Technique for Detecting Serial Topographic Changes in the Optic Disc and Peripapillary Retina Using Scanning Laser Tomography

Balwantray C. Chauhan, 1,2 J. Wade Blanchard, 3 David C. Hamilton, 3 and Raymond P. LeBlanc 1

PURPOSE. To describe and evaluate a new statistical technique for detecting topographic changes in the optic disc and peripapillary retina measured with confocal scanning laser tomography.

METHODS. The 256 × 256-pixel array of topographic height values obtained with each image from the Heidelberg Retina Tomograph (Heidelberg Engineering, Heidelberg, Germany) was divided into an array of 64 × 64 superpixels, where each superpixel contained 16 (i.e., 4 × 4) pixels. An analysis of variance technique was developed to analyze each superpixel with three baseline and three follow-up images. The performance of the technique was tested with and without adjustment for spatial correlation of topographic values using computer simulations and with real data from a normal control subject and a patient with progressive glaucomatous disc change.

RESULTS. Computer simulation with fixed population means and variance, and varying spatial correlation showed a monotonically increasing number of superpixels with significant test results (false positives), with 20% false-positives when the spatial correlation was 0.8 (the approximate median value in real patient data). The number of false-positive results was similar (17%) in serial images of a normal subject. When corrected for spatial correlation, the number of false-positives was independent of the level of spatial correlation and remained at the expected value of less than 5% in both simulations and real data. Although the number of significant test results in the patient with progressive glaucoma decreased after correction for spatial correlation, the change was readily apparent. Statistical power to detect mean differences in topographic values ranging from 0.5 to 4.0 SDs in computer simulation showed low power for changes of 1 SD or less, but increased dramatically with larger changes.

CONCLUSIONS. This technique has a high level of sensitivity to detect changes in the optic disc while maintaining a high level of specificity. (Invest Ophthalmol Vis Sci. 2000;41:775–782)

Open-angle glaucoma results in deterioration of the visual field and in structural changes involving the optic disc. Changes at the optic disc can manifest as an enlargement or deepening of the optic cup and formation of localized notches in the neuroretinal rim. 1 Quantitative clinical assessment of the optic disc for the follow-up of glaucoma is frequently performed with cup-to-disc ratio estimates. 2 With stereo or monoscopic photography and planimetry, it is also possible to compute the neuroretinal rim area, whereas stereophotogrammetry 3,4 and stereovideographic techniques 5 allow volumetric estimation of the optic cup.

Confocal scanning laser tomography has been introduced as an adjunct or alternative method for the clinical evaluation of the optic disc. The technique has been described elsewhere, 6,7 but briefly, confocal sections of the optic disc, where the focal plane of the laser and the detector plane are optically conjugate, are obtained. The focal plane of the laser is changed incrementally to obtain the sections. The optical setup ensures that the information contained in a given image section is derived largely from the focal plane of the laser. After the confocal sections are aligned and processed, topographic heights of discrete locations in the scanned area are estimated.

Confocal scanning laser tomography allows reproducible estimates of the topographic measurements at individual measured locations, or pixels. 8–10 The most important advantages of the technique are ease of operation, rapid image acquisition and processing times, and, unlike conventional photography, the ability to obtain images with natural pupils in most patients. 11

The ability of detecting change due to disease depends largely on the test–retest variability of measurements. When the variability of measurements is high, there is little statistical confidence in detecting small changes over time. If, however, the variability is low, small changes can be detected with confidence. Previous studies with confocal scanning laser tomography have shown that local variability measurements de-
pend critically on topographic gradients and whether the measurements are made on blood vessels.\textsuperscript{10,12} Because these anatomic features are unique to each eye, local variability estimates should be made for each eye to gauge whether corresponding local topographic differences between two sets of images separated by time are statistically significant.

The purpose of this study was to describe a statistical technique for detecting changes in the optic disc and the peripapillary retina using a commercially available device, the Heidelberg Retina Tomograph (Heidelberg Engineering, Heidelberg, Germany). The robustness of the technique was tested with computer simulation as well as with real follow-up data from a glaucoma patient with progressive disc change and a healthy control subject.

**MATERIALS AND METHODS**

**Equipment**

The Heidelberg Retina Tomograph uses a low-intensity diode laser (wavelength, 670 nm) and a confocal optical setup to obtain 32 equally spaced confocal sections, each with a transverse resolution of 256 × 256 pixels, in the z-axis perpendicular to the optical axis. The image acquisition time is 1.6 seconds.

The 32 sections in each scan are aligned for horizontal and vertical shifts to compensate for any eye movements during image acquisition. The reflectivity profile for each aligned pixel is determined by plotting reflectivity as a function of distance along the z-axis. For a given pixel, the area under the reflectivity profile is the sum of reflectivity, and the position of maximum reflectivity along the z-axis is assumed to be the topographic height value. After the calculations have been made for all pixels, reflectivity and topographic images of the scanned area are determined. Therefore, the result of each scan is a topographic image representing the topographic height of each pixel from the focal plane of the eye.

Typically, in each session multiple scans are obtained in a subject (usually three\textsuperscript{9}) from which a mean topographic and reflectivity image are computed after horizontal, vertical, and rotational alignments are made with a correlation procedure using the reflectivity images. Alignment for depth and tilt are made using the topographic images.

**Statistical Analysis**

The 256 × 256-pixel array of each topographic image was divided into a 64 × 64-array of superpixels with each superpixel containing 16 (i.e., 4 × 4) topographic values. Pooling over a larger area allows more reliable estimates of test–retest variability using three baseline scans.\textsuperscript{10} In this case, there are 48 measurements (3 × 4 × 4) per superpixel. Because the size of a superpixel in a 10° × 10°-scan is approximately 47 × 47 \( \mu \)m, variability estimates based on the 48 measurements will be influenced by the topography of the imaged structure in the superpixel. Therefore, although the pooling has minimal effect on a superpixel situated in an area with flat topography, in an area with steep contours such as the optic cup edge, the variability estimates are increased. To remove the topographic component in variability estimates, we subtracted the respective topographic measurement in each pixel from its respective mean across the three images to determine the adjusted value. After this process, we computed three corrected images from which estimates of test–retest variability were made. These estimates are expressed as variability maps and have been described in detail elsewhere.\textsuperscript{10}

Assuming that there are \( I \) baseline and follow-up images and that the topographic measurements from a superpixel of 4 × 4 pixels is a vector of length 16 indexed by \( l \), an analysis of variance model for each superpixel in vector form is:

\[
\begin{align*}
\mathbf{b}_{il} &= \mu + T_i + L_l + TL_{il} + I(T)_{il} + e_{il} \\
\mathbf{b}_{il} &= \mu + T_i + L_l + TL_{il} + I(T)_{il} + e_{il}
\end{align*}
\]

where \( b_{il} \) represents the topographic height value at time of examination \( t \), at location in image array \( l \), in image \( i \) (\( t = 1, 2; l = 1, \ldots, 16; i = 1, \ldots, I \)). \( \mu \) represents the overall mean topographic value; \( T_t \) the time effect, which allows for differences between the baseline and follow-up examinations; \( L_l \) the location effect, which allows for differences among pixels; \( TL_{il} \), the time by location effect, which allows the time effect to differ by location; \( I(T)_t \) the image within time effect, which allows for variability in the images at baseline and at follow-up; and \( e_{il} \) the error effect assumed to be independently normally distributed with a mean of zero and variance \( \sigma_e^2 \).

With respect to equation 1, determining the significance of temporal changes in topographic values within each superpixel corresponds to testing for the main effect of time and the interaction of time and location simultaneously leading to the test statistic

\[
F = \frac{MS[\text{NUM}]}{MS[\text{DEN}]} \tag{2}
\]

where the mean squared numerator, or \( MS[\text{NUM}] = \{SS[T] + SS(TL)/16; \text{the mean squared denominator}, \text{involves } \{SS(IT) + SS(I)/16 \}/SS[I] \}

where \( SS[T] \) is the sum of squares associated with time; \( SS(TL) \) is the sum of squares associated with time by location interaction; \( SS(IT) \) is the sum of squares associated with image within time; \( SS[I] \) is the sum of squares associated with the residual or error; and degrees of freedom

\[
\nu = 2(I - 1)16.
\]

If \( T_t \), \( TL_{il} \) and \( I(T)_{il} \) are assumed to be fixed effects, then the statistic in equation 2 has an \( F \) distribution with 16 and \( \nu \) degrees of freedom. (See Appendix A for a detailed description of the test statistic).

Given that topographic values in neighboring pixels are likely to be correlated, accounting has to be made for this spatial dependence. If \( I(T)_{it} \) is a random variable with a mean of zero and variance \( \sigma_{I(T)}^2 \) that are independent of \( e_{it} \), then the model (equation 1) allows for spatial correlation between topographic values within a superpixel. The correlation between \( h_{it} \) and \( h_{i'it} \) (two locations within the same superpixel) is

\[
\rho = \frac{\sigma_{I(T)}^2}{\sigma_{I(T)}^2 + \sigma_e^2}
\]

and is referred to as the intraclass correlation coefficient. In the present context \( \rho \) measures the degree of spatial dependence between topographic values within a superpixel.

The effect of spatial correlation is to decrease the amount of available information. The Satterthwaite correction\textsuperscript{13} appropriately reduces the degrees of freedom of the approximating \( F \) distribution for the test statistic to
Sensitivity estimates were made by randomly generating baseline and follow-up images in computer simulations. Simulated baseline and follow-up images were randomly generated by keeping the population mean and variance of topographic values within a corresponding aligned superpixel identical at baseline and follow-up. Ten thousand simulations were performed for each level of spatial correlation ranging from 0 (no spatial correlation) to 0.99 (almost perfect spatial correlation), and the number of superpixels showing significant change (false positives) at $\alpha = 0.05$ was recorded. We compared the results with and without adjustment for spatial correlation.

Sensitivity estimates were made by randomly generating baseline and follow-up images where the mean change between corresponding aligned superpixels varied from 0.5 to 4 SDs, whereas the variance was kept identical. Ten thousand simulations were performed, for the same levels of spatial correlation used for the specificity estimates. The number of significant differences that were detected at $\alpha = 0.05$ were recorded.

RESULTS

The computer simulation experiments showed that when using identical baseline and follow-up images with no spatial correlation of topographic values in a superpixel, the number of significant test results (false-positive rate) for both unadjusted and Satterthwaite-adjusted results was less than 0.05 (Fig. 1). As the level of spatial correlation increased, however, the false-positive rate for the unadjusted test results increased monotonically with increasing spatial correlation. The false positive rate for the adjusted test was independent of spatial correlation and remained at the expected value of less than 0.05.

Because the test procedure depended critically on correction for spatial correlation, the computer simulation results for sensitivity are presented only for the adjusted test procedure.
tions are inadequate estimates of intraindividual variability, because one image yields only one estimate of the index, and the loss of spatial information not captured by summary measurements. Many of the limitations of summary indices for complex data sets are also present in the serial analysis of computed perimetry results, where change in visual field status has to be determined.

In this study we have described a new method for the serial analysis of topographic images based on an empiric probabilistic approach. Variability estimates for small discrete areas (superpixels), each comprising a square of $4 \times 4$ pixels, are estimated. An analysis of variance is conducted on aligned superpixels based on a calculation of serial pointwise differences. The degrees of freedom value used to calculate the $F$ statistic is adjusted to account for the spatial correlation between topographic values of pixels within a superpixel.

Our study showed that the correlation between the topographic values in neighboring pixels is considerable with a median intraclass correlation coefficient of 0.8, which, if not corrected for, would yield erroneous results in the follow-up of patients and healthy control subjects. Computer simulations using baseline and follow-up images with the same population mean and variance showed that as the spatial correlation increased, the number of false-positive test results also increased, with more than 20% false-positive results when the spatial correlation approached 0.8. Analyzing serial images of a normal subject, we found a similar number of false-positive results. When we applied the Satterthwaite correction to adjust the test statistic, the number of false-positive results was invariant to increases in spatial correlation in computer simulation and reduced the number of false-positive results in the same normal subject to levels expected by chance alone.

**Figure 1.** The number of significant test results (false-positive rate) obtained by comparing sets of three baseline and three follow-up confocal scanning laser tomographic images using computer simulation. The population mean and variance of topographic values of corresponding superpixels at baseline and follow-up were identical. The results are shown as a function of spatial correlation in unadjusted and Satterthwaite-adjusted test procedures, with each data point representing the mean of 10,000 simulation runs.

**Figure 2.** The number of significant test results (power) obtained by comparing three baseline and three follow-up confocal scanning laser tomographic images. The topographic height values in the follow-up images were changed by 0.5 to 4.0 SDs from baseline. The results are shown as a function of spatial correlation in Satterthwaite-adjusted test procedures, with each data point representing the mean of 10,000 simulation runs.
Computer simulation also allowed us to determine the power of our proposed method to detect changes as a function of change in topographic values between serial images and spatial correlation. For changes of 2 SDs or more, there was adequate power to detect changes in superpixels where the spatial correlation is 0.6 or less; however, increasingly larger changes are required in superpixels that have a high degree of spatial correlation. The correction for spatial correlation is likely not to inhibit the ability to detect clinically meaningful glaucomatous disc change.

Because scanning laser tomography can be performed rapidly and without pupil dilation in most patients, it is feasible...
to obtain optic disc images at each clinical visit. The availability
of a series of serial images would enhance the ability of the
clinician to assess optic disc progression. A form of serial
change analysis in which each follow-up image is compared to
a baseline image is clearly possible. Using these data, it is
possible to determine locations where repeatedly significant
change from baseline has occurred and assign probabilities to
these changes, both in terms of size of change (number of
clustered superpixels) and time (temporally overlapping signif-
icanat superpixels).

We have described a statistical technique for the serial
analysis of optic disc topography with scanning laser tomogra-
phy with computer simulation and case examples. The validity
and merit of the technique in a clinical situation, of course,
require data from a larger number of subjects. Additionally, the
guidelines for the clinical use of this and other techniques for
determining change require some convergence between statis-
tical significance and clinical significance. The performance of
this technique in a prospective longitudinal cohort of glau-
coma patients and normal control subjects and comparing

![FIGURE 5. Left optic disc photographs of a glaucoma patient observed longitudinally, showing disc progression; baseline photograph (left) and follow-up photograph (right) 3 years later show changes in the temporal and inferior disc (arrows).](image)

FIGURE 6. Left optic disc and peripapillary retina of the patient shown in Figure 5 imaged with confocal scanning laser tomography. Baseline reflectivity (A) and topographic (B) images were recorded the same time as the baseline photograph in Figure 5. Follow-up reflectivity (C) and topographic (D) images recorded 3 years later and recorded at the same time as the follow-up photograph in Figure 5. The variability map (E) shows highest variability along some vessels (scale represents 90% confidence interval in micrometers). The difference (follow-up minus baseline) map (F) shows change throughout the neuroretinal rim, especially inferiorly and temporally (scale represents change in micrometers). Uncorrected probability map (G) with 48.0% of the superpixels showing significant change, whereas the Satterthwaite-corrected probability map (H) showed 25.0% of the superpixels with significant change (scales represent probabilities).
results with the topographic indices, conventional optic disc photography, and visual field progression are beyond the scope of this article and will be the subject of a future publication.

References


APPENDIX A

Equation 1 gives an expression for $b_{hl}$, the topographic height value at time $t$, location $l$, and image $i$. In this appendix we give an explicit expression for, and justification for, the test statistic given in equation 2.

The test statistic is a ratio of two terms $MS_{\{NUM\}}$ and $MS_{\{DEN\}}$. $MS_{\{NUM\}}$ is a scaled measure of the total squared deviation between the estimated mean heights at the two times, whereas $MS_{\{DEN\}}$ is a scaled measure of the total squared deviation of individual heights around their means. Under the null hypothesis of no mean differences among the images at each time and of no mean difference between the two times—that is, no $T$, $TL$, or $I$ effects—the $F$ ratio is approximately equal to one. Large values of $F$ give evidence against the null hypothesis of no difference and suggest that the differences are real.

Assuming no differences between images at time $t$, the mean heights are $\mu_{tl} = \mu + T_t + I_t + TL_{tI}$. These means are estimated by the average heights at each location $l$

$$\hat{\mu}_{tl} = \sum_{i=1}^{I} b_{1li}$$

and

$$\hat{\mu}_{2l} = \sum_{i=1}^{I} b_{2li}$$
and their difference gives information about the magnitude of the time and time by location interaction effects. Squaring and summing these differences gives the total squared difference

\[ SS\{NUM\} = \sum_{t=1}^{16} (\hat{\mu}_{it} - \hat{\mu}_{i2})^2 \]

If \( T_t, TL_t, I(T) \) are assumed to be fixed effects, then each difference has variance \( 2\sigma^2 / I \), and \( SS\{NUM\} / (2\sigma^2 / I) \) has a \( \chi^2 \) distribution with 16 degrees of freedom. The numerator of the test statistic is

\[ MS\{NUM\} = \frac{SS\{NUM\}}{16} \]

Under the same assumptions, the deviation \( b_{iti} - \hat{\mu}_{iti} \) estimates the error term \( e_{iti} \) and their sum of squares

\[ SS\{DEN\} = \sum_{t=1}^{2} \sum_{i=1}^{16} \sum_{l=1}^{I} (b_{iti} - \hat{\mu}_{iti})^2 \]

has expected value \( \sigma^2 \), where \( \nu = 2(I - 1)16 \), so an estimate of \( \sigma^2 \) is given by

\[ MS\{DEN\} = SS\{DEN\} / \nu \]

The quantity \( SS\{DEN\}/\sigma^2 \) has a \( \chi^2 \) distribution with \( \nu \) degrees of freedom, independent of \( SS\{NUM\} \). It can be shown that

\[ SS\{DEN\} = SS\{I(T)\} + SS\{e\} \]

where \( SS\{I(T)\} \) is the total variation among images within different follow-up times, and \( SS\{e\} \) is the total residual variation.

It follows that, under the assumptions, the test statistic (equation 2) has an \( F \) distribution with 16 and \( \nu \) degrees of freedom. If \( I(T) \) is assumed to be a random effect, however, then the Satterthwaite approximation is used to approximate the distribution of the test statistic as an \( F \) with \( f_N \) and \( f_D \) degrees of freedom, as in equation 3.

**APPENDIX B**

This is an investigation of the effect of spatial dependence on the degrees of freedom as given by the Satterthwaite correction. Both degrees of freedom in equation 3 depend on the data through the mean squares. Their approximate expectations under the null hypothesis of no change are

\[ E[f_N] \approx \frac{16^2}{(1 + 15\rho)^2 + 15(1 - \rho)^2} \]

\[ E[f_D] \approx \frac{\nu^2}{2(I - 1)15(1 - \rho)^2 + 2(I - 1)(1 + 15\rho)^2} \]

where \( \rho \) is the degree of spatial correlation and \( \nu = 2(I - 1)16 \). If there is no spatial dependence (\( \rho = 0 \)), then the expected number of degrees of freedom are \( (16, \nu) \), which agrees with the classic result. As \( \rho \) increases, the number of degrees of freedom in both the numerator and denominator decrease in a monotonic fashion. In the limit, as \( \rho \) approaches 1, \( E[f_N] \) approaches 1 and \( E[f_D] \) approaches \( 2(I - 1) \).