Optimizing and Validating an Approach for Identifying Glaucomatous Change in Optic Nerve Topography

James C. H. Tan and Roger A. Hitchings

**Purpose.** To determine and validate optimal parameters for analysis in a previously described approach for identifying glaucomatous optic nerve progression by scanning laser tomography.

**Methods.** Thirty-degree sectors of rim area, as defined by an experimental reference plane, were analyzed for change with respect to different statistical limits of variability (80%, 90%, 95%, 98%, 99%, and 99.9%) in the longitudinal image series of 62 eyes from 30 ocular hypertension converters and 32 normal control subjects. A criterion requiring that change is repeatable in two of three consecutive tests (the 2-of-3 criterion) was compared with a single-test strategy not requiring confirmation, and four other plausible criteria. The influence of these various parameters on sensitivity and the false-positive rate was evaluated. The same series were also assessed for change by the known method of computer-generated probability maps.

**Results.** More sectors were identified as progressing in converter eyes than in control eyes at every limit of variability. With stricter limits of variability and a requirement of confirmation, fewer sectors were identified as changing, especially in control eyes. The 2-of-3 criterion had the most favorably balanced sensitivity and false-positive rates: these were, for the 90% limit of variability, 90.0% and 6.2%, respectively, and for the 95% limit, 83.3% and 3.1%, respectively. Confirmed rim loss in converter eyes was most frequent in the disc poles and corresponded with the field hemisphere of conversion in 80%. Probability maps detected significant and repeatable change in 26 (86.7%) of 30 converter eyes and 14 (43.8%) of 32 of control eyes. The 2-of-3 criterion had the most favorably balanced sensitivity and false-positive rates.

**Conclusions.** This study was conducted to optimize and validate an approach for identifying progression. The method distinguished eyes with glaucomatous change from unchanging control eyes. (Invest Ophthalmol Vis Sci. 2004;45:1396–1403) DOI:10.1167/iovs.03-0025
rence plane to define the inner edge of the rim. This reference plane is customized to the morphology of each ONH, and it defines rim area more reproducibly than conventional reference planes. The reasoning behind its design is described in detail elsewhere. 14-15 The position of the reference plane (REFpos) relative to each ONH is kept constant in any image series and is defined as

\[
\text{REFpos} = \text{MHC} + \text{LOW}_{\text{sv}} + R
\]  

(1)

where MHC is the mean height of the contour line, LOW_{sv} is the average of the ONH contour line’s 5% lowest height values calculated from the constituent topographies of a baseline image, \( R \) is the level of the reference plane below LOW, where variability is least, previously determined in longitudinal data as \( R = 100 \mu m \). 14

The outer extent of the rim coincides with the contour line, marking the inner margin of the scleral ring of Elschnig. In this study, the same observer drew the contour line in each subject’s baseline topography image (T). Mean topography images are derived on computer from triplets of single topography images (HRT software ver. 2.01; Heidelberg Engineering). Contour lines are exported to other mean and single topography images in each series. Only images with mean pixel SD < 50 \( \mu m \) are used and grainy images with a honeycombed appearance are excluded.

**Testing Different Limits of Variability.** The limits of variability define the smallest amount of change we can expect to detect above test variability. Measurement variability in each sector of an eye exceeds its limit of variability, and enters the zone of change is taken relative to the baseline profile (see Figure 1). Lower limits of variability for each sector are plotted in any image series. This represents a rim area at different points in time in a common graph. Lower limits of variability for each sector are plotted relative to the baseline profile, with the region beneath the lower limits termed the zone of change. Rim area in a sector that diminishes, ability. We had previously arbitrarily defined \( Y \) by a probability of 0.05. 13 In the present study, we evaluated the effect of a range of variability de ned for each sector of an eye. We had previously arbitrarily defined \( Y \) by a probability of 0.05. 13

\[
\text{VARLIM}_s = \frac{\sum |X_s|}{\sum \text{X}_s^2}/(n-1)
\]  

(2)

where \( a \) is the sector number (corresponding to the order of a sector’s location on the ONH circumference between 0° and 360°), \( \delta \) is the sector rim area difference between pairs of same-visit single topography images, \( s \) is the ith value of \( \delta \), \( X_s \) is the mean of observations of \( \delta \), \( n \) is the number of observations of \( \delta \), and \( Y \) is the value of the \( \delta \)statistic for degrees of freedom for \( \delta \), corresponding to a chosen two-tailed probability. We had previously arbitrarily defined \( Y \) by a probability of 0.05. 13 In the present study, we evaluated the effect of a range of probability levels for \( Y \) due to random error. \( Y \) due to random error. We had previously arbitrarily defined \( Y \) by a probability of 0.05. 13

**Correspondence between Morphologic Change and Serial Perimetry.** The spatial relationship between neuroretinal rim loss and visual field change was then examined by associating confirmed rim loss in an ONH hemisphere (superior or inferior) with the presence of confirmed perimetric change representing conversion in the opposite visual field hemisphere.

**Comparison with Detection of Change by Computerized Probability Maps.** The same image series were also analyzed by probability maps (HRT software ver. 2.01b-MS, 1999 with probability map analysis) as described in detail by Chauhan et al. 4,5 elsewhere. The nature of analysis in this version of software is identical with that in the present HR2 Explorer for Windows software (Reuter M, Heidelberg Engineering, personal communication, May 2003). Briefly, the software condenses 100 HRT images with 256 × 256 pixels to arrays of 64 × 64 superpixels, each a grid of 4 × 4 pixels representing topographical height. Each successive follow-up visit is compared statistically with a common baseline visit, with data from each visit derived from its three single topography images, and spatial associations between superpixels statistically accounted for as previously described. 13,14 Computerized color-coded probability maps show regions in which a significant \( (P < 0.05) \) increase (red) or decrease (green) in topographical height is calculated to have occurred. Additional filtering to remove isolated significant superpixels is also possible (Zinser G, personal communication, June 2003). We analyzed our data with and without this filtering. An eye was judged to have significant change if it had at least one cluster of 20 contiguous red superpixels present in the same location in three consecutive images. 13

An Approach for Detecting Glaucoma Progression 1397

**Testing Different Criteria for Verifying Change.** We had established a criterion for verifying change, requiring that any change must be demonstrated in two of three consecutive tests before being attributed to progression. 13 This was to ensure that identified change was consistent with true disease-induced change and not variability. In the present study, we compared this 2-of-3 criterion with (1) a strategy that does not require confirmation (the “single strategy”) at different limits of variability. Sectors were called progressed sectors for a strategy if they met that strategy’s criterion for change. The number of progressed sectors arising by different limits of variability in each strategy was then compared with other plausible criteria: (2) 2-of-2 consecutive tests, (3) 3-of-3 consecutive tests, (4) two adjacent sectors in a single test, and (5) two adjacent sectors in 2-of-3 consecutive tests. Receiver operating characteristic (ROC) curves were plotted to evaluate how well each criterion distinguished eyes with glaucomatous change from the unchanged eyes of normal control subjects for each limit of variability. Any progressed sectors detected in control eyes were considered false positives.

**Excluding Changes in Image Size.** We screened image series to ensure they were free of magnification changes over time. The fit of exported contour lines to the ONH margin was subjectively examined in each follow-up mean topography image and compared with fit in the baseline image. Fit was considered poor if the ONH’s transverse dimensions (x-y axis) differed from that of the exported contour line. In such cases, distances between landmarks (vessel bifurcations in the peripapillary region) were compared between the follow-up and baseline image using the software’s “interactive measures” function, always after having exported the contour line. Seven repeat measurements were made in each image, and the median value for each image was determined. Percentage of change was the ratio of the median measurements at baseline to those in follow-up images. Measurement change exceeding an arbitrary cutoff of 5% above baseline was tantamount to changed size. Series having these changes were excluded from analysis for progression. The same person performed all checking (ICHCT).

Collation of Data. Measurements of rim area in sequential mean topography images are plotted as rim area profiles. Profiles are plots of rim area by angular location round the ONH (0° to 360°, with 0° temporal, 90° superior, 180° nasal, and 270° inferior) from the same image series. This represents a rim area at different points in time in a common graph. Lower limits of variability for each sector are plotted relative to the baseline profile, with the region beneath the lower limits termed the zone of change. Rim area in a sector that diminishes, ability. We had previously arbitrarily defined \( Y \) by a probability of 0.05. 13

An Approach for Detecting Glaucoma Progression 1397

Correspondence between Morphologic Change and Serial Perimetry. The spatial relationship between neuroretinal rim loss and visual field change was then examined by associating confirmed rim loss in an ONH hemisphere (superior or inferior) with the presence of confirmed perimetric change representing conversion in the opposite visual field hemisphere.

Comparison with Detection of Change by Computerized Probability Maps. The same image series were also analyzed by probability maps (HRT software ver. 2.01b-MS, 1999 with probability map analysis) as described in detail by Chauhan et al. 4,5 elsewhere. The nature of analysis in this version of software is identical with that in the present HR2 Explorer for Windows software (Reuter M, Heidelberg Engineering, personal communication, May 2003). Briefly, the software condenses 100 HRT images with 256 × 256 pixels to arrays of 64 × 64 superpixels, each a grid of 4 × 4 pixels representing topographical height. Each successive follow-up visit is compared statistically with a common baseline visit, with data from each visit derived from its three single topography images, and spatial associations between superpixels statistically accounted for as previously described. 13,14 Computerized color-coded probability maps show regions in which a significant \( (P < 0.05) \) increase (red) or decrease (green) in topographical height is calculated to have occurred. Additional filtering to remove isolated significant superpixels is also possible (Zinser G, personal communication, June 2003). We analyzed our data with and without this filtering. An eye was judged to have significant change if it had at least one cluster of 20 contiguous red superpixels present in the same location in three consecutive images. 13

Converter eyes meeting this criterion were considered true positives, whereas normal control eyes with corresponding findings were considered false positives.
Criteria for Selecting Subjects

Sixty-two longitudinal image series from 32 normal control eyes and 30 age-matched ocular hypertension converters were analyzed. Each group had been expanded from their original 20 subjects presented in a previous study, all of whom were included in the present study. Each subject contributed images from only one eye to control subjects and the eye that had converted in the subjects with converters. Converter subjects and normal control subjects had regularly attended the Ocular Hypertension and Early Glaucoma Research Clinic at Moorfields Eye Hospital and had undergone imaging on at least six separate occasions over a minimum of 5 years. This study adhered to the tenets of the Declaration of Helsinki, having appropriate institutional review board approval and subjects’ informed consent.

Normal control eyes were taken to be unchanging. Normal subjects were volunteers comprising spouses or friends of hospital patients, hospital staff, or members of external nonmedical social organizations. They repeatedly had intraocular pressure (IOP) less than 22 mm Hg, serially normal and reliable Humphrey 24-2 visual fields (Carl Zeiss Meditec) with Advanced Glaucoma Intervention Study (AGIS) 18 visual field indices and global rim area.

Converters were assumed to have gluama hypoplasia in the eye that had converted. They initially had a diagnosis of ocular hypertension with IOP consistently 22 mm Hg or more in one or both eyes without IOP-lowering treatment, open angles on gonioscopy, initially normal Humphrey 24-2 visual fields with an AGIS scores of 0, determined after a learning period of three consecutive tests, refractive errors less than ±6 D, and age of more than 40 years. ONH appearance was not taken into account for entry into the study. Image data from all available test visits was used in testing normal series to identify change.

Converters had, over the course of monitoring, shown development of visual field abnormalities, according to AGIS criteria (score > 0), that was reproducible in the same location on three consecutive tests. A glaucoma expert independently confirmed this. Other possible causes of visual field defects were excluded. Patients meeting these criteria were treated medically to lower IOP. ONH appearance was not part of the criteria for inclusion. Converter and normal control groups had equivalent durations of total follow up (control eyes: 5.7 ± 0.85 years, converter eyes: 6.2 ± 0.52 years) and number of imaging sessions. Table 2 shows study subjects’ demographic information. For converters, there was a significant difference (P < 0.05) between final and initial visual field indices and global rim area.

Table 1 shows the excluded eyes (subject 2). These observations were based on distances measured between vessel landmarks in the peripapillary retina outside the ONH margin, and possible explanations for this include factors that affect image magnification, such as lens changes or altered eye-scanner distance, although the possibility of an enlarged scénal canal cannot be totally ruled out. No normal control eyes were excluded from the study. Table 2 shows study subjects’ demographic information. For converters, there was a significant difference (P < 0.05) between final and initial visual field indices and global rim area.

Figure 1 shows that many more converter sectors exceeded their limits of variability to be counted as progressed sectors compared with normal control eyes. This was true at each limit of variability, and also whether (Fig. 1B) or not (Fig. 1A) confirmation was necessary. Stricter limits of variability resulted in fewer progressed sectors in either strategy. Judged by the number of progressed sectors, converter and control groups were more widely separated at stricter limits and in the confirmation strategy than in the single strategy. With the single strategy, converters had 216 and 52 progressed sectors with the 80% and 99.9% limits of variability, respectively.

### Table 1. Description of Three Eyes in Which Altered Image Dimensions Were Documented

<table>
<thead>
<tr>
<th>Subject</th>
<th>Period between Images (y)</th>
<th>Initial Refraction (D)</th>
<th>Refraction Change Over Time (± D)</th>
<th>IOP: Initial–Later (mm Hg)</th>
<th>Measured Distance Change (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.0</td>
<td>+1.5</td>
<td>−1.75</td>
<td>29–24</td>
<td>+6</td>
</tr>
<tr>
<td>2</td>
<td>3.5</td>
<td>−7.5</td>
<td>+0.75</td>
<td>24–21</td>
<td>−15</td>
</tr>
<tr>
<td>3</td>
<td>4.0</td>
<td>+3.0</td>
<td>−3.5</td>
<td>26–32</td>
<td>+7</td>
</tr>
</tbody>
</table>

Subject 2 had high myopia and did not strictly meet inclusion criteria for this study but is included for completeness.

### Table 2. Subjects’ Demographics

<table>
<thead>
<tr>
<th></th>
<th>Normal Controls (median, Q1, Q3)</th>
<th>Converters (median, Q1, Q3)</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Initial</td>
<td>Final</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>14.5 (12.0, 18.5)</td>
<td>15.0 (13.5, 17.0)</td>
<td>0.76</td>
</tr>
<tr>
<td>IOP (mm Hg)</td>
<td>0.58 (−0.26, 1.00)</td>
<td>0.24 (0.39, 0.390)</td>
<td>0.77</td>
</tr>
<tr>
<td>Visual field MD (dB)</td>
<td>1.06 (0.11, 1.42)</td>
<td>1.07 (0.52, 1.58)</td>
<td>0.56</td>
</tr>
<tr>
<td>Visual field CPSD (dB)</td>
<td>1.46 (1.27, 1.70)</td>
<td>1.48 (1.32, 1.70)</td>
<td>0.10</td>
</tr>
<tr>
<td>Rim area (mm²)</td>
<td>26.0 (24.0, 28.0)</td>
<td>26.0 (29.0, 29.0)</td>
<td>0.465</td>
</tr>
<tr>
<td></td>
<td>−0.58 (0.08, −1.46)</td>
<td>−2.54 (−3.86, −0.77)</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>1.65 (1.01, 2.14)</td>
<td>3.22 (2.30, 4.34)</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>1.28 (1.06, 1.48)</td>
<td>1.24 (0.99, 1.42)</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Normal subjects: n = 32; age 65 (59.0, 71.0); converters: n = 30; age 66.1 (55.0, 69.8). Rim area was analyzed by the experimental reference plane. MD, mean deviation; CPSD, corrected pattern standard deviation.

* Significance testing by the Wilcoxon test.
Control eyes had 65 and 6 progressed sectors with the 80% and 99.9% limits of variability, respectively. The ratio of converters to control eyes for the number of progressed sectors was 5:1 for the 80% limits of variability rising to 9:1 for the 99.9% limits. With the 2-of-3 criterion, converters had 137 and 24 progressed sectors with the 80% and 99.9% limits of variability, respectively, and control eyes had 6 and 1 progressed sector with the 80% and 95% limits of variability, respectively. No progressed sectors were found in normal eyes for limits of variability between 98% and 99.9%. Ratio of converters to control eyes for the number of progressed sectors was 23:1 for the 80% limits of variability but 76:1 for the 95% limit.

Figure 2 shows that the ROC curve for the 2-of-3 criterion lies almost horizontally to the left of the curve for the single strategy. The former criterion had fewer false positives, but its sensitivity was not greatly different from that of the single-test strategy for limits of variability between 80% and 95%. Compared with the other criteria, the sensitivity of the 2-of-3 criterion was higher at every confidence limit tested. The 3-of-3 criterion and two adjacent sectors in the 2-of-3 criterion did not have any false positives whatsoever, but this occurred at the expense of sensitivity. Each achieved best sensitivity with the 80% limit of variability: 76.7% in the former and 60.0% in the latter (Fig. 2). For the 2-of-3 confirmation strategy, the 95% limit of variability had a sensitivity of 83.3% (25/30) and a false-positive rate of 3.1% (1/32), corresponding to a specificity of 96.9%; the 90% limit yielded a sensitivity of 90.0% (27/30) and false-positive rate of 6.2% (2/32), corresponding to specificity of 93.8%.

Figures 5, 4, and 5 are examples of converter eyes evaluated by different limits of variability simultaneously and the 2-of-3 criterion. In Figure 3, two inferior sectors are confirmed to have repeatedly exceeded their 90% but not 95% limits of variability. Suspicions of change in a third adjacent sector is raised by its sequentially depressed profile, although its limit of variability is not exceeded and change cannot be verified statistically. Figure 4 shows a different eye, also with inferior rim loss, but in which change has exceeded the stricter 95% and 99% limits of variability. In Figure 5, superior and inferior rim changes exceed the 95% and 99% limits of variability.

Figure 6 shows that converter rim area loss, as identified by 90% limits of variability and confirmed by the 2-of-3 criterion, was most frequent in the inferior hemisphere of the ONH, especially between 240° to 300°. In the superior and nasal hemispheres of the ONH, rim loss was most common superotemporally between 60° and 90°. Change in the nasal rim was uncommon; 81.5% (22/27) of converter eyes showed development of confirmed field abnormality that was matched by confirmed rim loss in an opposite ONH hemisphere. Half of converter eyes (14/27) had confirmed rim loss in both ONH hemispheres but in 10/14 of these eyes, confirmed perimetric abnormality developed in only one field hemisphere.

In probability map analysis without filtering for isolated significant superpixels, confirmed change was identified in 26 (86.7%) of 30 converters and 14 (43.8%) of 32 normal control eyes. Changed clusters were located in converters in the ONH only in 5 (19.2%) of 26, in the peripapillary only in 5 (19.2%) of 26, and in both in 16 (61.5%) of 26. In control eyes, changed clusters were only in 3 (20.0%) of 15, in the peripapillary only in 3 (26.7%) of 15, and in both in 8 (53.3%) of 15. Of the three converter eyes that did not have repeatable rim loss when analyzed by our approach using 90% limits, two also did not have significant topographical change detected by probability maps. Reanalysis after filtering showed confirmed change in 24 (80.0%) of 30 converters and 12 (37.5%) of 32 control eyes.

![Figure 1](image1.png) **Figure 1.** The number of sectors exceeding different limits of variability to be counted as progressed sectors in converter and control groups for the single-test strategy (A) and 2-of-3 criterion (B).

![Figure 2](image2.png) **Figure 2.** ROC curves for the single-test strategy, 2-of-3 criterion, 2-of-2 criterion, and the two adjacent sectors in a single-test criterion plotted for limits of variability between 80% and 99.9%. ROC curves for the 3-of-3 criterion and the two adjacent sectors in 2-of-3 consecutive tests criterion were vertical along x = 0, indicating no false positives. Only their highest sensitivity is plotted, corresponding to their respective 80% limits of variability.
DISCUSSION

We have previously tested this analytical approach and found it to be sensitive and specific.\(^{13}\) The present study was not to confirm previous results but to further develop the concepts we introduced and determine optimal parameters for analysis. No data from previous subjects have been reapplied to this assessment: each image series is independently evaluated by its own individualized reference plane and against its own estimates of variability. The larger sample provided a reasonable basis for optimizing the method.

Fine-tuning the various analytical parameters of the approach by testing different limits of variability and criteria for confirming change yielded the following. First, stricter limits of variability within each criterion reduced the number of progressed sectors in both converter and normal eyes. Normal eyes were affected proportionately more, however, indicating a relative reduction in false positives. Second, the 2-of-3 criterion had fewer false positives but reasonably preserved sensitivity compared with the single strategy. The 2-of-3 criterion was more sensitive and specific than the other criteria we tested. Only the criteria requiring change in two adjacent sectors in 2-of-3 tests and 3-of-3 consecutive tests had better specificity, but their respective sensitivities of 60% and 77% were not optimal. Third, as with stricter limits of variability, the 2-of-3 criterion also reduced the number of progressed sectors in normal eyes, resulting in false positives being eliminated for confidence limits of 98% and above. Sensitivity for the 98% limit, however, was only 67%. ROC curve analysis of the 2-of-3 criterion indicated that the optimal balance of sensitivity and specificity was achieved at either the 90% or 95% limits of variability. The 90% limit had a sensitivity of 90% and false-positive rate of 6.2%; the 95% limit had a sensitivity of 83.3% and false-positive rate of 3.1%. The 90% limit was more sensitive to change but the 95% limit had marginally fewer false negatives. Separation between converters and control eyes for the 95% limit of variability of the 2-of-3 criterion was high: converters had 76 times as many progressed sectors as normal control eyes.

Our approach is designed to deal with variability in three separate steps: experimental reference plane for intervisit variability,\(^{14,15}\) limits of variability for intravisit variability,\(^ {13}\) and verification criteria for any remaining random variability.\(^{13}\) Images from any number of visits, not just pairs of visits, can be assessed simultaneously. Its limits of variability take into account all within-visit single topography images from all test visits, using all available data in a series. For example, an image series comprising three standard test visits, each with three...
single topography images, yields nine point-estimates for calculating limits; but in a fourth visit, say, six images are acquired, 15-point estimates become available from that visit alone, yielding 27 estimates for calculating limits of variability (see the Methods section).\textsuperscript{13} Hence, larger samples for estimating intravisit variability can quickly be obtained as needed and flexibly, without unduly burdening testing resources. Limits also have an inbuilt mechanism for statistically adjusting for sample size so that they widen with smaller samples to reflect increased uncertainty and vice versa.

We showed how sequential data can be evaluated by graded limits of variability to give extra information on the nature of progression in each eye. The stricter the limit of variability, the greater the magnitude of change needed to exceed it, and the more likely it is that measured change is not due to variability alone. Thus, the exceeding of stricter limits of variability reflects a greater probability of change than when less strict limits are exceeded (illustrated in Figs. 4, 5). Conversely, that glaucomatous change could be too small to exceed test variability can also be judged subjectively based on the quantitative analysis, as illustrated in Figure 3. Thus, the approach’s method of objective analysis combined with the option of qualitative appraisal provides a composite framework that is potentially useful for guiding clinical decisions.

Having an empiric basis for duplicate testing is important for determining the usefulness of a test, but also because duplicate testing affects clinical resource allocation. For example, a 3-of-3 criterion\textsuperscript{5} for confirming change requires 50\% more testing than a 2-of-2 criterion.\textsuperscript{11,12} With analysis by our approach, we found that most tentative change could be confirmed by just two consecutive tests. Allowing a third test in the event of a second test’s failing to confirm tentative change increased sensitivity but did not compromise specificity, as evident in the relative positions of the ROC curves for the 2-of-2 and 2-of-3 criteria in Figure 2.

In probability maps, verification of change requires clusters of at least 20 significant superpixels in the same location on three consecutive tests.\textsuperscript{5} Applying this to our data set, we found reasonable sensitivity but many false positives. Filtering for randomly occurring significant superpixels reduced false positives though also sensitivity. It could be that it is hard to adequately estimate and account for both intravisit and intervisit variability just from three images at baseline and at follow-up. Also, topographical height over the ONH fluctuates relative to regions of the peripapillary surface (used for z-axis zero-referencing) in a way that is complex, unpredictable, and not easily accounted for.\textsuperscript{17} We studied only the red superpixels

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig5.png}
\caption{Top: rim area profiles with 95\% and 99\% limits of variability of a converter. Change at sectors 90\° to 120\° and 240\° to 270\° is repeatable, exceeding both the 95\% and 99\% limits of variability (long arrows). Sector change at 60\° to 90\° (short arrow) repeatedly exceeds the 95\% but not 99\% limits of variability. The bottom series of images highlight these sectors.}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig6.png}
\caption{Frequency distribution of sectors identified as progressing in converter eyes as assessed by the 90\% limits of variability and 2-of-3 criterion. (A) Frequency distribution of progressed sectors in the ONH; (B) diagram of the ONH showing the frequency distribution of progressed sectors. N, nasal; T, temporal.}
\end{figure}
representing decreased topographical height; how clusters of green superpixels should be interpreted is not clear to us.

The tradeoff of introducing criteria to improve test specificity is that the severity of minimally detected change is not the same at all points along an ROC curve, nor for different criteria of duplicate testing. This should be noted when interpreting our ROC curve analysis in Figure 2. Stricter limits of variability within an ROC curve can be expected to detect change that is more severe, possibly when underlying disease is also worse, compared with less strict limits. Likewise, each ROC curve in Figure 2 probably represents a different level of change corresponding to a different stage of disease, especially when compared with the single strategy. Our subjects did not prospectively undergo duplicate testing. Instead, their preexisting longitudinal measurements were examined for repeatable change. The degree to which our results had “bias” depends on the underlying rates of progression and how quickly duplicate testing was completed in each series. Other comparable studies having criteria for repeatable change have also used retrospective assessment, and their findings should be interpreted accordingly. Still, our findings on testing by the 2-of-3 criterion are in line with what is predicted by theoretical modeling—namely, that the 2-of-3 criterion markedly improves specificity but does not appreciably compromise sensitivity.

Our validation data set may not have fully represented the whole spectrum of disease severity. We tested only converters because unequivocally telling that they had progression—converter visual fields simply had to change from normal to abnormal—is much simpler and more unambiguous than judging progression in eyes already having established glaucomatous field defects. Whereas converters might be considered to have early glaucoma, we found that some already had quite extensive field loss. Nevertheless, rim loss could be identified in these eyes, and Figure 5 shows an example. Because variability is accounted for separately in each part of each ONH, and because the experimental reference plane is customized to suit the morphology of each ONH, we do not expect morphologic variations to pose problems to our identifying change.

Most converters (80%) had detectable rim loss that matched the field hemisphere in which conversion occurred. Why the remainder did not have matching rim loss is unclear. The concepts underlying field conversion and our method of detecting rim loss are different: visual sensitivity loss had simply to reach threshold, whereas for significant rim loss to be identified, change had to exceed variability. Their measurements are also scaled differently: The units of rim area are on a linear scale, but the decibel scale of visual sensitivity is logarithmic. Thus, it could be that visual sensitivity thresholds were reached before corresponding neuroretinal rim loss exceeded the statistical limits of variability. Some eyes had statistical rim loss that did not have matching confirmed field loss, and it is possible that morphologic change predated white-on-white field defects in these eyes, as has been reported by several investigators. In eyes with ocular hypertension, Kamal et al. have reported detecting rim changes by scanning laser tomography before visual field conversion. In glaucomatous eyes, Chauhan et al. have reported topographical changes by scanning laser tomography before detecting field progression by Humphrey Statpac-2 analysis. Our finding that change was most frequent in the poles, especially inferiorly, agrees with previous observations of disc photographs in ocular hypertension converters and early glaucoma.

We have sought to empirically determine and validate optimal parameters for analysis in the analytical approach. We found that the limits of variability of 90% or 95% provided reasonable cutoffs for identifying progression when the 2-of-3 criterion was used for verification, and our results compared favorably with the alternative technique of probability maps. Characteristics of identified rim loss broadly corresponded to serial perimetry and also published descriptions based on serial photography of the disc at a similar stage of disease. In our approach, the assessing of change by a range of graded limits of variability can give extra information on the nature of change and is potentially useful clinically. Having investigated these issues, we have now turned our attention to evaluating this approach in glaucoma suspects ocular hypertension as well as eyes with diverse presentations of glaucoma of varying severity.

References


6. Quigley HA, Pease ME. Change in the optic disc and nerve fiber layer estimated with the glaucoma-scope in monkey eyes. J Glaucoma. 1996;5:106–121.


