Reproducibility of Digital Image Analysis for Measuring Corneal Haze After Myopic Photorefractive Keratectomy

MIGUEL J. MALDONADO, MD, VICENTE ARNAU, PHD, RAFAEL MARTÍNEZ-COSTA, MD, PHD, AMPARO NAVEA, MD, PHD, FRANCISCO M. MICO, MS, ANGEL L. CISNEROS, MD, AND JOSÉ L. MENEZO, MD, PHD

• PURPOSE: To evaluate the usefulness of digital image analysis for quantifying corneal haze by determining the reproducibility of its measurements at the corneal plane.

• METHODS: In a prospective study, 20 randomly selected eyes that had undergone myopic photorefractive keratectomy were photographed focusing the slit beam on their anterior corneal surface. Each photograph was examined using computer image analysis techniques that detect the edge of the reticular pattern of the image. Quantification of the difference between two areas, treated and adjacent untreated cornea, each containing 3,750 pixels with a resolution of 256 gray levels, was performed. Intra-analyzer variation was determined by evaluating the photographs obtained by two analyzers under standard conditions on four separate visits. Interanalyzer variation was calculated using one measurement and the mean of the four measurements.

• RESULTS: The pooled standard deviation of the measurements for the analyzers was 0.63 and 0.62 gray levels (coefficient of variation, 4.1% and 3.3%). An association between less severe haze measurements and higher reproducibility scores was found ($r = .42; P = .007$). The mean interanalyzer variation was smaller for the average of four measurements, $0.55 \pm 0.37$ gray levels, than for one measurement, $0.94 \pm 0.73$ gray levels ($P = .014$).

• CONCLUSIONS: Good reproducibility for haze measurements by digital image analysis of the differences between the treated and adjacent untreated corneal areas was obtained. When the average of four measurements was used instead of a single measurement, interanalyzer reproducibility increased significantly. This new technique may be used to quantify and analyze corneal haze after myopic photorefractive keratectomy.

The ophthalmic community would benefit from a method that could accurately quantify corneal haze after excimer photorefractive keratectomy. Most large clinical studies of photorefractive keratectomy report haze using a subjective classification based on arbitrary scales. Such scales are limited by their lack of objectivity, accuracy, and reproducibility between centers.

We developed a new method to quantify corneal haze after excimer laser photorefractive keratectomy. The method uses digital image analysis to quantify the difference in opacification between treated and untreated areas of the post-photorefractive keratectomy cornea. In our study, we determined whether...
repeated haze measurements by digital image analysis may be sufficiently reproducible to detect subtle progression of corneal opacification after myopic photorefractive keratectomy.

**PATIENTS AND METHODS**

TWENTY SUBJECTS WHO HAD UNDERGONE PHOTOREFRATIVE KERATECTOMY to correct their myopia were enrolled in a prospective study after giving informed consent. All the patients were randomly recruited from an ongoing trial evaluating photorefractive keratectomy for severe myopia using a MEL-60 excimer laser (Aesculap-Meditec, Heroldsberg, Germany). For each individual, one eye was randomly selected for study. Measurements were taken at a mean (±SD) of 24 (±13.4) weeks (range, 4 weeks to 1 year) after the excimer laser treatment. The patients were, on average, 34.7 years old (range, 24 to 52 years) and the proportion of male subjects was 55% (11 of 20). All the subjects had an attempted correction of −6 to −22 diopters (mean, −10.7 diopters), and neither combined astigmatic corrections nor retreatments were included in this series. Details on the ablative procedure and postoperative care have been described elsewhere. Briefly, the excimer laser delivered a fluency of 250 mJ/cm² at the cornea with a repetition rate of 20 Hz, using a slit-scan mode of 7 × 1 mm to correct severe myopia. The optical zone diameter in all eyes was 5 mm, with a 1-mm wide tapered transition zone. Postoperatively, 0.25% fluorometholone was administered five times a day during the first month and then slowly tapered over the next 5 months.

We used image analysis to quantify the information in the photographs of treated corneas (Figure 1). Our procedure for obtaining the standard anterior segment photographs was to use a 75 SL (slit-lamp) camera (Carl Zeiss, Oberkochen, Germany) and Agfachrome 200-ASA slide films (Agfa-Gevaert AG, Leverkusen, Germany). The conditions for our standardized photographic technique (Appendix 1) were a 1.5-mm-wide, 8-mm-long slit without diffuser lens; a 45-degree angle of illumination from the temporal aspect of the cornea (Figure 2); constant photograph magnification; and high flash intensity without background illumination. Care was taken to keep corneal light reflection out of the central field. Such conditions had been previously found to better highlight the corneal opacification and allow further bidimensional analysis of the treated and untreated areas. A color CCD (charge coupled device) camera (XC-711P, Sony, Tokyo, Japan) captured the photographs of the corneas under a standardized illumination of 1,900 lux. The CCD camera was focused on the geometric center of the color photograph of the treated cornea, and the image acquisition process was visually controlled on a monitor that displayed the picture in gray levels (Figure 1). The output of the CCD camera was fed into a PIP-512/1024 video digitizer (Matrox Electronic Systems Limited, Quebec, Canada).
Figure 2. Standardized photographic technique with 1.5-mm-wide, 8-mm-long vertical slit and an angle of 45 degrees between the observation and the illumination axes from the temporal side of the eye. Patients were asked to look at a fixed target to ensure primary position of gaze, and the slit beam was focused on the anterior corneal surface. The instrument’s zoom magnification changer was adjusted to a magnification factor of 1 (range of the instrument, 0.1 to 1.5). The area of the cornea that is most adequate for quantification is highlighted (C) and corresponds to a portion of the corneal plane that is focused directly, has an iris background free from illumination, and lies over the entrance pupil.

Figure 3. Edge detector masks (3 x 3 pixels) oriented in the row and column direction are applied to every pixel to ascertain which is brighter than its surrounding area. Two images are created, one for the horizontal edge detection mask and another for the vertical edge detection mask. These two filtered images are then added to yield the final gradient image.

To define the regions to be measured, six squares in the treated cornea (Figure 4, middle right; red squares) and another six squares in the adjacent untreated cornea (Figure 4, middle left; green squares) were delineated interactively by means of a cursor on the gradient image displayed on the personal computer monitor. Each of these regions was 25 x 25 pixels wide. In delineating the six regions of the adjacent untreated cornea, the intense reflection of the slit beam on the iris anterior surface that was

bec, Canada), which converted the scanned image into monochrome digital form. The monochrome image was recorded at 512 x 512-pixel resolution and 8-bit gray scale (256 gray levels per pixel).

Automated analysis of the images was performed by implementing original filtering algorithms based on the edge detector masks previously described by Prewitt9,10 (Figure 3). We used the absolute value of the gradient function to enhance the edge of the reticular discontinuities on the image. In this way, the rough areas in the image resulting from intensity differences between adjacent pixels were highlighted. Gradient images (Figure 3 and Figure 4, middle left) were generated by using the two masks, one that responded more strongly to edges oriented horizontally and the other that responded more strongly to edges oriented vertically (Figure 3). Convolving these masks with an image resulted in two filtered images. These two images were then added to yield the complete gradient image (Figure 3). The average time for this whole analysis was approximately 5 seconds.
situated nasally was avoided (Figure 4, top right), and the regions were selected so as not to coincide with the iris crypts, lash, or corneal light reflections. These untreated regions were used as a control for the measurements of the treated zone of the same eye and were designed to minimize the bias originating from changes in brightness attributable to the proximity of the light source, which was situated temporally (Figure 4, top right).

Using commercially available software (Quattro Pro, version 1.0, Borland International, Inc, Scotts Valley, California), we calculated the mean value of the gray level intensities of the 625 pixels contained within each square region delineated on the gradient image. In addition, we produced color-coded three-dimensional surface plots using Matlab with Simulink software (version 4.0, The MathWorks, Inc, Natick, Massachusetts), which graphically represented the intensity and distribution of gray levels of the gradient image (Figure 4, bottom). Absolute haze measurements were calculated by subtracting the average of the values obtained for each of the six untreated regions (U1 through U6) from the average of the six treated regions (T1 through T6).

Each patient was photographed under the above-mentioned standard conditions by two of us (M.J.M. and R.M.-C.) on four separate visits within a week. At each of the four visits, each eye was photographed three times by both examiners. The best photograph obtained by each examiner on each visit was chosen for analysis. The selection criteria for the best photograph included correct alignment of the eye to ensure primary position of gaze, proper focus of the anterior corneal plane, and situation of the corneal light reflection out of the central area to be quantified. The relative coordinates of the pixel position of the regions analyzed (the red and green squares in Figure 4) with respect to the entrance pupil on the first visit were recorded by each examiner for use as a reference on the three remaining visits. The reproducibility associated with repeated measurements (n = 4 per analyzer) of each eye (n = 20) was determined, and then the average reproducibility was computed. To evaluate the variability of the values, we calculated the standard deviation and the coefficient of variation of the four haze measurements for each of the 20 patients. To increase the statistical power of the observations, we combined the variances for each patient to calculate a pooled standard deviation for examiner 1 and another for examiner 2 (intra-analyzer variation for examiners 1 and 2). A small pooled standard deviation and a small coefficient of variation corresponded to a high degree of reproducibility. Pearson correlations and simple linear regression analysis were performed to determine whether a relation existed between haze intensity and reproducibility.

The interanalyzer variation was calculated as the mean absolute difference between readings of examiner 1 and examiner 2, using both the first measurement and the mean of the four measurements. The results were compared with the paired Student's t test. Agreement between analyzers was also assessed by computing the Pearson correlation coefficient. A P value of less than .05 was considered statistically significant.

Figure 4. (Top left) Standardized slit-lamp photograph of a cornea (right eye) 8 months after a −6-diopter correction with photorefractive keratectomy. (Top right) Arrows point to the limits of the slit beam focused on the anterior surface from the patient's temporal (T) side. The tear reflex of the light source must lie temporal to the horizontal limits of the portion of the cornea (C) that is to be analyzed (bar). The intense reflection of the slit beam on the iris anterior surface that is situated nasally (N) is to be avoided. (Middle left) After digitizing the image, the two-mask convolution process shown in Figure 3 was performed, yielding this gradient image, which highlights the roughest areas of the cornea corresponding to discontinuities in the original image. (Middle right) Six regions on the treated cornea over the entrance pupil (red squares) and another six regions on the adjacent untreated cornea (green squares) were selected for quantification. Analysis of the differences between treated and untreated regions in this image provided an overall quantitative measurement of haze of 40.1 gray levels. (Bottom) Color-coded three-dimensional surface plots are graphic representations of haze and correspond to the six regions symmetrically distributed over the entrance pupil (treated cornea, T1 through T6) and the other six on the adjacent untreated cornea (U1 through U6) that were quantified.
RESULTS

The random selection of the eyes analyzed yielded a wide range of degrees of corneal opacification (range, 3.9 to 79.0 gray levels; mean [±SD], 27.7 [±19.2] gray levels).

The pooled standard deviation for haze measurements was 0.63 gray level for analyzer 1 (range, 0.14 to 1.18 gray levels) and 0.62 gray level for analyzer 2 (range, 0.23 to 1.15 gray levels). The corresponding mean coefficient of variation for analyzer 1 was 4.1% (range, 1% to 22%), and for analyzer 2, it was 3.3% (range, 1% to 11%). Figure 5 shows that a difference of less than 2 gray levels from the mean of each of the four haze measurements was always obtained by both analyzers (analyzer 1: mean absolute difference [±SD], 0.53 [±0.45] gray level; range of the absolute values, 0 to 1.85 gray levels; range of the data values, −1.57 to 1.85 gray levels; and analyzer 2: mean absolute difference [±SD], 0.52 [±0.43] gray level; range of the absolute values, 0 to 1.9 gray levels; range of the data values, −1.72 to 1.9 gray levels).

A correlation between haze severity and variability of .42 was found significant (P = .007). Figure 6 illustrates that the less intense the corneal haze, the smaller the standard deviation of the repeated measurements.

Comparable corneal haze results were obtained by the two independent analyzers after doing a single measurement on each cornea. Mean (±SD) absolute interanalyzer differences were 0.94 (±0.73) gray level (range, 0.0 to 2.6 gray levels). However, when the mean of four haze measurements was calculated, the mean (±SD) absolute interanalyzer differences diminished to 0.55 (±0.37) gray level (range, 0.05 to 1.37 gray levels), and this reduction (mean, 0.39 gray level) was statistically significant (95% confidence interval, 0.08 to 0.69; P = .014). Therefore, the interanalyzer reproducibility results for haze measurements were, on average, 41.5% better for the mean of four measurements than for a single measurement.

On the whole, single haze measurements obtained by the two analyzers were not significantly different (paired Student’s t test; P = .61), nor was the average of four haze measurements performed by the two analyzers on each cornea (paired Student’s t test; P = .39). Figures 7 and 8 show, respectively, excellent linear correlations between the single measurements (r = .998; P = .0001) and the mean of four measurements (r = .999; P = .0001) performed by the two analyzers.

DISCUSSION

Digital image analysis was shown to be a useful tool for measuring excimer laser-induced corneal scarring. The results indicated that this method was
highly reproducible, with low inter- and intra-analyzer variability, for quantifying corneal haze. Under standard conditions, the pooled standard deviations were 0.62 to 0.63 gray level, with corresponding variabilities of 4.1% to 3.7%, respectively. The agreement between analyzers was excellent, with mean absolute differences of less than 1 gray level when a single measurement per cornea was compared. This interobserver reproducibility was even 41.5% higher when the mean of four measurements was calculated instead.

An objective haze measurement technique should prove useful in studies comparing different corneal laser techniques and the effects of pharmacologic agents on the corneal healing response. Until now, most authors have used subjective classifications based on arbitrary scales. Grading scales of haze have several important limitations as they are based on the assessments of ordinal level variables. Conversely, more recent objective methods seem to have improved sensitivity and resolution. Most of them analyze multiple points across a standardized slit-shaped section of the cornea or make linear scans to quantify corneal haze. The most important limitation of these methods is that, because of the reticular nature of the scarring and the heterogeneous distribution of the corneal opacities, there is no single point or line on the cornea that reliably represents the overall level of scarring in the treated area. Cherny and associates, using computerized Scheimpflug densitometry, were able to successfully quantify corneal haze, obtaining measurements that correlated reasonably well with the clinical assessment but subject to far less variability. To obtain measurements that reliably reflected the overall level of scarring in the treated area of the cornea, Cherny and associates developed the corneal opacity index, that is, the highest gray scale intensity at which the software demonstrated at least 2% of the analysis area to be opaque. Reports on the rest of the methods do not indicate that this potential limitation has been overcome. Our technique comparatively analyzed at the corneal plane six regions on the treated cornea and another six on the adjacent untreated cornea (Figure 4, middle right). Each of the six regions was 25 × 25 pixels wide and contained 625 pixels, thus covering an area of the cornea measuring 358 × 358 μm². This made for a total extension of 1,240 × 1,240 μm² across each cornea undergoing digital image analysis and quantification, which was far beyond the linear configuration of most alternative methods. Furthermore, with our technique, we have also been able to measure the distribution of the corneal opacities with reference to various anatomic landmarks and provide results on the uniformity of corneal haze over the entrance pupil. Allemann and associates, using sequential high-frequency ultrasound scans of the cornea, noted higher reflectivity in marginal areas and heterogeneous patterns that suggest nonuniform distribution of the opacities. They
did not, however, attempt to assess its magnitude. Likewise, Lohmann and associates noted in some cases that haze was inhomogeneous, showing focal scatter centers without providing an estimate of the quantity of this heterogeneity.

Appendix 2 summarizes the main advantages and disadvantages of published methods for quantifying corneal haze after excimer laser photorefractive keratectomy. In a previous study on digital image analysis for measuring corneal haze, we found that mean differences between scarred and clear areas for haze grades 0.5, 1, 2, 3, and 4 were 16.9, 26.6, 42.6, 60.4, and 76.4 gray levels, respectively ($r_s = 0.96; P = .0001$), which means that this method is sensitive to a wide range of haze severity and correlates well with the clinical assessment. Beyond that, in our system, the spatial resolution of $14.3 \times 14.3\ \mu m$ compares well with the 20- to 40-µm resolution of high ultrasound imaging and approaches the 7.4 × 10.9-µm resolution of on-line Scheimpflug imaging, which is far superior to that of conventional densitometry.

For an instrument to be accepted as capable of producing repeatable readings, it should be clear that the mean difference on repeated readings does not vary significantly from zero and that 95% of the differences are less than 2 SD from the mean difference. Figure 5 shows that our system met these criteria. The literature contains few other studies on the reproducibility of corneal haze measurements that could be used for comparison. Our coefficient of variation values of 4.1% and 3.3% for measurements of corneal opacification were similar to the 3.71% coefficient of variation that Cherny and associates, using computerized Scheimpflug densitometry, reported for corneal opacity index values of greater than 100, and slightly higher than the coefficient of variation of 1.14% for corneal opacity index values of less than 100. They are also comparable to the 3.5% intra-analyzer variability reported by Chang and associates. Although Lohmann and associates reported no systematic diurnal or gender-related variations in the measurements, with optimal signals obtained with a 40-degree separation angle between the CCD camera and the light source, the reproducibility data they provide are not detailed enough to be used for comparisons with ours. Our results using digital image analysis showed a correlation (Pearson correlation, $r = 0.42; P = .007$) between haze intensity and variability (Figure 6). Therefore, one should be particularly strict in applying the standard variables required for the measurements when a severe haze is encountered to obtain a more reproducible result.

Our mean interanalyzer variations were similar to the interanalyzer differences found by Cherny and associates using on-line Scheimpflug densitometry. Figures 7 and 8 show excellent linear correlations calculated for both of our analyzers. Lower interanalyzer variability should be anticipated when the average of several haze measurements is compared with a single measurement. Herein, using an average of four measurements instead of a single measurement improved mean interanalyzer reproducibility by 41.5%. Therefore, the authors recommend that the average of four measurements be calculated to obtain lower interanalyzer variability scores.

There are several sources of error attributable to the patient, the analyzer, and the instrument that can explain the variability observed with our technique. The most common reason for inadequate image quality is eye movement during the photographing process. This is less of a problem when the operator is experienced because the picture is taken quickly; the anterior corneal surface is focused on immediately with the cooperation of a patient who has no discomfort caused by prolonged light exposure, motionless eye fixation, or lack of blinking. At any rate, it

---

### Appendix 1

**Standardized Photographic Technique**

- Patient’s gaze fixed on target to ensure its primary position
- 1.5-mm wide, 8-mm long vertical slit
- 45-degree angle between observation and illumination axes
- Incidence of light from temporal side of patient’s eyes
- Slit beam focused on anterior corneal surface
- Instrument’s zoom magnification changer adjusted to magnification factor of 1
- Slide films transformed into glossy color prints (3.9 × 5.9 inches)
## Appendix 2

### Summary of Methods to Quantify Corneal Haze

<table>
<thead>
<tr>
<th>Published Method</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opacity lensometer&lt;sup&gt;11&lt;/sup&gt;</td>
<td>Based on a commercially available opacity lensometer; measurements are reliable</td>
<td>Designed to measure back-scattered light from the lens; lack of sensitivity in measuring lower grades of haze; currently not in use for haze measurement</td>
</tr>
<tr>
<td>Scatterometer&lt;sup&gt;12,16&lt;/sup&gt;</td>
<td>High sensitivity to measure clinically nondetectable haze; good correlation with visual acuity and contrast sensitivity; tested on humans and animals</td>
<td>Requires modified slit-lamp microscope; exact centering of ablation zone on optic axis is required; measurements are possible only from areas bisected by cornea's equator</td>
</tr>
<tr>
<td>High-frequency ultrasound&lt;sup&gt;13&lt;/sup&gt;</td>
<td>High definition and resolution; ultrasound images correlate with the number and activity of keratocytes; determines epithelial, scar, and total corneal thickness</td>
<td>Requires high-frequency transducers and a motion-controlled system; immersion technique is required; linear scans may not be representative of overall level of haze</td>
</tr>
<tr>
<td>Scheimpflug densitometry&lt;sup&gt;14,15&lt;/sup&gt;</td>
<td>High spatial resolution; good inter- and intra-analyzer reproducibility; corneal opacity index reliably reflects overall level of scarring&lt;sup&gt;14&lt;/sup&gt;</td>
<td>Scheimpflug photography equipment is needed; linear measurements may not be representative of overall level of haze&lt;sup&gt;15&lt;/sup&gt;</td>
</tr>
<tr>
<td>Circular projection of light&lt;sup&gt;17&lt;/sup&gt;</td>
<td>Suitable for nonfixating eyes of experimental animals; long-term stability of instrument performance; good intraobserver reproducibility</td>
<td>Limited to ablation diameters of less than 5 mm; reduced sensitivity due to background of scattered light from lens; may not be representative of overall level of scarring when it is unevenly distributed</td>
</tr>
<tr>
<td>Digital image analysis</td>
<td>Based on standard slit-lamp photography; provides quantitative assessment of not only haze intensity but also haze distribution; good inter- and intra-analyzer reproducibility</td>
<td>Requires image analysis equipment; photographic processing can increase variability; regions to be quantified are defined interactively</td>
</tr>
</tbody>
</table>

is advisable to take three photographs of each cornea and then use the best one for further analysis. The most frequent analyzer error is incorrect choice of the regions on the treated and adjacent untreated cornea for digital image analysis. This type of error decreases with the strict avoidance of undesirable brightness due to the vicinity of corneal reflection of the light source and is also reduced by recording the relative x- and y-coordinates of the regions for successive measurements, which should adhere to the distribution shown in Figure 4, middle right. Moreover, recording the standard conditions used and the relative coordi-
nates of the pixel position of the regions analyzed ensures that the subsequent recordings on the same eye can be performed with the same variables, thus enabling different investigators to do comparable analyses after a few months, without alterations of the setting of the recording. Another potential source of error is the presence of tilt or decentration of the intersection between the visual axis and the cornea on the one hand and the pupil contour on the other. This problem is greatly diminished by reinforcing the patient's primary position of gaze and making the temporal limit of the slit beam on the iris plane lie over the middle of the circumference of the pupil contour (Figure 4, top right). An important instrument-related potential source of error may spring from first photographing the cornea and then using a CCD camera to capture the photographic image. Although we captured the image of the prints under a standard illumination of 1,900 lux, the current image acquisition method may be further optimized by using an integrated video camera with frame-grabbing capacity as a digitizer, thus avoiding photographic processing. This camera-digitizer device would directly record the natural image of the cornea on the slit lamp and provide a digital output signal. However, since this has not been performed in our approach, it is clear that, as an alternative, it is best to use a consistent photographic method. The variables listed in Appendix 1 can be reasonably well controlled by a careful photographic method, and they should be taken into account when assessing the reproducibility of the measurements. The other variables can be completely controlled by the analyzer consistently using the same film type and slit-lamp camera, with the flash intensity always set at the fixed level.

One of the chief advantages of our method over previously published ones is that the final corneal haze value results from subtracting the measurements on the untreated area from those on the treated area for each cornea, thus minimizing any variation in the acquisition or processing of the image. A great advantage of computerized image analysis over human evaluation of photographs is that with the computer there is no variation in the sensitivity of the evaluation process. The software programs used to enhance the edges of the reticular opacification from the digitized photographs are not altered in any way during the measuring.

In conclusion, our pooled standard deviation and coefficient of variation results indicated good intra-analyzer reproducibility of haze measurements at the cornea obtained with digital image analysis. In addition, we obtained adequate interanalyzer reproducibility when a single measurement was calculated and excellent interanalyzer reproducibility when the average of four measurements was taken into account. Therefore, digital image analysis may be used as a noninvasive method for diagnosing and monitoring excimer laser-induced loss of corneal transparency (haze) and could provide a quantitative grading for comparison and follow-up. Clinical application of objective, quantitative, and reproducible methods, such as we have outlined, may have a significant impact on studies of the effectiveness and safety of myopic excimer laser surgery. Through cross-sectional and longitudinal studies, it will be possible to establish databases that accurately relate haze severity and distribution to clinical and refractive variables.

REFERENCES

11. Andrade HA, McDonald MB, Liu JC, et al. Evaluation of an opacity lensometer for determining corneal clarity following...