Comparison of Computer-Assisted Versus Manual Optic Nerve Head Pallor Measurement

Michael J. Cox* and Colm O’Brien†

Computerized methods of examining the optic nerve head offer the possibility of providing consistent, reliable, and accurate measurements of the nerve head. The reliability and accuracy of such instruments must be examined to confirm this. This research was undertaken to gain an insight into the accuracy of such a system, which takes its measurements by tracking the luminance boundaries of the pallor area and the optic nerve head after an observer marks the starting points. Measurements made using the computer-tracked boundary were compared with those made using a boundary tracked by an image-processing system of greater sophistication, the human visual system. The comparison was made with images from one eye from each of 30 subjects (10 normal, 10 ocular hypertensive, and 10 glaucomatous). The measurements were made by two observers. Each observer made each set of measurements twice, and the results were analyzed with a five-factor analysis of variance. The measurements made using the computerized method did not differ significantly from those made using the manual boundary tracking for the two observers, and they were highly correlated (r = 0.99 for the area pallor-disc ratio). Significant differences (P < 0.05) were found between the two observers in both the manual and computerized boundary tracking. The computerized method, however, did not find differences between the two sets of measurements made by each observer; the manual method did show such differences. The computerized method appeared to trace the luminance boundaries successfully in the image, and it might reduce errors related to the observer marking the starting points on different occasions. It is not able to remove the errors caused by different observers marking the starting points. For maximum reliability in measurement, the computerized system should be used by the same observer when assessing optic nerve heads. Invest Ophthalmol Vis Sci 33:3169–3173, 1992

In assessing the optic nerve head for changes that occur during the course of primary open-angle glaucoma (POAG), such as the increase in cup size,1 the vertical elongation of the cup,2 and the increase in pallor,3 it is necessary for the measurements to be repeatable and accurate. We compared measurements made with a computer-assisted optic nerve head analysis system4 with those made using a manual method. This automated system uses luminance and pallor measurement compared with the more common computerized systems that measure physical cupping from stereoscopic image pairs. If the repeatability and accuracy of the system is satisfactory, then it should be useful for following changes occurring in the optic nerve head during the course of POAG. Such changes precede changes in the visual field.5-7 Measurements can be made easily using this system that are difficult to make by direct visual observation, such as quantitation of the area of the pallor area and its brightness.

Materials and Methods

The automated system used to measure the optic nerve head has been described previously.4 Briefly, it consists of a Zeiss (Oberkochen, Germany) fundus camera that is attached to a charge-coupled device video camera and then to a computer-controlled video digitizer. This device stores the image of the optic nerve head in the computer, and with the assistance of an observer, measurements of the optic nerve head can be made. The observer marks four points on the boundary of the optic nerve head and the boundary of the pallor area of the optic nerve head. These points comprise the top, bottom, left-most, and right-most points of the boundary in question. These points are used to define a restricted search area for the computer (which then tracks around the relevant boundary). Using the boundaries tracked, the computer measures the vertical and horizontal pallor-disc ratios (ratio of the vertical or horizontal extent of the
pallor area to the vertical or horizontal extent of the nerve head, the area pallor–disc ratio (ratio of the area of the pallor area to the area of the nerve head), the disc eccentricity (ratio of the vertical to horizontal dimensions of the nerve head), the pallor area eccentricity (ratio of the vertical to horizontal dimensions of the pallor area), and the pallor ratio (ratio of the area of the pallor area multiplied by its mean brightness to the area of the neuroretinal rim multiplied by its mean brightness). In addition, it also divides the optic nerve head into four quadrants using St. Andrew’s cross. In each of these quadrants, a measurement is made of the area pallor–disc and pallor ratios.

The good reliability of this instrument when (1) an observer analyzes different images of the same optic nerve head or (2) repeatedly analyzes the same image of an optic nerve head was indicated by a mean coefficient of variation for vertical pallor–disc ratios in glaucomatous eyes of 3%. The reliability of the system when used by ten different observers is worse, however (at 10% for the same measurement). We compared the measurements made by the system using its boundary tracking with those made by a human observer marking the same boundary manually to make the measurements. The reasoning for this approach was imposed by the limitations of directly measuring the structure of the optic nerve head in vivo. To estimate the accuracy of the system, a model eye could have been used for collecting measurements that then could be checked directly. This approach was thought to be too unrealistic compared with the measurement of real eyes. The model’s accuracy would depend largely on the physical and optical limitations of the system and the model eye, and the physical limitations of this system are known. If measurements cannot be made directly and model eye measurements do not provide a realistic assessment, is there another approach that would define a standard for the measurements being made? These measurements rely on the luminance boundary tracked by the computer. If this boundary were tracked accurately, then the measurements would be accurate within the resolution of the system. The human visual system is a highly developed image processor and has sophisticated image segregation and boundary analytic properties. If we compared the simple image processing of the computer in boundary tracking with a manual method of indicating the boundary, then a determination of how well the computer was tracking the boundary could be made.

We tested 30 subjects. Informed consent for participation in this study was obtained from all subjects after the procedures had been explained to them. The resulting total of 30 images for analysis consisted of three groups of ten images each; the groups included normal, ocular hypertensive, and glaucomatous subjects. Normal eyes were defined as eyes with intraocular pressures less than 22 mmHg and with scores for mean defect, corrected loss variance, and short-term fluctuation in the normal range when measured with the G1 program on an Octopus (Interzeag, Schlieren, Switzerland) automated threshold perimeter. Ocular hypertensive eyes were defined as eyes with an intraocular pressure of 22 mmHg or more and with normal scores on the G1 program of the Octopus perimeter. The glaucomatous group was defined as eyes with abnormal scores on the G1 program of the Octopus perimeter for either mean defect or corrected loss variance. All these eyes were undergoing treatment in a hospital glaucoma clinic.

The boundaries of the optic nerve head and the pallor area were marked manually by two experienced observers. This involved setting four cardinal points for each boundary. Then points were selected moving clockwise, starting from the left-most point of the nerve head boundary, with the computer connecting the current point to the previous point by drawing an elliptical arc using the four cardinal points to define the center of the ellipse. The successive points marking the boundary could be placed as close together as required to define the boundary to the observer’s satisfaction. When the boundary formed a closed loop, it was complete. The positions of the marked boundary points were stored on floppy disks. These stored boundary points then were used to supply thefour cardinal points for the pallor area boundary and optic nerve head boundary (which the computer-assisted method of boundary tracking required). Thus the computer-assisted method started with a searcharea that was the same as that used to mark the boundary manually. With both the boundaries found using the computer and manual methods, the computer was used to measure the optic nerve head to compare the two methods of boundary tracking. The two observers were both masked for the type of optic nerve head presented, and both used the same randomized order of presentation. Both observers assessed all images twice. The experiment was designed using two observers and two separate measurements with each method to include instances in which an observer was variable or the computer-defined boundary was followed exactly by chance. This gave a total of three groups of ten images from two observers providing two measurements using two techniques (3 × 10 × 2 × 2 × 2 = 240 sets of parameters).

The measurements of the optic nerve head we considered were horizontal and vertical pallor–disc ratio, area pallor–disc ratio, pallor ratio, quadrant-based area pallor–disc ratios, and pallor ratios. Eccentricity measurements were not analyzed because they de-
pend entirely on the position of the four cardinal points on each boundary; this did not vary between the methods.

An analysis of variance\(^1\) was used to compare the measurements made manually by the observer with those made by the computer. The measurements were compared between the two observers, and the reliability of measurement also was assessed by comparing the first and second measurements. The analysis of variance included five factors: one between factor, the group (with three levels); one pseudo factor, the subject (with ten levels in each group); and three within factors, the observer (two levels), the session (two levels), the measurement (two levels), and the method (two levels). Before analyzing the pallor ratio-related parameters, the logarithm of the values was taken to normalize their distribution.

Results

The correlation coefficients between the computerized and manual methods of measurement for the first session of measurement, pooling across the observers and the three groups, are given in Table 1 for the area pallor-disc ratio measurement of the entire disc and its quadrants. The two methods of measurement give results that were highly correlated.

There were significant differences between the means for the two observers for every parameter. These are shown in Table 2 with the results of Scheffé multiple-comparison testing. This analysis confirmed the influence of different observers on measurements taken using this system for all parameters assessed. Even with only two observers, although the mean values were close, the analysis of variance was sensitive enough to detect significant differences.

Eight of the parameters showed significant differences between the means for the two sessions. These are shown in Table 3 with their relevant Scheffé multiple-comparison test results.

No parameters showed significant differences (at the 5% level) in the analysis of variance between the means of the measurements made using the two methods. Therefore, the measurements being made using the computer-assisted boundary tracking were the same as those indicated by the observers. Interactions occurred for the ln(pallor ratio) measurement (an observer-session interaction) and the superior area pallor-disc ratio measurement (a group-session-method interaction). Evaluating these interactions revealed a significant difference between the measurements with two methods for (1) ln(pallor ratio) measurements for the second observer \((P < 0.001)\) and (2) superior area pallor-disc ratio for the second session of measurements \((P < 0.05)\). Analyzing the interaction for the superior area pallor-disc ratio by the method showed a significant difference between the two sessions for the manual method \((P < 0.05)\) but not for the computerized method.

Discussion

Ideally, regardless of where the observer places the cardinal points, the computer should follow the same (true) boundary of the optic nerve head or pallor area. It can do this to some extent when the same observer marks slightly different starting points on different occasions. It cannot adjust fully for the relatively large differences between different observers because such deviations (Table 2) are more significant than differ-

<table>
<thead>
<tr>
<th>Parameter measured</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area pallor/disc ratio</td>
<td>0.99; (P &lt; 0.01)</td>
</tr>
<tr>
<td>Superior area pallor/disc ratio</td>
<td>0.98; (P &lt; 0.01)</td>
</tr>
<tr>
<td>Inferior area pallor/disc ratio</td>
<td>0.98; (P &lt; 0.01)</td>
</tr>
<tr>
<td>Nasal area pallor/disc ratio</td>
<td>0.95; (P &lt; 0.01)</td>
</tr>
<tr>
<td>Temporal area pallor/disc ratio</td>
<td>0.96; (P &lt; 0.01)</td>
</tr>
</tbody>
</table>

Table 1. Correlation coefficients between the manual and computerized measurement techniques for the area pallor/disc ratio of the entire disc and its quadrants (n = 60)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Correlation coefficient</th>
<th>(DF = 1.27)</th>
<th>(P &lt; 0.0001)</th>
<th>(Scheffe) (&lt; 0.00001)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertical pallor/disc ratio</td>
<td>F = 22.68</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizontal pallor/disc ratio</td>
<td>F = 29.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area pallor/disc ratio</td>
<td>F = 35.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superior area pallor/disc ratio</td>
<td>F = 25.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior area pallor/disc ratio</td>
<td>F = 10.11</td>
<td></td>
<td>(P &lt; 0.01)</td>
<td>(Scheffe) (&lt; 0.001)</td>
</tr>
<tr>
<td>Nasal area pallor/disc ratio</td>
<td>F = 37.89</td>
<td></td>
<td>(P &lt; 0.0001)</td>
<td></td>
</tr>
<tr>
<td>Temporal area pallor/disc ratio</td>
<td>F = 14.87</td>
<td></td>
<td>(P &lt; 0.001)</td>
<td>(Scheffe) (&lt; 0.001)</td>
</tr>
<tr>
<td>Pallor ratio*</td>
<td>F = 14.16</td>
<td></td>
<td>(P &lt; 0.01)</td>
<td>(Scheffe) (&lt; 0.01)</td>
</tr>
<tr>
<td>Superior pallor ratio*</td>
<td>F = 8.04</td>
<td></td>
<td>(P &lt; 0.01)</td>
<td>(Scheffe) (&lt; 0.01)</td>
</tr>
<tr>
<td>Inferior pallor ratio*</td>
<td>F = 16.14</td>
<td></td>
<td>(P &lt; 0.01)</td>
<td>(Scheffe) (&lt; 0.01)</td>
</tr>
<tr>
<td>Nasal pallor ratio*</td>
<td>F = 30.44</td>
<td></td>
<td>(P &lt; 0.00001)</td>
<td>(Scheffe) (&lt; 0.000001)</td>
</tr>
<tr>
<td>Temporal pallor ratio*</td>
<td>F = 14.32</td>
<td></td>
<td>(P &lt; 0.001)</td>
<td>(Scheffe) (&lt; 0.001)</td>
</tr>
</tbody>
</table>

* Natural logarithm of measured value used for tests. F, F ratio. DF, degrees of freedom.
ences caused by varying sessions of measurement by
the same observer (Table 3).

The computer and the manual method of boundary
tracking give largely the same boundaries. The
system marks the boundary that the observer per-
ceives caused by varying sessions of measurement by
the same observer (Table 3).

The computer and the manual method of boundary
tracking give largely the same boundaries. The
system marks the boundary that the observer per-
ceives, but different observers will have different per-
ceptions of the boundary. The ideal solution would be
to have a completely automated system, but this is
more difficult to design. In the future, it seems likely
that better algorithms will be produced and that vari-
ability will be reduced still further. However, there is
always likely to be a problem with some optic nerve
heads in which the boundary of the pallor area does
not have a sharp outline.

During the course of this study, the final measure-
ments all were made by the computer to minimize the
error from a human observer’s poor ability to esti-
mate distances, areas, brightness, and shape.12–14 The
computer can measure with an accuracy limited only
by its resolution of space and luminance after the
boundary has been found. In this way, we used the
manual measurements only as a check on the bound-
ary-detecting process of the computerized system.

The resolution of the equipment used (one pixel
equaled 5 μm on the retina) was well matched to both
the optical resolution limit of the Zeiss fundus camera
(6 μm)13 and that of the adult human eye (3–5 μm).16
The use of high-resolution video standards or differ-
ting equipment with a higher resolution17–19 would
do little to increase the overall resolution, and
hence the maximum theoretic accuracy, of such a
system.

Although much previous work has provided infor-
mation on the reproducibility of computerized optic
nerve head analysis systems for intrimage,20–23 inter-
image,20–27 and interobserver20,23,28 variability, little
has been reported on the accuracy of these measure-
ments. The Humphrey retinal analyzer (Allergan
Humphrey, San Leandro, CA) can produce topo-
graphic optic disc maps for which the height of a point

in the first image may not correlate well with the
height of the same point in a repeated image of the
same optic nerve head.39 The 95% confidence inter-
vals for measurements of depth at a single point were
166 μm for the nerve head and 205 μm for the neuro-
retinal rim in normal eyes. This interval was larger
(232 μm and 261 μm, respectively) for these two fea-
tures in glaucomatous eyes.30 The PAR (PAR Tech-
nology Corp., New Hartford, NY) IS 2000 (later the
Topcon Imagenet; Topcon Instrument Corp., Para-
mus, NJ) had better depth accuracy. However, this
was tested on a model eye and gave 83% of readings
within 100 μm of the true depth and 100%, within 200
μm of the true depth.31

The Rodenstock optic nerve head analyzer (Roden-
stock [UK] Ltd., Gravesend, Kent, UK) has been
compared with manual stereo photographic measure-
ment and planimetry on 49 eyes of 49 “patients”
(diagnosis unspecified).32 This showed that a single
manual estimation of the vertical cup–disc ratio
would have a 95% spread of more than 0.4 over a wide
range of videographic measurements. This result may
be related to the variability in the manual measure-
ment in which an assessment of the optic cup (topo-
graphic) boundary must be made by manual means,
rather than the variability in measurement using the
Rodenstock instrument.

The high correlation between the computerized
and manual boundary-tracking measurements of the
area pallor–disc ratio (up to 0.99, Table 1) contrasts
with the relatively poor correlation between comput-
erized boundary tracking and global pallor analysis
in the Rodenstock optic nerve head analyzer.33 The area
pallor–disc ratio we used and percentage of disc pallor
presented previously35 should be equivalent measure-
ments. In the earlier report, the highest correlation
coefficient between global pallor measurement and
the percentage of disc pallor by boundary tracking
was 0.46. We found that the process of boundary
tracking was an effective indicator of the observations

<table>
<thead>
<tr>
<th>Table 3. Analysis of variance results of session mean differences for optic nerve head measurement, with Scheffé multiple comparison results for session 1 = session 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertebral pallor/disc ratio F = 6.46</td>
</tr>
<tr>
<td>Horizontal pallor/disc ratio F = 4.83</td>
</tr>
<tr>
<td>Area pallor/disc ratio F = 6.16</td>
</tr>
<tr>
<td>Superior area pallor/disc ratio F = 2.28</td>
</tr>
<tr>
<td>Inferior area pallor/disc ratio F = 6.90</td>
</tr>
<tr>
<td>Nasal area pallor/disc ratio F = 1.90</td>
</tr>
<tr>
<td>Temporal area pallor/disc ratio F = 4.45</td>
</tr>
<tr>
<td>Pallor ratio* F = 6.25</td>
</tr>
<tr>
<td>Superior pallor ratio* F = 0.28</td>
</tr>
<tr>
<td>Inferior pallor ratio* F = 6.51</td>
</tr>
<tr>
<td>Nasal pallor ratio* F = 2.91</td>
</tr>
<tr>
<td>Temporal pallor ratio* F = 4.84</td>
</tr>
</tbody>
</table>

* Natural logarithm of measured value used for tests. F, F ratio. DF, degrees of freedom. NS, not significant at 5% level.
of the optic nerve head made by a human observer. Consequently, global pallor analysis, which does not correlate well with the boundary-tracking measurements, cannot be as good an indicator of the observations made by a human operator.

In conclusion, our results showed that the software was able to determine the boundaries of the optic disc and pallor area, producing results similar to those of a human observer. In addition, it can adjust for small changes in the placement of the cardinal points by a single user and give a more repeatable measurement than a manual method of boundary tracking.

**Key words:** primary open-angle glaucoma, ocular hypertension, image analysis, optic nerve head, boundary tracking

**References**