Published in final edited form as:
Cornea. 2013 November ; 32(11): 1475–1482. doi:10.1097/ICO.0b013e3182a02e0e.

Assessment of Signs of Anterior Blepharitis Using Standardized Color Photographs

Vatinee Y. Bunya[1], **David H. Brainard**[2], **Ebenezer Daniel**[1], **Mina Massaro-Giordano**[1], **William Nyberg**[1], **Eliza Windsor**[1], **Denise J. Pearson**[1], **Jiayan Huang**[1], **Maureen G. Maguire**[1], and **Richard A. Stone**[1]

[1]Department of Ophthalmology, Perelman School of Medicine, University of Pennsylvania, Philadelphia PA

[2]Department of Psychology, School of Arts and Sciences, University of Pennsylvania, Philadelphia PA

Abstract

Purpose—To describe a standardized technique for acquiring and viewing photographic images of eyelids, assess the reproducibility and validity of a grading protocol for signs of anterior blepharitis, and explore whether the signs depend on the eyelid or area of eyelid assessed.

Methods—Subjects with anterior blepharitis ranging from none to severe were examined by ophthalmologists at clinical sites. Digital images of the eyelids of subjects were acquired using a protocol that allowed calibration of color and luminance. Three ophthalmologists at a centralized reading center applied a novel protocol for grading features of anterior blepharitis from the digital images viewed on color-calibrated monitors. The agreement among graders was assessed with percent agreement and weighted kappa statistics (K_w), and the correlation of photographic and clinical gradings was assessed using Spearman correlation coefficients.

Results—Agreement among graders was excellent (K_w >0.80) on the number of eyelid margin vessels and was substantial (K_w between 0.61 and 0.80) for erythema, collarettes, number of engorged vessels, and number of lashes. Grading of photographic images and the clinical assessments of erythema and lid debris were moderately correlated (r =0.27 to 0.45). The grades for different features depended on whether the upper or lower eyelid, eyelid skin or lid margin, and central or lateral lid were assessed.

Conclusions—Application of a protocol to obtain and display calibrated digital images of eyelids supports standardized assessment of anterior blepharitis in clinical care and research studies.

Keywords

standardized photography; anterior blepharitis; grading; reading center

Introduction

Blepharitis, an inflammatory condition of the eyelids, occurs in a high proportion (>35%) of patients examined by ophthalmologists and optometrists.[1] Blepharitis can be classified by its location on the eyelids, i.e., anterior or posterior, although patients often exhibit some

Corresponding Author: Vatinee Y. Bunya, MD, Scheie Eye Institute, 51 N. 39th Street, Myrin Circle, Philadelphia, PA 19104, 215-662-9791, Vatinee.Bunya@uphs.upenn.edu.

Financial disclosure: The authors retain intellectual rights to the photography protocol for the possibility of future licensing.

features of both types. Anterior blepharitis is inflammation of the lid margin anterior to the gray line and is centered around the lash follicles.[2] Visible eyelid signs in anterior blepharitis include erythema of the skin and lid margin, development of abnormal blood vessels or telangiectasias, and the build-up of eyelash debris.

Clinical research on blepharitis has been hampered by varied terminologies and by uncertainty about which signs might be most diagnostic. Generally, the signs of blepharitis are assessed subjectively using reference scales, but such gradings can be adversely affected by variable disease/scale definitions, unspecified or uncontrolled examination conditions, and by inter-observer variability. In ophthalmology, the use of photographic images and well-defined assessment criteria have been adopted for many diseases such as diabetic retinopathy, macular degeneration, glaucoma, etc. [3-8] Photographic images allow for more objective and reproducible assessment than is possible in a clinical setting. Photos also permit evaluations by more than one grader and allow for the use of image analysis to study regions of interest. For blepharitis, developing such objective approaches could facilitate characterizing affected patients and determining disease status over time, key requirements for clinical trials.

To our knowledge, there are no validated blepharitis grading systems based on digital photographic assessment. One factor that may be contributing to the absence of photography-based assessment is that color (i.e., eyelid erythema) is crucial in identifying and quantifying blepharitis. The capture and display of color images with external photography is challenging. The intensity and composition of both ambient and photographic flash lighting influence the color characteristics of the resulting image. In addition, the color in images of the same target also may vary because of differences in the technical features of image capture among cameras, characteristics of color processing, and how images are displayed.

Here, we describe a technique to obtain, normalize, and view photographic images of the eyelids. We also developed a protocol to grade anterior blepharitis that includes both conventional signs, as well as signs not commonly assessed in conventional approaches to grading the disease. We assess the grading protocol as applied to digital images and explore the utility of less commonly evaluated signs of the disease.

Materials and Methods

Human Subjects and Clinical Examination

Subjects with or without a clinical diagnosis of anterior blepharitis and at least 18 years of age were recruited by ophthalmologists from five clinical centers located in Cleveland, Ohio; Torrance, CA; Artesia, CA; New Albany IN; and Philadelphia, PA. All subjects provided informed consent. The research protocol and informed consent form were approved by each center's local Institutional Review Board. The research adhered to the tenets of the Declaration of Helsinki.

Subjects were examined by ophthalmologists and graded on two features for each eye: erythema of the lower eyelid margin and the presence of lid debris on the upper and lower eyelids. Each feature was graded on a scale from 0 (normal) to 4 (very severe). A photographer trained in the protocol for image acquisition took photographs of the lids of each eye.

Image Acquisition

A standardized external photography protocol was developed to obtain images of the upper and lower eyelids. Photographs were taken using the EOS Rebel T2i camera (Canon U.S.A.,

Inc.) which has a built-in pop-up flash, and with a 100 mm macro lens equipped with image stabilization (Canon EF 100mm f/2.8 L IS USM). Photographers used an adjustable monopod for positioning the camera.

For each subject, a series of images was acquired in a single session. As part of the session, a color calibration image was obtained, consisting of the mini (2.25 × 3.25 in) Gretag Macbeth Color Checker Chart together with a white index card. These were placed at the approximate location of the subject for photography (Figure 1A) – that is, with the camera approximately 1 meter away from the chart and card. The Color Checker Chart contains 24 precisely-produced and colorimetrically stable colored squares and provides a reference color characterization. The index card was subsequently placed near the eye of each subject and included in each clinical photograph, thus allowing image-by-image white balancing that was referenced to the Color Checker Chart. This approach enabled conversion of the eyelid images to a standard color representation (e.g., CIE XYZ, CIE Lab, CIELAB hsl, sRGB) for viewing on a calibrated display, for future analysis, and to minimize variation in image color across centers, patients and over time.

Eyelid images were taken at 1:1.5 magnification. Two images of the upper eyelid were taken, one focused on the eyelid skin surrounding the base of the eyelashes with the subject looking straight ahead and the other focused on the lid margin with the subject looking up (Figure 2). Next, two images of the lower eyelid were taken, one focused on the eyelid skin surrounding the base of the lower lashes with the subject looking straight ahead, and the other focused on the everted lower lid margin with the subject looking straight ahead (taken while having the subject gently evert his or her own lower eyelid). All images contained a region showing a portion of the index card.

Image Processing

Image data were saved in Canon's CR2 raw file format. The public domain program dcraw (http://www.cybercom.net/~dcoffin/dcraw/, versions 8.99-9.05, command line options '-6 -W -o 5 -g 1 1') was used to extract the image data from the raw files and into a nominal CIE XYZ tristimulus representation.[9] This representation was then processed further by custom software written in MATLAB (Mathworks, Inc.). The images of the Macbeth Color Checker and the white index card were used to standardize all photographs for both luminance and color.

The image of the Color Checker Chart was analyzed to determine an affine linear transformation between camera data and CIE XYZ tristimulus coordinates. We used measurements of the surface reflectance function of the individual Color Checker Chart squares (obtained at http://www.babelcolor.com) and computed the target CIE XYZ tristimulus coordinates of the squares under a 5000 °K CIE daylight. We then extracted the nominal XYZ coordinates of each square from the calibration image and found the affine linear transformation that brought the nominal coordinates into agreement with the target coordinates, based on minimizing the summed CIELAB δE error between transformed and target coordinates. This transformation was applied to the eyelid images and produced calibrated XYZ images that were white balanced to 5000 °K illumination (Figure 1B).

One additional step was implemented to correct for slight image-to-image variation in illumination. A diagonal linear transformation was applied to the calibrated XYZ values for each image. This transformation brought the calibrated XYZ values for the white card obtained in each eyelid image into register with the calibrated XYZ values for the white card in the color calibration image (Figure 3).

Image Display

Images were displayed for grading on calibrated monitors (NEC 24" PA241W-BK-SV) and viewed in Photoshop (Adobe Systems, Inc.). To convert the XYZ images to RGB values for display, we first used the calibration head and software supplied with the monitors to place each of them into the same well-defined target state (white point of D5000, maximum luminance of 220 cd/m2, contrast of 500:1, and gamma of 2.2). We then used a spectrophotometer (PR-650, PhotoResearch, Inc.) to measure a set 30 spatially uniform test images displayed on one of the monitors using Photoshop. Each test image drove only one of the red, green, and blue monitor channels. Within channels, the image values ranged from 0 (minimum) to 255 (maximum). These monitor calibration data obtained with the PR-650 were used in conjunction with standard methods to convert between XYZ and RGB values in a manner that accounted for the monitor's channel spectral power distributions and gamma functions.[10] An overall scale factor, common to all images, was applied before conversion to RGB so as to bring the eyelid images into the monitor's luminance range.

Central Grading of Images

Processed photographs were graded independently by three ophthalmologists (MMG, ED, VB) who were masked to the grades assigned by the ophthalmologist who examined the patients in the clinical centers. All images were displayed on the color-calibrated monitors described above using Photoshop CS5 (Adobe corp.) at 50% magnification.

Region of Interest for Grading

For each eyelid image, the central two-thirds of the eyelid, defined by the horizontal borders of the cornea on each image, was considered. Custom programs were written for Photoshop to delineate the center two-thirds of each lid and divide it into five equal "subareas," using vertical lines to delineate each subarea (Figure 4). Subareas were labeled 1 to 5 starting with the most nasal subarea in each eye.

Features of Anterior Blepharitis Included in the Grading Scale

Three main eyelid features of anterior blepharitis were assessed in the photographs: erythema, vasculature, and debris. In addition, we recorded several parameters not typically assessed: the numbers of visible engorged blood vessels, eyelashes, short eyelashes, clumps, and microabscesses.

Erythema of the lid skin and lid margin was assessed for each lid relative to an unaffected area of facial skin near the eyelid. Lid skin was defined as the skin at the base of the eyelashes within one eyelid margin width from the anterior lid margin. Only the anterior half of the eyelid margin was graded. The area of unaffected skin (control skin) was defined as skin present three eyelid margin widths away from the anterior lid margin (Figure 5). Lid skin and lid margin were graded on a scale of 0-3 with 0 denoting redness equivalent to the redness of the control skin and 3 denoting markedly increased redness.

For the anterior eyelid skin and the anterior half of the eyelid margin, the number of blood vessels was counted in each subarea. For the eyelid skin, vessels were counted if they were readily visible in the image and the grader could easily trace the course of the vessel. For the lid margin, vessels were counted when they 1) originated from the anterior half of the lid margin and were approximately perpendicular to the line dividing the lid margin into anterior and posterior halves; 2) were perpendicular to and crossed the dividing line; and 3) were visible enough to trace easily. For vessels that crossed between 2 subareas, the vessel was only counted if it crossed the left border of a given subarea (to avoid repetitive counting). Subsequently, the number of engorged, dilated blood vessels visible on the eyelid skin and lid margin was counted in each subarea.

Eyelid debris, collarettes, clumps and microabscesses were counted when they were present in the images of the lid margin on the half of the eyelash toward the base. Collarettes were counted in each subarea. The number of full-length eyelashes and short eyelashes was counted in each image. Short eyelashes were defined as eyelashes that were less than half of the length of adjacent full-length eyelashes. Clumps were defined as a large piece of eyelid debris that involved the basal half of at least 3 contiguous eyelashes. Microabscesses were defined as visible ulcerations or small collections present on the eyelid near the base of the eyelashes.

Statistical analysis

The agreement between graders on scoring of each feature was summarized by the percent of exact agreement and the weighted kappa statistic. Weighted kappa takes into account the degree of disagreement of individual grades between observers. The following terms are often used to describe the degree of weighted kappa agreement: slight (0.00-0.20, fair (0.21-0.40), moderate (0.41-0.60), substantial (0.61-0.80), and almost perfect (0.81-1.00).[11] The association of assessments of erythema and lid debris between ophthalmologists who examined patients and ophthalmologists who graded photographic images was assessed with Spearman correlation coefficients. For the correlation analysis, the summary score from images for an eye was computed by first calculating the median score per subarea among the 3 graders. For erythema, the 5 scores corresponding to the 5 subareas of the lower lid margin were averaged. For lid debris, the 20 scores corresponding to the 5 subareas of the upper and lower skin and lid margin were averaged.

Results

Among the 43 patients whose photographs were assessed, the mean (standard deviation) age was 64 (15) years, 27 (63%) were female, and the majority (93%) were Caucasian. The features of the eyelids are displayed in supplemental Table 1; subarea scores are the median of the scores assigned by three graders. Relatively few eyes were scored with the highest level of erythema. Approximately half of the eyes had no apparent visible vessels in the eyelid skin. The majority of lids had at least one collarette, and at least one vessel in the lid margin was graded as engorged for approximately 25% of lids. The mean (standard deviation) number of lashes within the central two-thirds of the lower lid was 39.0 (9.0) for right eyes and 38.6 (8.8) for left eyes.

Inter-grader Reliability and Correlation of Grading Scores with Clinical Assessments

The agreement on subarea scores between pairs of graders is summarized by the percentage of scores that agreed exactly and the weighted kappa (K_w) statistic (Table 2). The kappa statistic indicated the best agreement (K_w >0.80) for number of vessels per subarea, with substantial agreement for erythema, collarettes, number of engorged vessels, and number of lashes. Agreement for the number of short lashes and the presence of clumps or debris was moderate.

The correlation between grading of the photographic images and the clinical assessments of two features by the ophthalmologists examining the patients was determined. For erythema, the Spearman correlation coefficient was 0.42 for right eyes and 0.27 for left eyes (p=0.01 and 0.08, respectively). For lid debris, the Spearman correlation coefficient was 0.40 for right eyes and 0.45 for left eyes (p=0.01 and 0.003, respectively).

Systematic Variation in Characteristics Within and Between Images of Patients

We examined the data for any systematic effects or patterns present within the image sets such as increased numbers of markers of severe disease present in left versus right eyes,

upper versus lower lids, differences depending on the view (e.g., upper lid straight versus upper lid looking up), or differences due to different subareas (Table 2). For erythema, we found that lower eyelids had higher scores than upper eyelids ($p<0.0001$) and that the eyelid skin had more erythema than the eyelid margin ($p<0.0001$). For vasculature, upper eyelids were more likely to have visible vessels than lower eyelids ($p=0.003$), and images with straight ahead views were more likely to have visible vessels than eyelid margin images ($p=0.0002$). Visible vessels in the eyelids were least likely present in the most nasal subarea (subarea 1). Visible collarettes were more likely in the lower lids than the upper lids ($p=0.002$) and were more likely to be visible in the straight ahead view compared to the eyelid margin view ($p=0.005$).

The features of five subareas were analyzed for each eyelid. There were no significant subarea differences for erythema, lashes, short lashes, visible collarettes, or clumps. However, there was a statistically significant pattern of more visible vessels being present in the middle subarea (subarea 3) and more temporal subareas (subareas 4-5) compared to the nasal subareas ($P =0.001$). Also, more engorged vessels were visible in the temporal subareas (subareas 4-5) ($P = 0.035$).

Redundancy of Information Among Images of Patients

The different views of the same eyelid, comparing upper to lower eyelids of the same eye, and assessing left and right eyes of the same subjects provided different information. For example, gradings of erythema differed 47% of the time between the lower and upper eyelid skin. Straight ahead versus everted view of lower lids gave different gradings of collarettes 38% of the time.

Discussion

We were motivated by a need for a system for obtaining photographs of blepharitis patients from multiple clinical centers and for performing reliable, masked grading at a central reading center. To achieve this goal we developed a photography protocol for standardized image acquisition as well as a grading protocol that included key signs of blepharitis.

Standardized Protocol for Image Acquisition and Display

For external disease of the eye and eyelids, there are currently no standardized, widely used imaging techniques. Here for anterior blepharitis, we designed a photography protocol to minimize the variability in images resulting from different environmental conditions at the time of image acquisition, from different cameras and acquisition procedures and from different conditions of viewing. We used readily available camera equipment, and our method should work with many cameras. Using internal and external standards, as described above, allows image acquisition at multiple sites with conversion of the resultant images into a calibrated representation in the CIE XYZ or other standard color spaces. Our methodology meets the practical constraints of clinical trials where patients may be seen at many locations and where color calibration equipment and personnel skilled in its use are not available at each site. The key feature of camera performance required for our current processing pipeline is that the camera sensor responses are linear with light intensity. We verified this for our Canon T2i camera and expect that it would be the case for many comparable digital cameras. Our methods could be adapted to cameras whose responses are not linear with intensity by characterizing and correcting for the non-linearity.[12]

We also developed methods to standardize the display of images. Although the basic principles driving our method are well established[10], two features of our implementation are worth noting. First, we chose commercially available monitors that were each equipped with

a colorimetric calibration head. This allowed us to use software supplied with the monitors to place each into the same colorimetric state, and then to perform more extensive calibration measurements only for a single monitor. Thus specialized light measurement equipment is not required at all viewing locations, making standardized grading at multiple sites feasible. Second, we calibrated the system with test images passed through the same software pipeline that was to be used for the display of the images to be graded. This choice ensures that any color transformations applied by the image display software are accounted for as part of the calibration process. In turn, this means that our procedures can be applied to cases where we do not have full control over the display software. For example, in clinical trials it is sometimes desirable to have a third-party obtain images directly from the clinical sites, mask clinical information (e.g., treatment group) and then transmit them via a server/client setup to a reading center for display. Our procedures are designed to accommodate this type of situation.

Grading System for Anterior Blepharitis

To the best of our knowledge, there are no established scales that focus specifically on the assessment of anterior blepharitis. Our grading system includes the features of blepharitis typically used in assessing the condition, such as erythema and lid debris, as well as less conventionally assessed features to learn their potential utility. We also assessed the reliability, reproducibility and validity of gradings under this new system.

Reliability and Validity—We found good overall agreement among graders and moderate correlation between the gradings of digital images and gradings by the clinicians who examined the patients. The level of agreement we had among graders was substantial for erythema and collarettes (range weighted kappas 0.62-0.76). This level of agreement was similar, for instance, to the agreement achieved by graders using the Early Treatment Diabetic Retinopathy Study (ETDRS) system for the signs of diabetic retinopathy (0.41-0.8).[3] As was the case for clinical fundus examinations in the ETDRS, there are many difficult-to-control clinical variables in assessing external and surface ocular disease that hamper the utility of clinical grading schemes. Basing a grading system on the present approach to image acquisition, color calibration and display of external eye images comprises a useful starting point to improve the assessment of ocular erythema not only anterior blepharitis, but also other ophthalmic diseases.

The moderate correlation of the gradings (Spearman correlation coefficient ≈0.40) of erythema and debris from the images and the gradings by clinicians indicates that the image grade is likely to be higher when the clinical grade is high, but that there still is variation between the image grade and the grade assigned by clinicians. Only moderate correlation is expected because of the absence of specific clinical definitions of the grade levels and because of the variability in incident lighting from office to office, as occurred in the present investigation.

Anatomic Variation in Signs—We found that the gradings did depend on the eyelid, view of the lid, and subarea of the lid. Lower eyelids appeared to be more erythematous and had more collarettes than upper eyelids; eyelid skin had more erythema than the eyelid margin, and blood vessels were less likely to be observed in the nasal subareas. Therefore, the lid, view, and area of the eyelid ideally should be specified when assessing signs of anterior blepharitis. Because detailed grading of subareas requires more time with only limited added information, a summary grading of the central two-thirds of the eyelid could suffice for most purposes.

Utility of digital photographic images and image analysis in ophthalmology

The use of photographic images with standardized color correction offers many advantages over clinical grading. Examining a patient at the slit-lamp is a complex multidimensional dynamic process with variable lighting intensity, angle of the slit beam, magnification, etc. The potential to standardize such variables underlies the adoption of photography for many areas of ophthalmic diagnosis. For blepharitis, our approach provides color calibrated, standardized images. With comprehensive grading forms, basing analysis on archived images allows for a more detailed and quantitative assessment than is practical in a clinical office setting. Assessment by multiple graders, the potential for adjudication of discrepancies from archived images, and the possibility of using image analysis tools make the use of photographic assessment of eyelid disease attractive.

Limitations

Our study did not include subjects with severe blepharitis so that some signs typically observed only in severe disease, such as microabscesses, could not be assessed dependably. In addition, patients were photographed on only one occasion so that the grading system's reliability over time and sensitivity to change could not be assessed.

Conclusions

We introduce a technique for obtaining and viewing photographic images of the eyelids, using methods of photography and color processing readily adaptable to research and multi-site clinical settings. We also developed and evaluated a grading scale for the signs of anterior blepharitis. The method permits centralized grading, which can minimize grading bias and enable more detailed data collection

Our grading system showed moderate correlation with clinical gradings by independent ophthalmologists and had good reproducibility overall. The development of more objective measures for photographic features such as redness would greatly strengthen the photographic assessment of anterior blepharitis and could also facilitate the study of other ocular conditions where the assessment of erythema is important.

Although the protocol was initially developed to study anterior blepharitis, it could also be applied to other eyelid diseases. In addition to use in clinical research, it could also have applications to clinical care and to developing photographic teaching sets.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Jose Boyer, PhD for his help with the project, Jason Vittitow PhD for assistance during the development of the photography protocol, and David Hwang MD for feedback regarding the grading system for digital images. We would like to acknowledge Tom Shannon for writing the Adobe Photoshop scripts. We would also like to thank Eli Smith and Marguerite Leone for technical assistance with figure preparation.

Finally, we would like to thank the photographers at the clinical sites for their contributions: Kenneth Sall MD, Marc Abrams MD, Regina Wildey, Lacie Bellamare, and Tammy Weber.

Financial Support: Supported by Inspire Pharmaceuticals, Inc., grants from the National Eye Institute K12 EY015398(VYB) and P30 EY01583(VYB, DHB, ED, MMG, WN, EW, DJP, JH, MGM, RAS), R01-EY017299 (RAS), the Paul and Evanina Bell Mackall Foundation Trust (RAS), and an unrestricted grant from Research to Prevent Blindness (VYB, MMG, RAS).

References

1. Lemp MA, Nichols KK. Blepharitis in the United States 2009: a survey-based perspective on prevalence and treatment. Ocul Surf. 2009; 7:S1–S14. [PubMed: 19383269]
2. Nelson JD, Shimazaki J, Benitez-del-Castillo JM, et al. The International Workshop on Meibomian Gland Dysfunction: Report of the Definition and Classification Subcommittee. IOVS. 2011; 52:1930–1937.
3. Early Treatment Diabetic Retinopathy Study Research Group. Grading diabetic retinopathy from stereoscopic color fundus photographs - an extension of the Modified Airlie House classification ETDRS Report Number 10. Ophthalmology. 1991; 98:786–806. [PubMed: 2062513]
4. Klein R, Davis MD, Magli YL, et al. The Wisconsin age-related maculopathy grading system. Ophthalmology. 1991; 98:1128–34. [PubMed: 1843453]
5. International ARM Epidemiological Study Group. An international classification and grading system for age-related maculopathy and age-related macular degeneration. Surv Ophthalmol. 1995; 39:367–74. [PubMed: 7604360]
6. Chylack LT Jr, Leske MC, McCarthy D, et al. Lens opacities classification system II (LOCS II). Arch Ophthalmol. 1989; 107:991–7. [PubMed: 2751471]
7. Parrish RK 2nd, Schiffman JC, Feuer WJ, et al. Ocular Hypertension Treatment Study Group Test-retest reproducibility of optic disk deterioration detected from stereophotographs by masked graders. Am J Ophthalmol. 2005; 140:762–4. [PubMed: 16226544]
8. Madow B, Galor A, Feuer WJ, et al. Validation of a photographic vitreous haze grading technique for clinical trials in uveitis. Am J Ophthalmol. 2011; 152:170–176 el. [PubMed: 21652026]
9. CIE. Colorimetry. third. Vienna: Bureau Central de la CIE; 2004.
10. Brainard, DH.; Pelli, DG.; Robson, T. Encyclopedia of Imaging Science and Technology. New York: Wiley; 2002. Display characterization; p. 72-188.
11. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159–174. [PubMed: 843571]
12. Vora PL, Farrell JE, Tietz JD, et al. Image capture: simulation of sensor responses from hyperspectral images. IEEE Transactions on Image Processing. 2001; 10:307–316. [PubMed: 18249621]
13. Pugh JA, Jacobson JM, Van Heuven WAJ, et al. Screening for diabetic retinopathy: the wide-angle retinal camera. Diabetes Care. 1993; 16:889–95. [PubMed: 8100761]

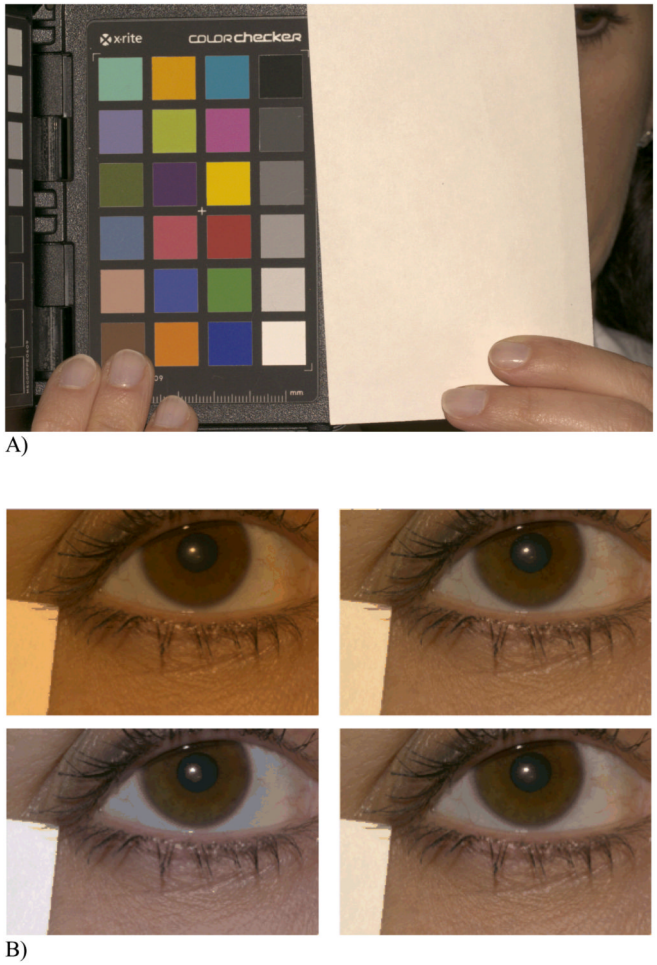

Figure 1.

A) Macbeth color checker (left) and index card (right). Image balanced to D5000 illuminant. B) Color balancing to different illuminants. Same image color balanced to four different illuminants (i.e., standard visible light sources) and then luminance corrected. Clockwise from upper left: CIE Illuminant A, 4000 °K daylight, 5000 °K daylight, CIE D65. For grading, images were color balanced to the 5000 °K daylight. We judged this color balance to be the best match to the clinical appearance of eyes when they were viewed using a slit lamp. The obvious visual differences between these images underscore the necessity of color balancing and luminance correction in the grading of images on parameters dependent on color and related properties.

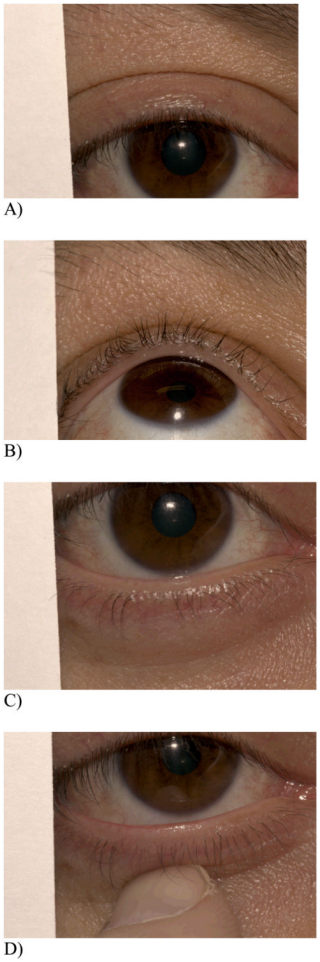

Figure 2.
4 eyelid views taken for each eye. A) Upper lid, looking straight; B) Upper lid, looking up; C) Lower lid, looking straight; D) Lower lid, everted.

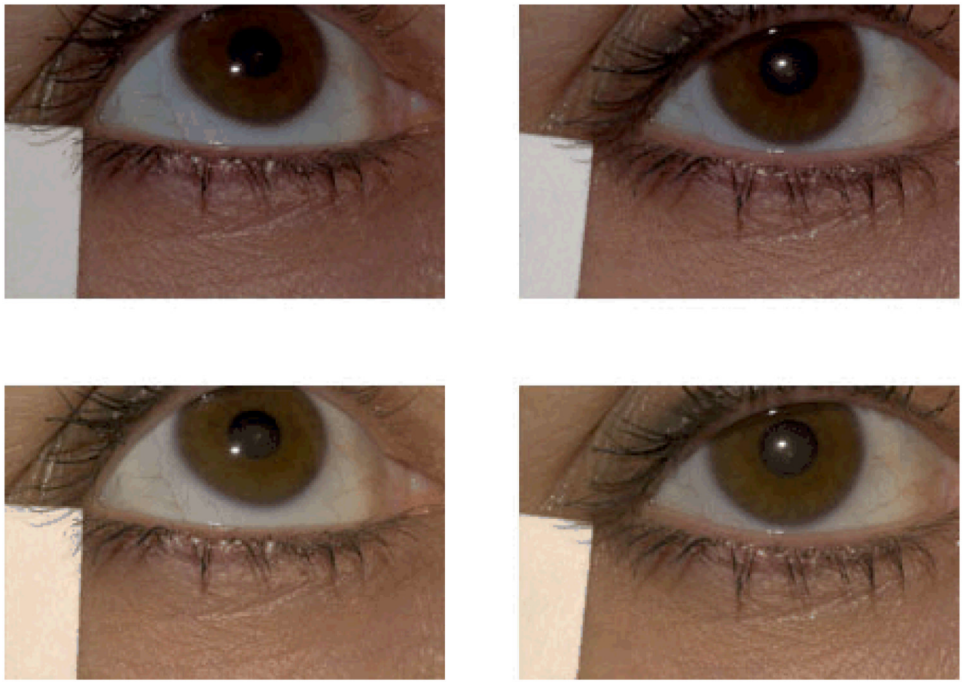

Figure 3.

Two images of the same subject, taken in the same session. The top row shows the images without color balancing or luminance correction. Note the luminance difference between the left and right images. The bottom row shows the result of color balancing and luminance correction. This corrects for much of the image-to-image variability, which is primarily variability in luminance. These images were color balanced to a D5000 illuminant, the color balance used in the grading. The colors in digital or print representations of these images may not precisely mimic the actual graded images unless the color management systems are appropriately matched. See also Figure 1B.

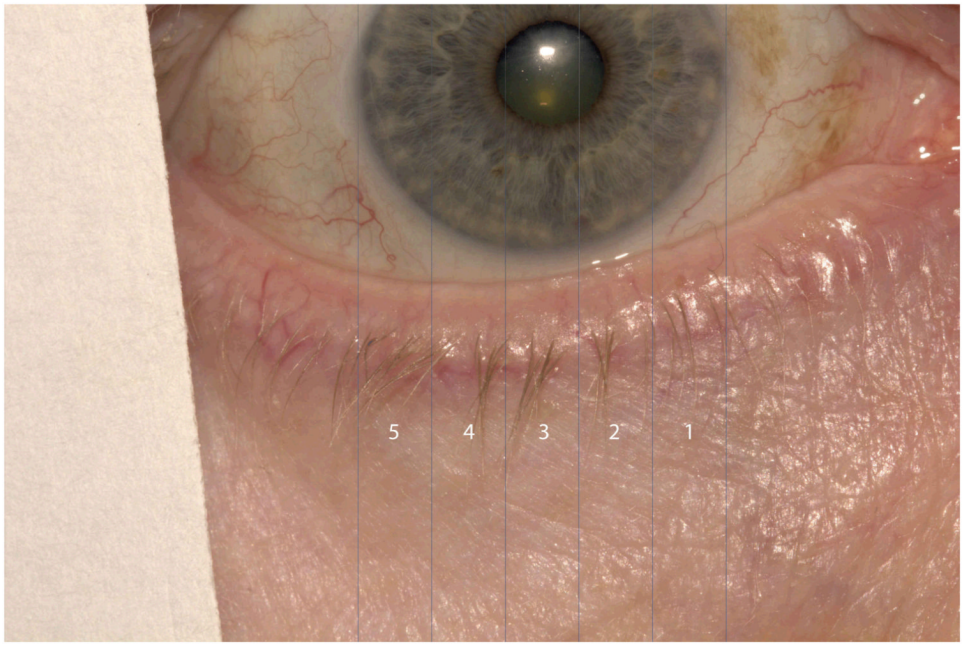

Figure 4.
Vertical lines delineating five subareas of central two-thirds of eyelid which were used for grading. Subareas were labeled #1-5, nasally to temporally for each eyelid.

Figure 5.
Rectangle delineates area of skin used as internal control for assessment of eyelid skin erythema and anterior eyelid margin erythema.

Table 1
Summary of agreement between grader pairs in gradings of features within subareas

		Grader Pair		
Feature (Subareas, N[#])	**Agreement Measure**	**A and B**	**A and C**	**B and C**
Erythema	Exact Agreement, n (%)	1039 (61.8%)	928 (54.8%)	1039 (60.8%)
1720	Weighted Kappa (CI[*])	0.73 (0.70, 0.75)	0.67 (0.64, 0.69)	0.72 (0.70, 0.75)
Vessels	Exact Agreement, n (%)	1313 (78.4%)	1281 (76.6%)	1357 (81.2%)
1720	Weighted Kappa (CI[*])	0.86 (0.84, 0.88)	0.84 (0.81, 0.86)	0.89 (0.87, 0.91)
Collarettes	Exact Agreement, n (%)	1005 (59.4%)	1017 (60.7%)	1142 (66.8%)
1720	Weighted Kappa (CI[*])	0.65 (0.61, 0.70)	0.62 (0.57, 0.67)	0.76 (0.72, 0.79)
Engorged	Exact Agreement, n (%)	746 (89.8%)	761 (91.8%)	770 (92.4%)
860	Weighted Kappa (CI[*])	0.58 (0.47, 0.70)	0.57 (0.39, 0.70)	0.75 (0.66, 0.82)
Lashes	Exact Agreement, n (%)	77 (17.9%)	178 (41.4%)	74 (17.2%)
430	Weighted Kappa (CI[*])	0.67 (0.61, 0.72)	0.86 (0.82, 0.89)	0.61 (0.57, 0.67)
Short Lashes	Exact Agreement, n (%)	215 (50.0%)	187 (43.5%)	229 (53.3%)
430	Weighted Kappa (CI[*])	0.51 (0.44, 0.61)	0.42 (0.30, 0.49)	0.58 (0.50, 0.63)
Clumps	Exact Agreement, n (%)	836 (97.2%)	839 (97.7%)	837 (97.4%)
860	Kappa (CI[*])	0.49 (0.31, 0.66)	0.54 (0.36, 0.71)	0.63 (0.49, 0.77)

[*] CI denotes 95% confidence interval

[†] Denominators for each grader pair may vary because of differing numbers of areas judged to be of adequate quality for grading.

Table 2

Multivariate analysis of the influence of lid, view and subarea on erythema, vessels, collarettes

	Mean (SE)	Odds ratio (95% CI)	
	Erythema	Vessels	Collarettes
Lid			
Lower eyelid	1.64 (0.08)	1.00	1.00
Upper eyelid	1.23 (0.10)	0.59 (0.43, 0.80)	1.63 (1.25, 2.13)
P value	<0.0001	0.003	0.002
View			
Eyelid margin	1.23 (0.09)	1.00	1.00
Eyelid skin	1.65 (0.09)	0.37 (0.24, 0.56)	0.73 (0.60, 0.89)
P value	<0.0001	0.0001	0.005
Subarea			
1	1.45 (0.09)	1.00	1.00
2	1.46 (0.09)	1.44 (1.13, 1.84)	0.89 (0.69, 1.16)
3	1.43 (0.09)	1.89 (1.46, 2.45)	0.72 (0.52, 1.02)
4	1.42 (0.09)	2.12 (1.52, 2.96)	0.95 (0.72, 1.25)
5	1.43 (0.09)	1.87 (1.44, 2.43)	1.00 (0.75, 1.33)
P value	0.84	0.002	0.19

* P values are from generalized linear models using the generalized estimating equations approach to account for correlations of repeated measures for one patient.