Variability of Automated Visual Fields in Clinically Stable Glaucoma Patients

Elliot D. Werner,* Benno Perrig,t Theodore Krupin,† and Kim I. Bishop†

The total variability of the visual field was measured in 20 patients with open-angle glaucoma who appeared to be clinically stable and well controlled on medical therapy. All patients had at least four visual fields performed on the Octopus 201 perimeter with at least 12 months follow-up since their first visual field. The four most recently performed visual fields were analyzed. Two different methods for calculating total variability were used. One was based on the variance of the threshold determinations and the other was based on the range. The average total variability per subject was 2.8 decibels (db) using the variance-based calculation and 5.1 db using the range-based calculation. Ninety-five percent of the test locations had a variability of less than 6 db by the variance-based calculation method and 13 db by the range-based calculation method. We discuss the possibility of using this type of data to develop criteria for detection of progressive visual field loss in glaucoma. Invest Ophthalmol Vis Sci 30:1083-1089, 1989

It is well recognized that the visual field of an individual often fluctuates on repeated testing. Techniques for measuring the fluctuation of the visual field over time have been developed by several workers. It has now been established that larger amounts of variability are found in the visual fields of patients with glaucoma and glaucoma suspects than in normal subjects.

Clinicians are primarily interested in learning whether or not their patients are getting worse so that appropriate glaucoma therapy may be applied. In the presence of large amounts of variability, the separation of true change in the visual field from fluctuation becomes an important problem. Various statistical techniques have been applied to the visual field in an effort to define progressive visual field loss by detecting statistically significant change. None of these techniques has been clinically validated. It has recently been shown that experienced clinical observers and standard statistical tests do not agree well with each other in detecting progressive visual field loss in glaucoma patients.

The purpose of this study is to determine the total amount of visual field fluctuation that may be expected over a long period of time in a sample of clinically stable glaucoma patients. Using this information we may be able to develop criteria for the amount of change in the visual field that must occur to be interpreted as significant when measured by automated perimetry.

Materials and Methods

The charts of all patients followed by either one of two of the authors (TK or EW) in the glaucoma service of the Scheie Eye Institute were reviewed. All subjects who met the following criteria were included in the study:

1. Diagnosis of primary open-angle glaucoma, pigmentary glaucoma or pseudoexfoliation with glaucoma;
2. At least four available automated visual field examinations with at least 12 months follow-up since the first visual field;
3. Presence of a definitive nerve fiber bundle visual field defect detected on manual (Goldmann) perimetry prior to the patient's first automated visual field examination;
4. Visual acuity of 20/40 or better with no change in visual acuity greater that one Snellen line during the follow-up period;
5. No intraocular pressure greater than 19 mm Hg at
any time since the initiation of the patient's present medical therapy;
6. No change in the appearance of the optic nerve head which could be detected by the treating ophthalmologist (TK or EW) after reviewing serial stereoscopic disc photographs obtained at yearly intervals on each subject;
7. No change in medical therapy for glaucoma during the follow-up period;
8. No laser or surgical treatment for glaucoma during the follow-up period and the 12 preceding months.

All subjects were phakic and had no other known ocular, neurologic or systemic disease likely to affect the visual field or other visual functions. All visual fields were performed on the Octopus 201 perimeter using program 32 with an undilated pupil.

A mean, sample variance and range were calculated for the sensitivity values expressed in decibels (db) for each test location in the visual field of each subject. Only the four most recently available visual field examinations were included in the calculations. Each subject was followed for at least 12 months after their first automated visual field examination. It is possible, however, in some subjects that the four most recent examinations that were used for analysis were obtained over less than 12 months.

The two test locations corresponding to the usual blind spot area were eliminated. There were, thus, a total of 74 test locations but 84 threshold values for each visual field because each of the two values for the doubly determined points were treated as separate test locations for purposes of statistical analysis. All visual field threshold data were entered into an electronic spreadsheet (1-2-3, Lotus Development Corporation, Cambridge, MA) on a personal computer for analysis.

Two methods were used to calculate the total variability per test location. In the first method, total variability per test location was defined as the square root of the sample variance (standard deviation) of the sensitivities for that location. Using this method based on variance, the total fluctuation of the entire visual field for each subject was defined as the square root of the mean variance of all subject's test locations, excluding test locations for which the threshold was always zero db.

Test locations for which the threshold was always zero were excluded because no variability can be measured if the threshold exceeds the maximal luminance available on the perimeter. Inclusion of test locations where the threshold was measurable during some of the examinations and zero at other times may have slightly underestimated the variability at those locations. We felt, however, that if the threshold was zero on some examinations but recordable on others, then the threshold for that location was probably varying around a mean greater than zero, and, therefore, the variability was measurable based on the available threshold determinations.

In the second method, total variability was derived from the range of the sensitivity values for each test location. We call this range variability. Range variability for a single test location was defined as the absolute difference between the highest and lowest sensitivity readings for that test location. With this method, total fluctuation of the entire visual field for a single subject was defined as the mean of the range variabilities of all test locations, excluding test locations where the sensitivity was always zero db.

Each visual field was divided into ten sectors roughly corresponding to the nerve fiber layer anatomy of the retina (Fig. 1). The number of threshold determinations in each sector varied between five and 13. A mean threshold was then calculated for each sector of each visual field for each individual subject.

This value is called the sector threshold.

Each sector threshold was treated as a single data point. The variance and range for each sector of each subject's visual field over time were calculated. As was done for individual test locations, total variability of the entire visual field was calculated in terms of either the square root of the mean variance or the mean range of all sector thresholds.
Results

Twenty subjects were found who satisfied the inclusion criteria. Each subject was followed for at least 12 months after their first automated visual field examination; however, the average time interval encompassed by the four most recent visual fields analyzed on each subject was 22.5 months (range 8-36 months). If both eyes of a subject qualified for the study, one eye was selected at random.

The average mean sensitivity of each subject's visual field for the entire sample was 15.8 db (range 5.4-24.7 db). The study thus included a sample of glaucoma patients ranging from minimal to severe damage (Fig. 2). Of the 1680 test locations in the sample, there were 176 whose sensitivity was always 0 db. These points were excluded from further calculations of variability. Of the 200 visual field sectors in the sample, there were nine sectors of four subjects whose sector threshold was always 0 db. These sectors were also excluded from further calculations of variability.

When using the calculation of variability based on variance, the mean total fluctuation per subject was 2.8 db with a range of 1.4-4.7 db. By using the range-based calculation, the mean total fluctuation per subject was 5.1 db with a range of 2.8-8.8 db.

The variability per test location was also calculated for the entire sample. If the square root of the variance of the four values for each test location is used as the measure of variability, 95% of the test locations had a variability of less than 6 db while 99% had a variability less than 9 db (Fig. 3). If the range of the four values for each test location is used as the measure of variability, 95% of the test locations had a variability less than 13 db while 99% of the test locations had a fluctuation less than 19 db (Fig. 4).

The total variability per sector threshold was also calculated. If the square root of the variance for each sector is used as the measure of variability, 95% of the sectors had a variability less than 3 db while 99% of the sectors had a variability less than 4 db (Fig. 5). If the range of the four values for each sector is used as the measure of variability, 95% of the sectors had a variability less than 7 db and 99% of the sectors had a variability of less than 8 db (Fig. 6).

The size of each subject's pupil was measured prior to each automated visual field test. As might be expected, there was a bimodal distribution of pupil size because some patients were using miotics and some were not. The mean pupil size for the sample was 3.0 mm with a range of 1.0 to 6.0. Nine of the subjects had an average pupil size of less than 2.0 mm while eight of the subjects had a mean pupil size greater than 3.5 mm. The other three subjects were in between.

We also measured the range over which the pupil size varied during the four visual field examinations. Three of the subjects showed no change in pupil size while another 12 showed less than 1.0 mm variation.
in the size of the pupil. The mean of the range of variation in pupil size for the sample was 0.8 mm. The largest change in pupil size in any of the subjects was 2.0 mm.

There was no correlation between the mean size of the pupil and the total fluctuation of the visual field as measured by the range based calculation described above (Pearson correlation coefficient = .223, \( P > 0.10 \), Spearman rank coefficient = \(-.214, P > 0.10\) ). There was also no correlation between the range over which the subjects' pupils varied and the total fluctuation of the field (Pearson correlation coefficient = .161, \( P > 0.10 \), Spearman rank coefficient = .082, \( P > 0.10 \) ).

**Discussion**

The ability of any measuring device to detect a true difference among several measurements depends on the reproducibility of its results. In the case of perimetry, reproducibility is confounded by the fact that the object being measured, the visual field, seems to be constantly changing. Furthermore, the magnitude of this change varies among individuals. This constant changing of the visual field has been termed fluctuation. Our calculation of fluctuation includes both long- and short-term components. We therefore use the term total fluctuation to describe the variability measured in this paper. A difference between examinations done at different times must be greater than the expected physiologic fluctuation to be recognized as a genuine alteration.\(^5,\( a, b \) Thus, any technique that is designed to detect either improvement or progressive deterioration of the visual field in glaucoma must take into account the magnitude of the expected fluctuation in glaucoma patients, not in normal subjects.

A detailed discussion of the sources of variability in visual field testing is beyond the scope of this paper. There are, however, two possible sources of variabil-
ity that we would like to address specifically. The first is pupil size. Visual fields were performed in subjects with an undilated pupil. The pupil size was recorded, but no attempt was made to control the pupil size. It is well recognized that a change in pupil size can have profound effects on the overall sensitivity of the visual field. In fact, the pupil size in our subjects did not vary much between examinations since the subjects' medical therapy was not altered during this study. Unless pupils are routinely dilated for perimetry, random changes in the size of a patient's pupil may be a source of variability. In our sample, however, neither the actual size of the pupil nor changes in pupil size between examinations seemed to correlate with the total fluctuation measured for the subjects.

Another possible source of variability is the presence of a learning effect in automated perimetry. These subjects all had experience with manual perimetry prior to their first automated visual field. We have previously shown that in such patients the effect of patient experience is very small and can only be demonstrated on the short-term fluctuation. Learning, therefore, probably did not contribute much to the variability measured in this study.

With one exception, previous studies of long-term fluctuation of glaucoma patients have generally evaluated multiple visual fields performed over a relatively short period of time. Our study reproduces the usual clinical situation in which a patient is examined and perimetry is performed at intervals of several months over a period of 1 or more years. The magnitude of total fluctuation in such circumstances should provide useful guidance in developing criteria to evaluate true change in the visual field.

In any sample of patients with chronic glaucoma followed for a long period of time, there will be some who are getting progressively worse. In order to evaluate true long-term fluctuation, these patients would have to be excluded. The patients included in this study were all clinically stable so far as we could determine. There was nothing about them that would have led any reasonable clinician to conclude otherwise. The sample should, therefore, provide the best available approximation of the expected total fluctuation in patients with chronic glaucoma who are receiving medical treatment.

It is possible, of course, that the visual fields of some of these subjects may be changing over time despite the clinical impression that their glaucoma is stable. It has recently been shown, however, that this determination is extremely difficult to make using standard statistical tools such as regression analysis, t-test or analysis of variance.

In general, calculations of fluctuation in automated perimetry have been based upon the variance of repeated determinations of threshold. Some authors, however, have based variability calculations on the range. We used both types of calculations, and, as one would expect, the range-based calculation of fluctuation gave values about twice that of the variance-based calculation.

In clinical practice there are certain advantages to using the range as the measure of variability. The range of a series of values is more readily obvious without calculation than the variance or standard deviation. If only a few visual fields are available on a patient, the range or difference between the first and last fields may be more useful for detecting progressive change in the field than a variance calculated from a small number of values.

Based on our entire sample of 1504 threshold determinations, one can now define the upper limit of change in visual thresholds that might be considered a result of random fluctuation. For example, using the range-based calculation of variability, a change of 13 dB or more in a single test location due to random fluctuation will be observed only 5% of the time. There are, however, 74 available test locations in the test pattern known as Octopus program 32. The probability that any number (r) of these 74 (n) test locations vary by at least 13 dB due only to fluctuation and not true progressive visual field change is given by the following binomial formula:

\[
P = \frac{n!}{(n-r)!r!} \times p^r q^{n-r}
\]

where \( n = \) number of points in the visual field tested (74 for program 32); \( r = \) number of points in which the specified change occurs (1–73), \( p = \) probability.
The sector threshold of 95% of the visual field sectors in our sample varied over a range of less than 7 db, and there are only ten sectors in each field. As was done for individual test locations, we can apply the same binomial formula for the sectors. In this case, however, n = 10 and r ≤ 10. The probability of the sector threshold of any one of the ten sectors (r = 1) varying over a range of 7 db, or more is .264, but the probability of two sectors (r = 2) in the same visual field varying by at least 7 db is 0.046.

Using our data, one can now develop rational criteria for progressive visual field loss in glaucoma based upon probabilities. Table 1 shows the number of visual field sectors that would have to change by different specified amounts to be considered outside the range of random fluctuation at the 0.05 and 0.01 probability levels. For example, if one sector varies by 8 db, then this is expected to be due to physiologic fluctuation less than 5% of the time, or if three sectors vary by 7 db, this is expected to be due to fluctuation less than 1% of the time, etc. The magnitude of these sector-based variabilities are very similar to those reported by Hoskins et al.31 These guidelines should be useful both for clinicians who follow patients with automated perimetry as well as for researchers using the visual field to evaluate risk factors or therapy in glaucoma.

Key words: glaucoma, automated perimetry, visual fields, variability, long-term fluctuation

<table>
<thead>
<tr>
<th>Table 1. Number of visual field sectors whose sector threshold would have to change by various specified amounts to be considered outside the expected range of random variability at the 0.05 and 0.01 probability levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of sectors</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

* Range of variability (db).

point will vary by the specified amount or more (.05 or .01), and q = probability point will vary by less than the specified amount (.95 or .99).

Thus for one test location (r = 1), this probability is .891. In other words, even though the probability of a single test location varying over a range of 13 db or more is 5%, the probability of observing any one of the 74 test locations varying by at least this amount due to random fluctuation alone is nearly 90%. It is not until we reach seven test locations varying simultaneously over a range of 13 db or more (r = 7) that the probability of this being due to random fluctuation becomes less than 5%.

As a sample criterion of progressive change, one could say that if seven test locations anywhere in the visual field change by 13 db or more in the same direction, then one can be 95% certain that such a change is not due to random fluctuation.

We do not, however, recommend using data from individual test locations to develop criteria of change for several reasons. One is that our calculations assume that the fluctuation at each test location is independent of adjacent locations, which is probably not true. Second, we assume that there is a measurable threshold at all 74 test locations which is not always true. Third, our technique does not take into consideration the position of the test location in the visual field. Heijl et al have shown that the variability of test locations in the field is position-dependent in normal subjects.30 While this has not been studied in patients with glaucoma, it is likely that the variability of individual test locations is position-dependent in patients with glaucoma as well.

To eliminate the problems of dealing with single test locations, we divided the visual field into sectors. By treating the mean threshold of each sector as a single data point, we deal with groups of test locations. The location and shape of the sectors (Fig. 1) was chosen to reflect the nerve fiber layer anatomy of the retina and the usual patterns of visual field loss in glaucoma.

References