0.43) to the overall score on the VFQ was good (Fig 7 and Table VIII). The PCCs of the linear rating scale with the various domains of the VFQ with more than one question (continuous scores) were similar, ranging from 0.42 to 0.47, except for driving (0.30) and ocular pain (0.15). For the domains with only one question (Table IX), the linear rating scale correlated best with the general vision (multiple PCC [MPCC] of 0.55) and peripheral vision (MPCC, 0.43) and poorly with color vision (MPCC, 0.19).

The adjusted linear rating scale had uniformly lower correlation with the VFQ and its domains, with a PCC of 0.43 with the overall VFQ score, and PCCs ranging from 0.26 to 0.36 for the other domains of the VFQ with more than one question, except for ocular pain (0.16). It also had a pattern of correlation similar to the linear rating scale for the domains with only one question, except for general health, which had an MPCC of 0.54.

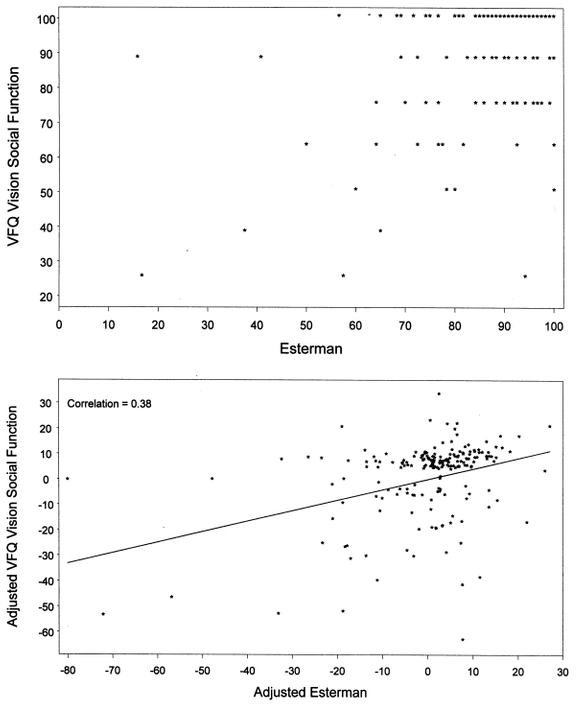

FIGURE 6A

Correlation of Esterman score and mean deviation in better eye with vision social function domain of VFQ. Top, Scatterplot of Esterman score versus score on vision social function domain. Bottom, Adjusted variable plot of Esterman score versus score on vision social function domain.

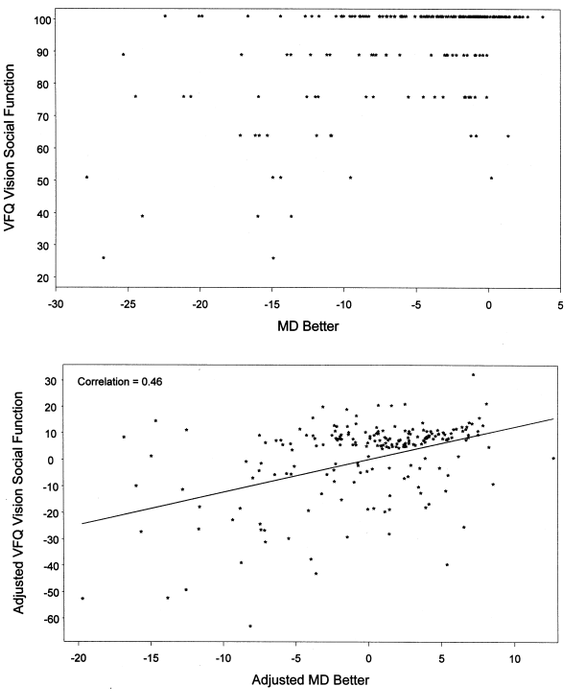

FIGURE 6B

Correlation of Esterman score and mean deviation in better eye with vision social function domain of VFQ. Top, Scatterplot of mean deviation in better eye versus vision social function domain score. Bottom, Adjusted variable plot of mean deviation in better eye versus vision social function domain score.

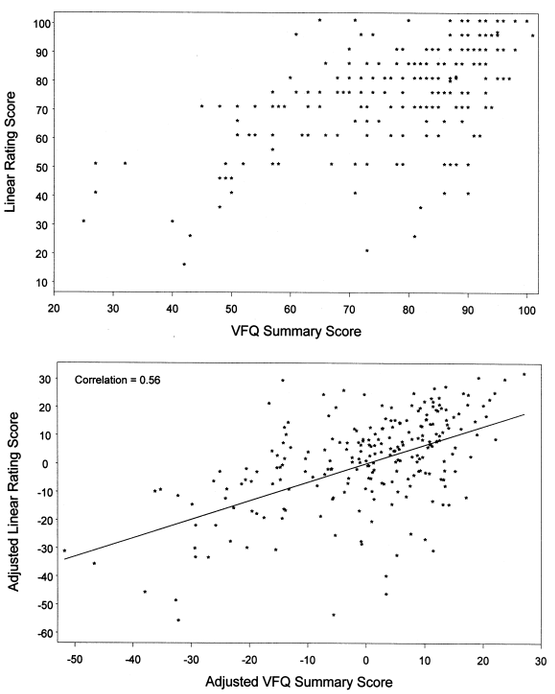

FIGURE 7

Correlation of VFQ summary score with linear rating score. Top, Scatterplot of VFQ summary score versus linear rating score. Bottom, Adjusted variable plot of VFQ summary score versus linear rating score.

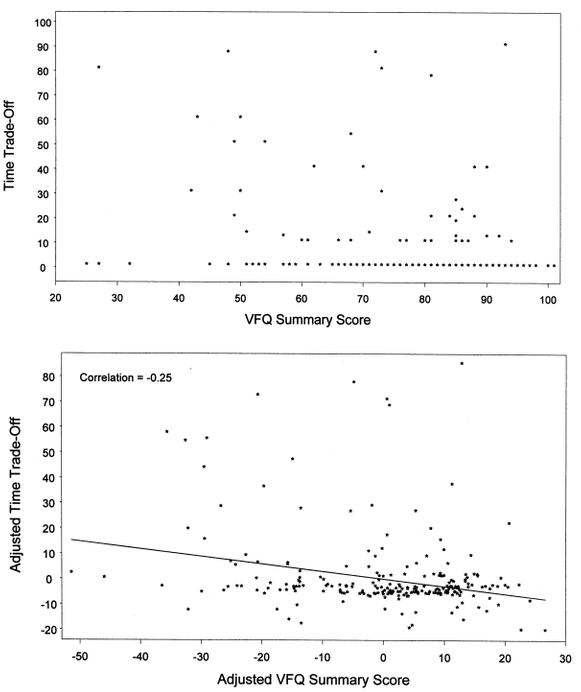

FIGURE 8

Correlation of VFQ summary score with time trade-off score. Top, Scatterplot of VFQ summary score versus time trade-off score. Bottom, Adjusted variable plot of VFQ summary score versus time trade-off score.

TABLE VIII: RELATIONSHIP OF LINEAR RATING SCORES AND TIME TRADE-OFF TO VFQ (DOMAINS WITH CONTINUOUS SCORES)

	LINEAR RATING	RATING OF ONE'S OWN GENERAL HEALTH	RATING OF BLINDNESS	ADJUSTED LINEAR RATING	TIME TRADE OFF
Overall VFQ score	0.56 (.0001)*	0.32 (.0001)	0.14 (.05)	0.43 (.0001)	-0.25 (.001)
VFQ subscales					
Ocular pain	0.15 (.048)	0.14 (.018)	0.04 (.54)	0.16 (.022)	-0.04 (.75)
Near vision	0.47 (.0001)	0.24 (.0002)	0.08 (.35)	0.33 (.0001)	-0.14 (.06)
Distance vision	0.43 (.001)	0.29 (.001)	0.11 (.10)	0.36 (.001)	-0.21 (.001)
Social function	0.46 (.0001)	0.21 (.04)	0.03 (.63)	0.28 (.0002)	-0.26 (.01)
Vision specific mental health	0.45(.0001)	0.16 (.016)	0.16 (.028)	0.30 (.0001)	-0.15 (.006)
Vision specific role difficulties	0.42 (.001)	0.19 (.02)	0.11 (.16)	0.27 (.0001)	-0.21 (.004)
Vision specific dependency	0.46(.0001)	0.19(.02)	0.06 (.39)	0.29(.0001)	-0.24 (.003)
Driving	0.30(.0001)	0.17 (.02)	0.06 (.29)	0.26 (.0001)	-0.25 (.0001)

*Partial correlation coefficient, with *P* value in parentheses

TABLE IX: RELATIONSHIP OF LINEAR RATING SCORES AND TIME TRADE-OFF TO VFQ (DOMAINS WITH ORDINAL SCORES)

	LINEAR RATING	RATING OF ONE'S OWN GENERAL HEALTH	RATING OF BLINDNESS	ADJUSTED LINEAR RATING	TIME TRADE-OFF
General health	0.35 (.0001)*	0.67 (.0001)	0.19 (.002)	0.54 (.0001)	0.25 (.003)
General vision	0.55 (.0001)	0.33 (.0001)	0.16 (.003)	0.43 (.0001)	0.32 (.0001)
Color vision	0.19 (.004)	0.16 (.01)	0.17 (.003)	0.14 (.02)	0.12 (.11)
Peripheral vision	0.43 (.0001)	0.28 (.0001)	0.09 (.06)	0.31 (.0001)	0.29 (.0005)

*Multiple partial correlation coefficient, with *P* value in parentheses.

The correlation of the time trade-off test with the overall VFQ score was fair (-0.25) (Fig 8 and Tables VIII and IX), and correlations with the domains were also lower than for the linear rating scales. Similar to the linear rating scales, correlation was lowest with the ocular pain and color vision domains.

The correlation between the subjects' assessment of their general health on the linear rating scale (second column from the left in Tables VIII and IX) and the overall VFQ score was fair (PCC, 0.32) but correlated highly with the general health question of the VFQ (MPCC, 0.67). The general health question reads, "In general, would you say your overall health is excellent, very good, good, fair, or poor?" The subject's assessment of what a state of total blindness would be like (third column from the left in Tables VIII and IX) was poorly correlated with the overall VFQ score and with the domain scores.

CORRELATION OF UTILITY TESTS AND THE SF-36

The correlation of the linear rating scale with the domains of the SF-36 ranged from poor (mental function, PCC of 0.11) to fair (social function, PCC of 0.32) (Table X). Contrary to the findings with the VFQ, the adjusted linear rating scale had higher correlations with the domains of the SF-36 than did the linear rating scale. In particular, the correlations were good with general health (PCC, 0.53), physical function (PCC, 0.50), and vital function (PCC, 0.46). The time trade-off was poorly correlated with the domains of the SF-36, except for the social function domain, which had a PCC of –0.25.

The subjects' assessment of their general health on the linear rating scale (second column from the left) showed a good correlation with the general health (PCC, 0.61), physical function (PCC, 0.56), and vital function (PCC, 0.53) domains and poor correlation with the mental function (PCC, 0.16) domain.

The subjects' assessment of what a state of total blindness would be like (third column from the left) was poorly correlated with all domains.

NONLINEAR ASSOCIATIONS

To determine if there might be a nonlinear relationship (eg, a threshold visual field loss at which patient assessment of vision markedly decreased), we examined scatterplots of visual acuity and visual field scores as independent variables and QOL and utility scores as dependent variables. No threshold or nonlinear relationships were detected.

TABLE X: RELATIONSHIP OF LINEAR RATING SCORES AND TIME TRADE-OFF TO SF-36

	LINEAR RATING	RATING OF ONE'S OWN GENERAL HEALTH	RATING OF BLINDNESS	ADJUSTED LINEAR RATING	TIME TRADE OFF
General health	0.29 (.0001)*	0.61(.0001)	0.16 (.03)	0.53(.0001)	-0.08 (.21)
Role–physical	0.24 (.002)	0.32 (.0001)	0.07 (.35)	0.29 (.0001)	-0.15 (.17)
Physical function	0.28 (.0001)	0.56 (.0001)	0.14 (.04)	0.50 (.0001)	-0.10 (.27)
Role–emotional	0.17 (.04)	0.21 (.001)	0.01 (.83)	0.20 (.004)	-0.15 (.17)
Social function	0.32 (.0001)	0.44 (.0001)	0.08 (.32)	0.37 (.0001)	-0.25 (.005)
Pain	0.15 (.01)	0.39 (.0001)	0.11 (.06)	0.30 (.0001)	-0.06 (.26)
Mental function	0.11 (.04)	0.16 (.004)	0.05 (.38)	0.14 (.03)	0.02 (.94)
Vital function	0.28 (.0001)	0.53 (.0001)	0.15 (.04)	0.46 (.0001)	-0.14 (.07)

*Partial correlation coefficient, with *P* value in parentheses.

DISCUSSION

This study confirms and adds to the growing literature on QOL in glaucoma patients. The unique aspect of this study is the addition of utility testing to patient evaluation. To the best of our knowledge, this has not previously been done in glaucoma subjects and suspects.

Two utility tests were studied: the linear rating scale and the time trade-off test. On the linear rating scale, we asked subjects to rate their vision on a blindness to ideal vision scale (a number we referred to as the linear rating score) as well as to rate their general health and the state of blindness on a death to ideal health scale (adjusted linear rating score). This "cascading" technique theoretically allows the transformation of the perception of vision loss to a death and perfect health scale, which then can be directly compared with utility values derived from other disease states, such as angina and arthritis. For instance, Torrance and Feeny[53] have reported utility values for mild angina as 0.90, home dialysis as 0.64, and blindness as 0.39. Bass and associates[45] reported a utility value of 0.68 for vision in a cohort of patients about to undergo cataract surgery. On the adjusted linear rating scale, we found that glaucoma subjects and suspects rated the utility of their vision as 0.71 and that the blind rated the utility of their visual state as 0.54. Given different patient populations, variation in technique, and the fact that our "blind" patients still had some residual vision, our results seem comparable to those of other investigators.[45,53] In fact, our glaucoma suspects and patients rated the state of blindness with a utility of 0.38, remarkably close to that reported by Torrance and Feeny[53] and to the value of 0.33 reported by Bass and associates.[45] The correlations of the "adjusted" linear rating scores with tests of visual function and QOL measures scores were similar to the "unadjusted" linear rating scores.

In contrast to the linear rating scale, the findings on the time trade-off test were less informative. Eighty-six per cent of glaucoma subjects or suspects in the current study were unwilling to give up time for improved vision, calling into question the discriminative power of the time trade-off test in this population. This is in contrast to the findings of Brown and colleagues,[8] who reported that 43% of 81 patients with visual acuities of 20/20 or 20/25 were willing to give up at least some portion of their remaining life to achieve perfect vision on the time trade-off test and that, on average, patients with 20/20 acuity in the better eye were willing to give up 8% of their remaining life. Reasons why some of their subjects with good acuity were willing to trade time while our subjects would not include the following: the quality of vision in some of Brown's patients might have been poor because of distortion or decreased contrast sensitivity; the peripheral vision in those patients was poor; the worse eye function was of unrecognized importance; and the time trade-off questions were posed to the patients in different ways in that study and in the current study.

One of the major hypotheses tested in this study was that visual function as assessed on a binocular visual field test would correlate well with patients' evaluation of their vision on utility testing. This hypothesis was not borne out by the data. In fact, the correlation of the Esterman binocular visual field test with both the linear rating scale (PCC, 0.17), the adjusted linear rating scale (PCC, 0.17) and the time trade-off test (PCC, –0.14) was weak. Furthermore, the other clinical measures we examined fared little better. The visual function measures that correlated most highly with the utility tests (but still only with fair correlation) were the logMAR visual acuity in the worse eye (PCC of –0.30 for linear rating and 0.22 for time trade-off) and CPSD in the worse eye (PCC of –0.19 for linear rating and 0.26 for the time trade-off).

There are several possible reasons why I might have failed to find a strong correlation between either binocular visual field testing or monocular tests of vision and patients' assessment of their vision. First, the utility test might simply not have been able to distinguish persons with better vision from those with worse vision. This does

not seem likely for the linear rating scale because we also studied a small number of patients without ambulatory vision and a small number of patients with normal vision to determine the ceiling and floor effects of our utility instruments. The fact that the normals had higher scores on the utility instruments (eg, mean of 90 ± 8.0 on the linear rating scale) and the blind much lower scores (mean of 15.6 ± 15.3 on the linear rating scale) than our glaucoma suspects and patients (mean of 75.7 ± 17.6 on the linear rating scale) provided encouragement that the tests had the potential of correlating with varying degrees of visual field loss. Furthermore, the reproducibility of the linear rating scale was good.

A second reason is that there may not be a close relationship between visual function and patient perception of that function. Perhaps, in contrast to the implications of the findings of Gutierrez and associates,[34] early visual field loss really does not affect patients' assessment of their vision and, in fact, is more of an all-or-none phenomenon, with patients only noticing marked visual field loss. However, we found no evidence for such a threshold effect.

Third, there might be a strong correlation between visual function and its perception, but the best tests for evaluating either function or the value of that function to the patient have not been developed. In terms of functional tests, we suspect that the reason that the correlation between the Esterman test and other parameters is only poor to fair is the lack of a broad range of values on the Esterman test in this study. Most of the scores on the Esterman binocular visual field test were clustered in the 80% and above range, which would make it extremely difficult to find a strong and meaningful correlation between the Esterman score and QOL and utility scores, even if one existed. This same clustering of scores toward 100 (perfect) on the Esterman test was also observed by Parrish and colleagues[33] and by Harris and Jacobs.[54] The latter investigators suggest that the stimulus intensity used in the Esterman test could be decreased to expand the useful range of scores. We are in the process of developing binocular visual field tests that will hopefully distribute the scores more widely and hence be more sensitive to varying degrees of impairment.

In terms of improving the utility tests, we are evaluating a modification of the time trade-off test that has shown preliminary promise.[55] In this variant, the patient is asked how many hours of wakefulness per day he or she would be willing to sacrifice in order to have an ideal state of vision for the rest of the day. This may be an easier concept for the patient, and hence more accurate, than giving up years of life to attain better vision.

In addition to correlating the results of clinical tests of visual function with the utility tests, we also correlated the clinical tests with the QOL instruments. Scores for the VFQ, a vision-directed QOL instrument, were highest for the "normals," intermediate for glaucoma suspects and patients, and lowest for the blind subjects, but these groups did not differ on the scores for the SF-36, a generic QOL instrument. Although this study was not designed to determine if these QOL instruments could distinguish between persons with and without glaucoma and/or visual impairment, our findings in this regard are similar to those of Parrish and coworkers,[33] Gutierrez and associates,[34] and Sherwood and coinvestigators[31] for glaucoma and Cole and associates[46] for optic neuritis. The study of Wilson and colleagues[32] differed in that they found that glaucoma subjects scored worse than glaucoma suspects and controls on most domains of the SF-36, but the differences were small.

We also correlated the results of the Esterman binocular visual field test with VFQ and the SF-36. It is interesting to compare our results with the 2 other studies (Parrish and coworkers,[33] Mills[19]) that have evaluated binocular visual fields. Parrish and coworkers[33] used the same Esterman binocular visual field as we did. Our correlations between the Esterman and the VFQ were similar, if perhaps a little weaker, than theirs. Our correlation with the overall VFQ score was 0.32, whereas their correlations (their Table V), after correction for visual acuity, were in the same range for most subscales. Similar to Parrish and coworkers, we found that correlation of the binocular visual field test with the SF-36 subscales was universally poor, an anticipated finding, since the SF-36 does not target visual problems. Although Parrish and coworkers used the self-administered 51-item VFQ and we used the 25-item test given by an interviewer, it is unlikely that this difference accounts for the minor differences between our findings, because Cole and associates[46] have shown that the 51-item and 25-item VFQs give similar results.

Mills,[19] in his AOS thesis, also investigated the correlation between binocular field of vision and QOL instruments in newly diagnosed subjects with open angle glaucoma in the CIGTS. He combined the monocular visual field scores from each eye to arrive at a simulated binocular score, but this did not correlate well with either the vision-related QOL instrument (VAQ), or the generic QOL instrument (SIP) used in the CIGTS. Furthermore, when the Esterman test was integrated into the CIGTS protocol about 4 years into the study, he compared those Esterman scores with the subjects' responses to the QOL instruments upon entry into the study. Again the correlation was poor. Mills speculated that the lack of correlation between visual function and responses on the QOL instruments might have been due to the mild nature of the visual loss in the CIGTS patients at the beginning of the study. Since visual field loss was not a requirement for entrance into the study, 30% of subjects had no visual field loss at all. Subjects needed to

have 20/40 or better vision in both eyes to be eligible. If the visual function of the subjects were clustered toward the normal, it would be difficult to show a strong correlation between visual function and QOL.

Although the correlations between the clinical tests of visual function and the QOL and utility tests were not great, some clinical tests appeared more highly correlated overall than others. The vision test that appeared to correlate best overall with the utility tests and QOL instruments was the visual acuity in the worse eye (weakest correlation –0.10, strongest correlation –0.32). The linear rating score and the overall VFQ-25 score correlated as highly with the visual acuity in the worse eye (PCCs of –0.30 and –0.32, respectively), as with any other measure of visual function tested. This finding differs from those of Brown[8] and Steinberg and associates[24] in the cataract PORT study, who found that the relationship between acuity and QOL and utility instruments was stronger for the better eye than the worse eye. However, Bass and colleagues[45] 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. Furthermore, Turano and Rubin[56] studied the correlation between clinical measures of vision in glaucoma patients and their walking speed through an obstacle course and found that the mean deviation of the visual field in the worse eye had the strongest correlation. They could not explain their counterintuitive result. Analyzing the aggregate visual function and QOL/utility tests, correlations appeared to be about equal for the better and worse eyes. It should be emphasized that in none of these studies was the correlation with either better or worse eye particularly good.

Examining the opposite question of which QOL or utility instrument had the best correlation with the battery of clinical parameters of visual function that we tested, the overall VFQ score appeared to be best. Within the VFQ, the social function domain had the best correlation of all with the clinical parameters (0.38 with the Esterman and 0.46 with the mean deviation in the better eye) and was higher in general than the overall VFQ score. The social function domain comprises the following 2 questions:

- Because of your eyesight, how much difficulty do you have seeing how people react to things you say?
- Because of your eyesight, how much difficulty do you have visiting with people in their homes, at parties, or in restaurants?

This correlation supports the recent report of Klein and colleagues[57] that visual difficulties in common daily activities, such as reading a menu in dim light and finding a movie seat, were commonly reported by adults.

The final set of correlations examined was between the utility tests and the QOL instruments. As might be expected, the correlation between the linear rating scale was particularly good with the overall VFQ score and with the subscale of general vision. Given the similarity of the question asked in the linear rating scale and that in the general vision subscale, it is somewhat surprising that the correlation is not even greater. Subscales of the VFQ that correlated much more weakly with the linear rating scale were ocular pain, driving, and color vision. This finding is expected, since pain in particular should have little to do with an assessment of vision and lends credence to the concept that the VFQ and linear rating scale are testing similar things.

Although the correlation of the linear rating scale with the domains of the SF-36 was not good, the adjusted linear rating scale had higher correlations in general, with good correlations with general health, physical function, and vital function. Furthermore, subjects' assessment of their general health on the linear rating scale showed good correlation with these same domains but poor correlation with the mental function domain, suggesting that subjects tended to give the same answer to similar questions asked in 2 different ways.

CONCLUSIONS AND RECOMMENDATIONS

The search goes on to understand better the significance to our patients of vision loss from glaucoma. The Esterman binocular visual field is a short test available for automated perimetry that is readily accepted by patients. I theorized that if the Esterman test correlated well with patients' assessment of their own visual function, it might be useful for determining the impact of glaucoma on our patients' lives. I tested this hypothesis but found that the correlation was generally weak.

One explanation that may in part explain the weak correlation is that the Esterman test, as currently configured for automated perimetry, is insensitive to early and moderate degrees of vision loss. This would make it difficult to correlate with any measure of glaucoma damage, whether functional or quality of life. To test the validity of this recommendation, my colleagues and I are in the process of testing several alternative binocular visual field tests that have been modified from the Esterman test. These tests employ stimuli that appear to be closer to the threshold values for normals at peripheral points than the 10-dB stimulus used in the Esterman. It is our hope that one of these tests, or a combination of these tests, will correlate strongly with patient responses on the utility and QOL instruments. This correlation could take the form of a linear relationship or one in which no change in response to these instruments is seen until there is a threshold amount of vision loss.

However, a lack of correlation between clinical tests

of visual function and QOL or utility tests does not necessary mean that either the clinical tests or the QOL and utility tests are invalid. Each test may simply be testing different effects of a disease upon patients, and therefore, because they do not highly correlate, the tests are providing complementary information. It is unlikely that one test, be it visual function, assessment of QOL, or utility measurement, will yield the entire truth about the effect of glaucoma damage on our patients. Rather, it is more likely that any one test will yield only partial truths about the impact of this chronic disease upon patients. Clinical investigators should improve and refine existing tests and continue to develop newer tests to advance our understanding of the impact of glaucoma. Ultimately, however, it will be the ability of clinicians to integrate the results of multiple disparate tests with discussions with the patient that will determine how well they understand the effect of glaucoma on that patient. This understanding is critical in formulating a therapeutic strategy for each patient.

ACKNOWLEDGEMENTS

The author would like to acknowledge Rhonda Miller, COA, who performed all the interviews with incredible spirit and good humor; Elizabeth Johnson for superior statistical support; Drs Irvin Pollack, Donald Abrams, Arthur Schwartz, Howard Weiss, and Ronald Richter for letting us interview patients from their practices; Drs Irvin Pollack and Paul Lichter for nominating me; and Drs Harry Quigley, David Friedman, Nathan Congdon, and Donald Zack for their input in study design and suggestions for improving the manuscript. Lastly, I would like to thank my wife, Risa, and our children, who, whether they knew it or not, gave me the time to write this thesis.

REFERENCES

1. Steinberg EP, Tielsch JM, Schein OD, et al. National study of cataract surgery outcomes. Variation in 4-month postoperative outcomes as reflected in multiple outcome measures. *Ophthalmology* 1994;101:1131-1140.
2. Mangione CM, Phillips RS, Lawrence MG, et al. Improved visual function and attenuation of declines in health-related quality of life after cataract extraction. *Arch Ophthalmol* 1994;112:1419-1425.
3. Javitt JC, Steinberg EP, Sharkey P, et al. Cataract surgery in one eye or both. A billion dollar per year issue. *Ophthalmology* 1995;102:1583-1592.
4. Javitt JC, Wang F, Trentacost DJ, et al. Outcomes of cataract extraction with multifocal intraocular lens implantation: Functional status and quality of life. *Ophthalmology* 1997;104:589-599.
5. Mangione CM, Gutierrez PR, Lowe G, et al. Influence of age-related maculopathy on visual functioning and health-related quality of life. *Am J Ophthalmol* 1999;128:45-53.
6. Scott IU, Schein OD, West S, et al. Functional status and quality of life measurement among ophthalmic patients. *Arch Ophthalmol* 1994;112:329-335.
7. Brown GC, Sharma S, Brown MM, et al. Utility values and age-related macular degeneration. *Arch Ophthalmol* 2000;118(1):47-51.
8. Brown GC. Vision and quality of life. *Trans Am Ophthalmol Soc* 1999;97:473-511.
9. Brown GC, Brown MM, Sharma S. Difference between ophthalmologists' and patients' perceptions of quality of life associated with age-related macular degeneration. *Can J Ophthalmol* 2000;35:127-133.
10. Sharma S. The cost-effectiveness of grid laser photocoagulation for the treatment of diabetic macular edema: Results of a patient-based cost-utility analysis. *Curr Opin Ophthalmol* 2000;11:175-179.
11. Brown MM, Brown GC, Sharma S, et al. Utility values and diabetic retinopathy. *Am J Ophthalmol* 1999;128:324-330.
12. Rose K. Quality of life in myopia. *Br J Ophthalmol* 2000;84:1031-1034.
13. Freitas C, Oliveiros BM, Marques E, et al. Effect of photorefractive keratectomy on visual functioning and quality of life. *J Refract Surg* 1995;11(Suppl):S327-334.
14. Vitale S, Schein OD, Meinert CL, et al. The refractive status and vision profile: A questionnaire to measure vision-related quality of life in persons with refractive error. *Ophthalmology* 2000;107:1529-1539.
15. Courtright P. Visual function among corneal disease patients waiting for penetrating keratoplasty in British Columbia. *Ophthalmic Epidemiol* 1998;5:13-20.
16. Melia BM, Moy CS, McCaffrey L. Quality of life in patients with choroidal melanoma: A pilot study. *Ophthalmic Epidemiol* 1999;6:19-28.
17. Cruickshanks KJ, Fryback DG, Nondahl DM, et al. Treatment choice and quality of life in patients with choroidal melanoma. *Arch Ophthalmol* 1999;117:461-467.
18. Lee BL, Wilson MR. Health-related quality of life in patients with cataract and glaucoma. *J Glaucoma* 2000;9:87-94.
19. Mills RP. Correlation of quality of life with clinical symptoms and signs at the time of glaucoma diagnosis. *Trans Am Ophthalmol Soc* 1998;96:753-812.
20. Mangione CM, Berry S, Spritzer K, et al. Identifying the content area for the 51-item National Eye Institute Visual Function Questionnaire: Results from focus groups with visually impaired persons. *Arch Ophthalmol* 1998;116:227-233.
21. Mangione CM, Lee PP, Pitts J, et al. Psychometric properties of the National Eye Institute Visual Function Questionnaire (NEI-VFQ). NEI-VFQ Field Test Investigators. *Arch Ophthalmol* 1998;116:1496-1504.
22. Linder M. Validity of the visual function index (VF-14) in patients with retinal disease. *Arch Ophthalmol* 1999;117:1611-1616.
23. Cassard SD, Patrick DL, Damiano AM, et al. Reproducibility and responsiveness of the VF-14. An index of functional impairment in patients with cataracts. *Arch Ophthalmol* 1995;113:1508-1513.
24. Steinberg EP, Tielsch JM, Schein OD, et al. The VF-14. An index of functional impairment in patients with cataract. *Arch Ophthalmol* 1994;112:630-638.
25. Mangione CM, Phillips RS, Lawrence MG, et al. Improved visual function and attenuation of declines in health-related quality of life after cataract extraction. *Arch Ophthalmol* 1994;112:1419-1425.
26. Mangione CM, Phillips RS, Seddon JM, et al. Development of the "Activities of Daily Vision Scale." A measure of visual functional status. *Med Care* 1992;30:1111-1126.
27. Lee BL, Gutierrez P, Gordon M, et al. The Glaucoma Symptom Scale. A brief index of glaucoma-specific symptoms. *Arch Ophthalmol* 1998;116:861-866.
28. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473-483.
29. Tarlov AR, Ware JE Jr, Greenfield S, et al. The Medical Outcomes Study. An application of methods for monitoring the results of medical care. *JAMA* 1989;262:925-930.

30. McHorney CA, Ware JE Jr, Lu JF, et al. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care* 1994;32:40-66.
31. Sherwood MB, Garcia-Siekavizza A, Meltzer MI, et al. Glaucoma's impact on quality of life and its relation to clinical indicators. A pilot study. *Ophthalmology* 1998;105:561-566.
32. Wilson MR, Coleman AL, Yu F, et al. Functional status and well-being in patients with glaucoma as measured by the Medical Outcomes Study Short Form-36 questionnaire. *Ophthalmology* 1998;105:2112-2116.
33. Parrish RK, Gedde SJ, Scott IU, et al. Visual function and quality of life among patients with glaucoma. *Arch Ophthalmol* 1997;115:1447-1455.
34. Gutierrez P, Wilson MR, Johnson C, et al. Influence of glaucomatous visual field loss on health-related quality of life. *Arch Ophthalmol* 1997;115:777-784.
35. Barber BL, Strahlman ER, Laibovitz R, et al. Validation of a questionnaire for comparing the tolerability of ophthalmic medications. *Ophthalmology* 1997;104:334-342.
36. Gordon MO. The Ocular Hypertension Treatment Study: Design and baseline description of the participants. *Arch Ophthalmol* 1999;117:573-583.
37. Kassirer JP. Incorporating patients' preferences into medical decisions. *N Engl J Med* 1994;330:1895-1896.
38. Froberg DG, Kane RL. Methodology for measuring health-state preferences—IV: Progress and a research agenda. *J Clin Epidemiol* 1989;42:675-685.
39. Froberg DG, Kane RL. Methodology for measuring health-state preferences—II: Scaling methods. *J Clin Epidemiol* 1989;42:459-471.
40. Feeny DH, Torrance GW. Incorporating utility-based quality-of-life assessment measures in clinical trials. Two examples. *Med Care* 1989;27:S190-S204
41. Green C. Valuing health-related quality of life. A review of health state valuation techniques. *Pharmacoeconomics* 2000;17(2):151-165.
42. Torrance GW, Thomas WH, Sackett DL. A utility maximization model for evaluation of health care programs. *Health Serv Res* 1972;7:118-133.
43. Furlong W, Feeny D, Torrance GW, et al. Guide to design and development of health-state utility instrumentation. McMaster University CHEPA Working Paper Series 90-9, 1990:1-141.
44. Torrance GW. Utility approach to measuring health-related quality of life. *J Chronic Dis* 1987;40:593-603.
45. Bass EB, Wills S, Scott IU, et al. Preference values for visual states in patients planning to undergo cataract surgery. *Med Decis Making* 1997;17:324-330.
46. Cole SR, Beck RW, Moke PS, et al. The National Eye Institute Visual Function Questionnaire: Experience of the ONTT. Optic Neuritis Treatment Trial. *Invest Ophthalmol Vis Sci* 2000;41(5):1017-1021.
47. Esterman B. Functional scoring of the binocular field. *Ophthalmology* 1982;89:1226-1234.
48. Parrish RK. Visual impairment, visual functioning, and quality of life assessments in patients with glaucoma. *Trans Am Ophthalmol Soc* 1996;94:919-1028.
49. Mills RP. Esterman disability rating in severe glaucoma. *Ophthalmology* 1986;93:371-378.
50. Crabb DP, Viswanathan AC, McNaught AI, et al. Simulating binocular visual field status in glaucoma. *Br J Ophthalmol* 1998;82:1236-1241.
51. Advanced Glaucoma Intervention Study. 2. Visual field test scoring and reliability. *Ophthalmology* 1994;101:1445-1455.
52. Fisher LD, Van Belle G. *Biostatistics: A Methodology for the Health Sciences*. New York: John Wiley & Sons; 1993.
53. Torrance GW, Feeny D. Utilities and quality-adjusted life years. *Int J Technol Assess Health Care* 1989;5:559-575.
54. Harris ML, Jacobs NA. Is the Esterman binocular field sensitive enough? In: Mills RP, Wall M, eds. *Perimetry Update 1994/1995*. Amsterdam, Netherlands: Kugler; 1995:403-404.
55. Buckingham JK, Birdsall J, Douglas JG. Comparing 3 versions of the time trade-off: Time for a change? *Med Decis Making* 1996;16:335-347.
56. Turano KA, Rubin GS, Quigley HA. Mobility performance in glaucoma. *Invest Ophthalmol Vis Sci* 1999;40:2803-2809.
57. Klein BE, Klein R, Jensen SC. A short questionnaire on visual function of older adults to supplement ophthalmic examination. *Am J Ophthalmol* 2000;130(3):350-352.