Objective Assessment of Corneal Staining Using Digital Image Analysis

Youn Sook Chun,1 Woong Bae Yoon,2 Kwang Gi Kim,2 and In Ki Park3

1Department of Ophthalmology, Chung-Ang University College of Medicine, Chung-Ang University Hospital, Seoul, South Korea
2Biomedical Engineering Branch, Division of Convergence Technology, National Cancer Center, Goyang, South Korea
3Department of Ophthalmology, Kyung Hee University College of Medicine, Kyung Hee University Hospital, Seoul, South Korea

Correspondence: In Ki Park, Department of Ophthalmology, Kyung Hee University College of Medicine, Kyung Hee University Hospital, #1 Hoegi-dong, Dongdaemun-gu, Seoul 130-702, Republic of Korea; ipark@khu.ac.kr.
Kwang Gi Kim, Biomedical Engineering Branch, Division of Convergence Technology, National Cancer Center, Goyang, Korea; kimkg123@gmail.com.
Submitted: September 5, 2014
Accepted: October 25, 2014

PURPOSE. To validate a new objective digital image analysis technique to evaluate corneal staining.

METHODS. One hundred photographs of corneal staining from various ocular surface diseases in 100 patients were quantified by a new strategy: a combination of the difference of Gaussians (DoG) edge detection for morphologic properties of corneal erosions and the red-green-blue (RGB) systems and hue-saturation-value (HSV) color model for detection of color. To enhance the image, we adopted a median filter, Otsu thresholding, and contrast-limited adaptive histogram equalization (CLAHE). To validate the diagnostic value of this new strategy, the same photographs were also graded by two independent clinicians using the Oxford scheme and the National Eye Institute/Industry (NEI)-recommended guidelines. The correlation between the average subjective grade and objective image analysis measurement was evaluated using the Pearson’s correlation coefficient.

RESULTS. The new algorithm showed a strong correlation with the clinical grading scale in the Oxford scheme and the NEI-recommended guidelines (R = 0.850 and 0.903, P < 0.001, respectively). The repeatability of the objective measurement was excellent (R = 0.994).

CONCLUSIONS. The new algorithm showed excellent correlation with the traditional subjective clinical grading scales. It may be useful for objective assessment of corneal staining, independent of disease conditions.

Keywords: clinical grading, corneal staining, correlation, objective measurement

Fluorescein staining is an important diagnostic tool for the detection of ocular surface disease. The corneal staining pattern and extent provide important information for characterizing disease, assessing severity, and monitoring clinical response to therapy.1 Currently used clinical grading systems for ocular surface disease have some limitations such as inter- and intraobserver variability, unequal steps, biased reference descriptions for severity, or restriction to specific conditions.2–5 Corneal staining classified by the international Dry Eye WorkShop (DEWS) is even vague and subjective.6 These ambiguities would be an obstacle to multicenter research.

The advantages of digital image analysis are that it can provide accurate and objective results and a permanent record, it can offer more sensitive and reliable assessment than subjective grading,7 it can detect subtle abnormalities in ocular pathology, and it can be applied independent of the competence level of the examiner due to its automatic analysis program. Previous reports related to digital image analysis of anterior segments have concentrated mainly on conjunctival hyperaemia.2,5,7,12 There have been few attempts to quantify the severity of corneal staining by objective measurement using computer software and even those were derivatives from research for conjunctival hyperaemia.2,7,13

The objective analysis of corneal fluorescein staining is very fastidious. Not only corneal erosion but also preocular tear film stained with fluorescein must be considered and differentiated via image analysis even though they are both the same green color. Complicated and multidimensional subjective interpretations including extent, type, and depth must be considered for an objective measurement of corneal staining. The quality of the images is also an important factor for successful objective analysis. Quality is affected by a number of factors, such as the timing of the snapshot in relation to blinking, the concentration of instilled fluorescein, and the time elapsed after fluorescein instillation.

Subjective grading of corneal staining images may be judged by two strategies. One is the chromaticity of bright green color emitted from sodium fluorescein and the other is corneal staining morphology, such as the area occupied by the staining, the number of staining points, or the distribution pattern of the staining. Previous studies, using digital analysis, applied a color extraction algorithm using red-green-blue (RGB) systems,2,7 edge detection,2,7 or operator dependent thresholding techniques13 to detect conjunctival hyperemia and corneal staining collaterally. However, there have been conflicting results in these studies.2,7,15

In the present study, we applied a new strategy: the combination of an RGB system and a hue-saturation-value (HSV) color model14 for detection of color, and difference of Gaussians (DoG) edge detection to establish the morphologic properties of corneal erosions. To enhance image processing, we applied contrast-limited adaptive histogram equalization (CLAHE).15 CLAHE is a widely used technique for contrast enhancement of low-density images. It segments the original
Objective Assessment of Corneal Staining

Materials and Methods

This study compares an objective grading score obtained from digital analysis of corneal staining images with a subjective score determined by two clinical graders using the same images and two different grading schemes.

Image Collection

We selected 100 high-quality photographs of the anterior segment of 100 eyes from the database of the Department of Ophthalmology, Chung-Ang University Hospital. Photographs of each cornea were taken after fluorescein instillation, with the eye in the primary position, and were selected on the basis of satisfactory illumination, focus, and high resolution. There were no selection criteria related to the severity and cause of corneal staining. Photographs were taken using a Haag-Streit BM 900 slit-lamp microscope (Haag Streit AG, Bern, Switzerland) in combination with a Canon EOS 20D digital camera (Canon, Tokyo, Japan). Pictures were interfaced to a personal computer and saved as a JPEG file (3008 × 1960 pixels, 24-bit, RGB).

Generally, a fluorescein-impregnated strip was wet with a single drop of sterile saline, and when the drop saturated the impregnated tip, any excess fluid was shaken off. The lower eyelid was pulled down, and the strip was gently touched onto the lower tarsal conjunctiva only once. The patient was asked to gently blink to distribute the dye across the ocular surface. A photograph of the entire cornea was taken immediately after gentle blinking. We used slit-beam lighting with the maximum width (30 mm) of the white light source, a blue excitation filter, 10× magnification, and a diffusion lens at a 10 to 30° oblique angle (with the light source on the midpoint between the pupil margin and limbus). An automated digital camera system set the aperture, exposure time, and shutter speed to the local region, such that overamplification is eliminated and the characteristics of the low-density histogram regions are clearly defined. To eliminate signal noise, we adopted a median filter and binarization using Otsu thresholding in the background image.16 Except for the RGB system, the other algorithms were strategies used for the first time for the detection of corneal staining.

The purpose of this study was to propose a new objective digital imaging analysis technique and to compare it with traditional subjective grading techniques of corneal staining to determine its utility for the objective grading of corneal staining.

Subjective Grading of Corneal Staining Images

In order to validate the diagnostic accuracy of this new strategy, two independent, experienced ophthalmologists graded the photographs with two widely used standardized grading criteria: the Oxford grading scale17 and the National Eye Institute (NEI)-recommended guidelines.18 Each clinician independently graded the photographs displayed on their own monitor, using their clinic-room illumination without any time limitation. They did not receive any special training regarding the grading techniques or scores used in the study. Previously reported standardized grading criteria for the two grading systems were provided, and the observers evaluated the images according to these criteria. The mean grade of the two graders for photograph based on each grading scale was defined as the standard score of the photograph.

Briefly, the Oxford grading scale divides corneal staining into six groups according to severity from 0 (absent) to 5 (severe). The examiner compares the overall appearance of the patient’s corneal staining with a reference figure, simulating the pattern of staining encountered in dry eye disease. No attempt is made to count the dots or to assess the position or confluence of the dots. The examiner selects the appropriate grade that best represents the state of corneal staining intuitively. The grading system recommended by NEI divides the cornea into five zones (central, superior, temporal, nasal, and inferior) and for each zone, the severity of corneal fluorescein staining is graded on a scale from 0 to 3 based on reference figures. Therefore, the maximum score is 15.

Objective Image Analysis of Corneal Staining

Development of the software program and imaging analysis were performed using an Intel Core i7 3, 20 GHz, 16 GB personal computer. The main programming language for the software was Microsoft Visual C++ .NET (2010, Microsoft, Redmond, WA, USA), and optimized algorithms were obtained from the open source computer vision (Open CV) software library. A flow diagram of the multiple steps of corneal staining digital imaging analysis is shown in Figure 1. After the picture was loaded within the user interface, corneal segmentation was performed using the automatic Daugman’s method15 and the imaging process proceeded according to the algorithm, and finally the ratio of the number of pixels that were identified as corneal erosions to the total number of pixels in the cornea was presented as a percentage of the resulting image.

Corneal Segmentation. For the objective assessment of corneal erosion, we concentrated on the corneal region excluding the eyelid, cilia, and conjunctiva. To detect the circular margin and the area of the cornea, we used Daugman’s method.20 Originally, Daugman’s integro-differential equation was developed for iris segmentation and recognition, yet we applied it to corneal segmentation, where it considers the cornea to be a circle. The position of the pupil can, in general, be detected easily based on pixel resolution value information. Using the pupil as the standard, the continued change of position of the midpoint and length of the radius within the pupil can be determined using Equation 1, as in Figure 2. Then,
Maximal (circumference change rates. Using a circular edge detector, these values are expressed as radius change rates, they become $x$ and the radius at $r$ models the corneal area. The greatest change rate. When the circumference change rate is holding was chosen (Fig. 3D). The algorithm assumes that the erosion edges while removing noise.

To reduce image noise, median filtering was applied (Fig. 3C). The pre-processing step is important to preserve corneal background images by applying the median value per area to it possible to not only to eliminate noise, but also to create areas corresponding to erosions were extracted.

First, median filters, used mainly to eliminate noise effects of eye drops. Contact lens wear, allergic keratoconjunctivitis, and adverse various ocular disease conditions, such as dry eye syndrome, contact lens wear, allergic keratoconjunctivitis, and adverse effects of eye drops.

The average time taken to process a single image was 6.25 seconds for Oxford grading and ≤ 10 seconds for the NEI-recommended guidelines.

**RESULTS**

The eligible 100 anterior segment photographs covered various ocular disease conditions, such as dry eye syndrome, contact lens wear, allergic keratoconjunctivitis, and adverse effects of eye drops.

Contrast-limited adaptive histogram equalization partitions the image into small contextual regions and applies histogram equalization to each one. By limiting the maximum and minimum ranges of equalization operations on individual tiles and by matching with multiple Gaussian transformation functions, overequalization of the final image is eliminated. The enhanced image was thresholded and divided into erosions and background.

Finally, the images were converted into HSV images and erosions were detected. The existing RGB color model is constructed to fit physical equipment rather than human vision. In the HSV color model, which is more cognitively designed and more similar to human vision than the RGB color model, the H channel corresponds to the color value, the S channel corresponds to the chromaticity, and the V channel corresponds to brightness. Images to which CLAHE filters had been applied were converted into the HSV color model (Fig. 3F), and erosions were detected in the H channel through the green area partition (Fig. 3G).

Edge Detection Using DoG. When detecting erosions using the color model, to complement the misdetection of even faint tear areas instead of erosions depending on the image, cornea erosion areas are detected in the G channel of the RGB color model through DoG.

By using Equation 2, which is a Gaussian operation generally used for smoothing, disparate variances were assigned, and the differences between them were used to detect edges. For erosion detection, edge images extracted with DoG were binarized through the Otsu thresholding, open and close operations were performed for the noise elimination process, and the edges of the erosions were highlighted. Figure 3H shows an image obtained by applying DoG edge detection to the original image, which consists of the RGB image.

$$\text{Dog} (x, y) = \frac{e^{-\frac{(x^2+y^2)}{2\sigma_1^2}}}{2\pi\sigma_1^2} - \frac{e^{-\frac{(x^2+y^2)}{2\sigma_2^2}}}{2\pi\sigma_2^2}$$ (2)

Finally, the resulting image (Fig. 3I) was created after all of the processing steps were combined.

**Statistical Analysis**

The objective corneal staining score was defined as the ratio of the number of pixels that were identified as corneal erosions to the total number of pixels in cornea.

The relationship between the subjective grading scales and the objective scores of the new strategy was evaluated using the Pearson’s correlation. To determine the reliability of this program, two objective measurements, which were performed 3 weeks apart, were compared with each other using ordinary least squares linear regression.

Statistical analyses were performed using SPSS software version 19.0 (PASW, version 19.0; SPSS, Inc., Chicago, IL, USA). The alpha level (type I error) was set at 0.05.
Digital image analysis could detect corneal erosions very successfully and finely (Figs. 4A, 4B). Obstacles such as eyelashes (Figs. 4C, 4D) or fluorescein pooling (Figs. 4E, 4F) were well differentiated using this new strategy. Digital image analysis could enhance the corneal erosions; therefore, it could detect the erosion more sensitively than clinicians could feel and detect them (Figs. 4G, 4H). In most images, fluorescein-stained corneal erosions were successfully separated from the less well-defined areas of the overlying fluorescein-stained tear film (Figs. 4B, 4D, 4F, 4H).

However, in few cases, the low quality of images could have been an issue. The new strategy was too sensitive; it misidentified lid reflection or pupil margins as corneal erosions (Figs. 5A, 5B). In addition, it inadequately recognized cases of densely confluent stains. It easily detected linear erosions, but only detected the boundary of the confluent stains, instead of detecting the stains one by one (Figs. 5C, 5D).

The reliability of this program between the first and second measurements, separated by 3 weeks, was excellent ($Y = 0.978X + 0.071$, $R = 0.994$, $R^2 = 0.988$, $P < 0.001$ by linear regression; Fig. 6A).

There was a significant correlation between the Oxford grading scale and the objective measurement ($R = 0.850$, $P < 0.001$; Fig. 6B). There was also a strong correlation between the NEI-recommended system and the objective measurement ($R = 0.903$, $P < 0.001$; Fig. 6C). The NEI-recommended systems showed stronger correlations with the automated techniques than the Oxford grading scale.
DISCUSSION

The newly applied objective assessment algorithm for corneal staining showed an excellent correlation with the traditional subjective clinical grading scales. Among them, the NEI-recommended system showed a stronger correlation with the automated technique than the Oxford grading scale. The reliability of this objective digital image analysis was also very excellent.

Corneal staining is a useful method to evaluate the integrity of ocular surfaces, and its objective assessment outcomes are needed for multicenter studies or large-scale clinical trials. Despite the fact that objective corneal staining assessment is important for recording pathology, monitoring therapeutic

![Representative photographs and their processed images. (A, B) Corneal erosions were detected very successfully and finely. Obstacles such as eyelashes or shadows (C, D), or fluorescein pooling (E, F) were well differentiated. (G, H) Digital image analysis detected the erosion more sensitively than the clinician's detection. In most of images, fluorescein stained corneal erosions were successfully separated from tear film.](https://iovs.arvojournals.org/pdfaccess.ashx?url=/data/journals/iovs/933678/)
response, and comparison of research intervention, it has not been developed for practical use in clinics.

Although few studies regarding digital image analysis for corneal staining have been published, there has been some controversy on this topic. Wolffsohn and Purslow\textsuperscript{7} reported that color extraction and thresholding techniques were negatively correlated ($R = -0.54$ to $-0.72$) with the subjective grading scale of the Cornea and Contact Lens Research Unit (CCLRU)\textsuperscript{22} and only edge detection showed a positive correlation ($R = 0.85$) with CCLRU. In contrast, Pritchard et al.\textsuperscript{13} reported that an observer-dependent thresholding technique showed high correlation ($R = 0.92$) between the objective method and the subjective CCLRU scale. Peterson and Wolffsohn\textsuperscript{2} reported that a combination of edge detection and RGB color extraction showed a better correlation than an individual algorithm ($R = 0.86$ in combination versus 0.36 in each separate algorithm) between the objective method and the subjective CCLRU scale.

All three previous studies evaluated corneal erosions related to contact lens wear and they compared objective measurement with subjective assessment using CCLRU\textsuperscript{22} or Efron\textsuperscript{23} grading scales, which were designed for contact lens-related complications. Corneal staining related to contact lens complications has particular configurations such as 3, 9 o’clock staining or central macular punctate erosion due to hypoxia. Reasonably, objective measurement might be oriented to detect these specific configurations. There is a need to validate the objective assessment of corneal staining for variable ocular disease conditions, such as dry eye syndrome, allergic keratoconjunctivitis, superior limbic keratoconjunctivitis, adverse effects of eye drops, and so on. In this study, we analyzed corneal staining from a variety of conditions not limited to contact lens wear, and showed the effectiveness of objective assessment for various corneal staining features.

Although the correlations between the objective assessment and the two subjective grading systems were excellent,
the NEI-recommended systems showed a stronger correlation than Oxford grading. We inferred the reasons as follows: the Oxford scheme was developed to compare patient eyes with dry eye syndrome with reference images intuitively, without counting dots or assessing the location. Therefore, subjective grading of images implicating other ocular surface disorders might induce further variation from objective assessment. On the contrary, the NEI-recommended system divides the cornea into five zones and grades the severity of corneal staining individually in each zone, with the resulting scores then summed. Therefore, it can encompass the intensity and the area of corneal erosion at the same time and it is in line with the strategy of the newly applied algorithm.

Our results were consistent with those of the report of Peterson and Wolffsohn. Although they used a different image analysis and subjective grading including CCLRU and Efron, they obtained similar results. They determined that subjective judgment showed a 86% to 88% correlation with their image analysis using a combination of edge detection and green color extraction. Also, they emphasized that subjective judgment was dependent on the area covered by the staining rather than the green hue and it was more dependent on the staining extent than the staining depth. This means that more subjective grading, such as depth or type, could show a weaker correlation with objective assessment of corneal staining.

We introduced a new strategy for the evaluation of corneal staining with a combination of occupied area by corneal staining using DoG edge detection and green color extraction using RGB and HSV systems. To highlight corneal erosion, we also adopted a median filter, Otsu thresholding, and CLAHE. Pritchard et al. reported a better correlation ($R = 0.92$) with staining intensities than ours ($R = 0.903$); however, their objective technique used an observer-dependent thresholding technique; therefore, a subjective element was not avoidable. Our results were a little better than those of Peterson and Wolffsohn’s report, which similarly involved edge detection and color extraction of staining. We think the reasons for this are that we focused on the pre-processing of the image to detect the staining more sensitively using CLAHE and kernel size adjustment, and we used a two-color model with RGB and HSV at the same time.

As shown at Figure 5, our program could not perfectly detect confluent erosions. Despite the two types of color models that were applied, the green color of the inner part of the confluent erosions was not completely included. It detected the boundary of the confluent erosions as one large area. We suggest reasons for this as follows. First, the edge detection of each of the confluent erosions did not work because the erosions were so closely spaced that the contrast of the erosion margin was very low and ambiguous. Second, we composed our algorithm with an “AND” operation of color and edge detection. Therefore, the erosion detection was impossible if one of the color or edge detections did not work. In this sense, we inferred that edge detection might be the more important detection mechanism.

After loading the picture within the user interface, the corneal margin was drawn using the Daugman’s method with an automatic process. Special training or manipulation was not needed to operate the software program. Thus, the reliability of this objective assessment was nearly perfect ($R = 0.994$). We believe that this proves the accuracy and objectivity of this program.

Our study has some limitations. The objective strategy could not account for the human eye’s detailed perception of corneal staining morphology characteristics, such as coalescence and dispersion. Although this is a weakness of the digital image analysis, it can be a strength for a universal assessment tool excluding subjective aspects. Another limitation is that our program could not differentiate the inner part of confluent erosions due to the low contrast of the erosions and their ambiguous boundaries. Further studies are needed on this topic to help differentiate confluent erosions. The other limitation is that a yellow barrier filter was not used when photographs were taken. Therefore, subjective grading may have suffered because of the lack of contrast in the primary images. If subjective grading was performed with high-contrast images using a yellow barrier filter, the correlation between subjective and objective assessment might be stronger.

We introduced a new strategy for evaluating corneal staining using both color extraction and the occupied area by staining. The reliability of this program was excellent. The digital method and the subjective approach used here are suitable for use in multicenter studies, such as clinical trials, as well as for analysis in grading centers. However, the subjective approach used here varies from real-time clinical grading via slit-lamp examination, where it is difficult to determine center-to-center variation in grading.

In conclusion, this novel digital imaging analysis technique may be useful in the evaluation of corneal staining objectively, independent of disease conditions and severity.

Acknowledgments
Disclosure: Y.S. Chun, None; W.B. Yoon, None; K.G. Kim, None; I.K. Park, None

References