Reliability of 4 Clinical Grading Systems for Corneal Staining

YEOUN SOOK CHUN AND IN KI PARK

• PURPOSE: To determine the interobserver and intraobserver reliability of 4 clinical grading systems for corneal staining.

• DESIGN: Retrospective, observational study.

• METHODS: One hundred twenty-two photographs of corneal erosions from variable ocular surface diseases were graded by 11 ophthalmologists. Each image was graded with 4 grading systems: the Oxford scheme, the National Eye Institute-recommended system, the area–density combination index, and the Sjögren's International Collaborative Clinical Alliance ocular staining score. Grading was repeated after 1 week to evaluate repeatability. Interobserver and intraobserver reliability were evaluated using intraclass correlation coefficients (ICCs). To determine the degree of agreement based on the severity of corneal staining, the relationship between the variance and the score using each grading system was evaluated with linear regression.

• RESULTS: Interobserver reliability for the 4 grading systems was excellent, with ICCs ranging from 0.981 to 0.991. The intraobserver repeatability of the 4 grading systems also was excellent, with ICCs ranging from 0.939 to 0.998. The National Eye Institute-recommended system showed the best reliability and repeatability. There was no definite correlation between variance and score in the Oxford scheme ($Y = 0.006X + 0.284$; $R^2 = 0.002$) or the Sjögren's International Collaborative Clinical Alliance ocular staining score grading system ($Y = -0.068X + 0.595$; $R^2 = 0.109$). However, there was a significant correlation between variance and score in the National Eye Institute-recommended system ($Y = 0.210X + 0.965$; $R^2 = 0.144$) and in the area–density combination index ($Y = 0.187X + 0.279$; $R^2 = 0.178$); the variance increased with the corneal staining score.

• CONCLUSIONS: The 4 grading systems may be useful for evaluation of corneal staining independent of disease conditions and grading individuals. (Am J Ophthalmol 2014;157:1097–1102. © 2014 by Elsevier Inc. All rights reserved.)

C ORNEAL STAINING IS A VALUABLE CLINICAL TOOL that assesses the integrity of superficial epithelial cell layers of the cornea and conjunctiva using slit-lamp microscopy. Vital staining by sodium fluorescein is the most common technique used to evaluate the cornea. Its staining pattern and extent provide important information to characterize disease, assess its severity, and monitor clinical response to therapy.[1] Corneal staining also is used to measure outcomes in clinical trials.

Various clinical grading methods for corneal staining have been introduced to compare the images of patients' eyes with reference images. Reference images can be black-and-white illustrations with a written description of each grade (Van Bijsterveld system,[2] Oxford scheme,[3] and National Eye Institute [NEI]/industry-recommended guidelines[4]), artistically rendered fine illustrations (Efron scale),[5] or a combination of verbal description and photography (Cornea and Contact Lens Research Unit [CCLRU] scale).[6] Some researchers have proposed a grading system using only a written description of the area and density of corneal erosion.[7,8] Recently, the Sjögren's International Collaborative Clinical Alliance (SICCA) ocular staining score (OSS)[9] was developed to evaluate the severity of keratoconjunctivitis sicca for Sjögren syndrome, using a written description of a modified Oxford scheme.

Originally, the Oxford scheme[3] and the NEI-recommended guidelines[4] were developed to determine the severity of dry eye syndrome. Grading systems such as the Efron[5] or the CCLRU[6] were oriented to contact lens-related disease and focused on conjunctival hyperemia.[10,11] As a result, the Oxford grading panel depicted the typical pattern of corneal erosions shown in dry eye syndrome, and CCLRU[6] reference photographs represented typical corneal and conjunctival images related to contact lens wear. Many researchers, however, adopted the Oxford scale to evaluate drug toxicity[12] or contact lens complications[13] because of its simplicity and ease of use.

Currently used grading systems have some limitations, including subjective judgment, unequal steps, biased reference descriptions of severity, and restriction to specific conditions, such as contact lens wear or dry eye syndrome. Therefore, intraobserver and interobserver variation may occur. However, to our knowledge, the few published reports regarding the reliability of corneal staining grading are limited to the CCLRU scale.[14]

The purpose of this study was to determine the intraobserver and interobserver reliability of 4 clinical grading systems for corneal staining for variable ocular surface

diseases: the Oxford scheme, the NEI-recommended system, the area–density combination index, and the SICCA OSS grading system.

METHODS

THIS WAS A RETROSPECTIVE, OBSERVATIONAL STUDY. Image acquisition, processing, and analysis were performed according to the tenets of the Declaration of Helsinki. This study was approved by the Chung-Ang University Hospital Institutional Review Board.

• IMAGE COLLECTION FOR GRADING SCALE: We used 122 anterior segment photographs from 122 eyes with appropriate illumination, fine-focus, high-resolution, and proper straight-fixation view from the database at the Department of Ophthalmology, Chung-Ang University Hospital. There were no selection criteria related to the severity of corneal staining. Photographs were obtained at ×10 magnification 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 JPG file (2544 × 1696 pixels; RGB, 16 MB).

A fluorescein-impregnated strip was wet with a single drop of sterile saline, and when the drop had saturated the impregnated tip, the excess drop was shaken and removed. The lower eyelid was pulled down, and the strip was gently touched onto the lower tarsal conjunctiva. The patient was asked to blink gently to distribute the dye across the ocular surface. A photograph of the entire cornea was obtained immediately after staining. We used slit-beam lighting with the maximum width (30 mm) of the white light source, a blue excitation filter, and a diffusion lens at a 10- to 30-degree oblique angle (with the light source on the midpoint between the pupil margin and limbus). An automated digital camera system set the aperture, shutter speed, and exposure time based on external lighting conditions.

• FOUR GRADING SCALES FOR CORNEAL STAINING: Eleven independent ophthalmologists with an average of 9.7 ± 4.6 years (range, 4 to 17 years) of clinical experience in various ophthalmic divisions were asked to grade the photographs. To determine interobserver reliability, each clinician independently graded photographs displayed on their own monitor, using their clinic room illumination without any time limitation. The 11 observers then evaluated the same images 1 week later to determine intraobserver reliability. They did not have any special training regarding the grading techniques or scores before beginning the study. Previously reported standardized grading criteria for the 4 systems[3,4,7–9] were provided to the observers individually, and they evaluated the images according to the criteria. The observers never met to discuss this study.

The Oxford grading scale[3] divides corneal staining into 6 groups according to severity: 0 = absent, I = minimal, II = mild, III = moderate, IV = marked, and V = severe. The examiner compares the overall appearance of the patient's corneal staining with a reference figure. No attempt should be made to count the dots or to assess the position or confluence of the dots. The examiner should select the appropriate grade that best represents the state of corneal staining.

The grading system recommended by the NEI Workshop on Clinical Trials in Dry Eyes[4] divides the cornea into 5 zones: central, superior, temporal, nasal, and inferior. For each zone, the amount of corneal fluorescein staining is graded on a scale of 0 to 3: 0 = normal or negative slit-lamp findings; 1 = mild or superficial stippling; 2 = moderate or punctate staining, including superficial abrasion of the cornea; and 3 = severe abrasion or corneal erosion, deep corneal abrasion, or recurrent erosion. The maximum score is 15.

The area–density combination index[7,8] is calculated as the area grade multiplied by the density grade. The severity of total corneal surface damage is graded for both the area (0 to 3) and density (0 to 3) of the lesion: area [A]0 = no punctate staining, A1 = less than one third, A2 = one third or two thirds, and A3 = more than two thirds; density [D]0 = no punctate staining, D1 = sparse density, D2 = moderate density, and D3 = high density with overlapping lesions. Therefore, the staining score can be 0, 1, 2, 4, 6, or 9.

The SICCA OSS grading system[9] is a modification of the Oxford grading scale. Punctate epithelial erosions (PEEs) are counted and scored: 0 = absent, 1 = 1 to 5 PEEs, 2 = 6 to 30 PEEs, and 3 = more than 30 PEEs. An additional point is added if PEEs occur in the central 4-mm diameter of the cornea, if any filaments are seen on the cornea, or if any patches of confluent staining including linear stains are found anywhere on the cornea. The maximum possible score is 6.

• STATISTICAL ANALYSIS: Interobserver and intraobserver reliability were evaluated using intraclass correlation coefficients (ICCs). ICCs are large when there is little variation within the observers' measurements.[15] An ICCs of 0.95 means that 95% of the variance in the outcome is from the photographs themselves, not from the grader. To determine the degree of agreement according to the severity of corneal erosion, the relationship between the variance and the score in each grading system was evaluated using linear regression. Statistical analyses were performed using SPSS software version 19.0 (PASW version 19.0; SPSS, Inc, Chicago, Illinois, USA). The α level (type I error) was set at 0.05.

RESULTS

THE ELIGIBLE 122 ANTERIOR SEGMENT PHOTOGRAPHS covered variable ocular disease conditions, such as dry