



Published in final edited form as:

Cornea. 2020 December ; 39(12): 1503–1509. doi:10.1097/ICO.0000000000002470.

Measurement reliability for anterior segment and keratitis morphology

Matthias F. Kriegel, MD^{1,2}, Jessica Loo, MEng³, Sina Farsiu, PhD^{3,4}, Venkatesh Prajna, FRCOphth¹, Megan Tuohy, MD¹, Kyeong Hwan Kim, MD¹, Autumn N. Valicevic, MS¹, Leslie M. Niziol, MS¹, Huan Tan, BS¹, Hamza A. Ashfaq, BS¹, Dena Ballouz, BS¹, Maria A. Woodward, MD MSc^{1,5}

¹Department of Ophthalmology and Visual Sciences, W.K. Kellogg Eye Center, University of Michigan, Ann Arbor, Michigan

²Augenzentrum am St. Franziskus Hospital Muenster, Germany

³Department of Biomedical Engineering, Duke University, Durham, NC 27708, USA

⁴Department of Ophthalmology, Duke University, Durham, NC 27708, USA

⁵Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor

Abstract

Purpose: To evaluate the reliability of manual annotation when quantifying cornea anatomical and microbial keratitis (MK) morphological features on slit lamp photography (SLP) images.

Methods: Prospectively enrolled patients with MK underwent SLP at initial encounter at two academic eye hospitals. Patients who presented with an epithelial defect (ED) were eligible for analysis. Features, which included epithelial defect (ED), corneal limbus (L), pupil (P), stromal infiltrate (SI), white blood cell infiltration at the SI edge (WBC), and hypopyon (H), were annotated independently by two physicians on SLP images. Intraclass correlation (ICC) was applied for reliability assessment; Dice Similarity Coefficients (DSC) were used to investigate the area overlap between readers.

Results: 75 MK patients with an ED received SLP. DSCs indicate good to fair annotation overlap between graders (L=0.97, P=0.80, ED=0.94, SI=0.82, H=0.82, WBC=0.83) and between repeat annotations by the same grader (L=0.97, P=0.81, ED=0.94, SI=0.85, H=0.84, WBC=0.82). ICC scores showed good intergrader (L=0.98, P=0.78, ED=1.00, SI=0.67, H=0.97, WBC=0.86) and intragrader (L=0.99, P=0.92, ED=0.99, SI=0.94, H=0.99, WBC=0.92) reliabilities. When reliability statistics were recalculated for annotated SI area, in the subset of cases where both graders agreed WBC infiltration was present/absent, intergrader ICC improved to 0.91 and DSC improved to 0.86, and intragrader ICC remained the same while DSC improved to 0.87.

Correspondence: Maria A. Woodward, MD MSc, Department of Ophthalmology and Visual Sciences, W.K. Kellogg Eye Center, University of Michigan, Ann Arbor, Michigan, 1000 Wall Street, Ann Arbor, Michigan 48105; Telephone: 734-763-6967, Fax: 734-232-2332, mariawoo@med.umich.edu.

Conflict of Interest: The authors have no proprietary or commercial interest in any of the materials discussed in this article.

Conclusion: Manual annotation indicates usefulness of area quantification in the evaluation of MK. However, variability is intrinsic to the task. Thus, there is a need for optimization of annotation protocols. Future directions may include using multiple annotators per image or automated annotation software.

Keywords

cornea; microbial keratitis; image analysis

Microbial keratitis (MK) manifests in diverse presentations. The non-uniform nature of MK morphology requires evaluating subtle differences across time for each patient. Ophthalmologists must be able to recognize the differences between patients and within patients over time to tailor treatment and improve outcomes. Currently, manual caliper measurements from the slit-lamp biomicroscope, free-text notes, and schematic drawings in the electronic health record are the standard method to measure and record MK morphologic features. However, the caliper tool measures linear dimensions of height and width of a three dimensional pathologic process. Some slit lamp biomicroscopes do not have a caliper measurement tool, so ophthalmologists use manual rulers. Even when the tool is available, ophthalmologists do not always record the slit lamp biomicroscope caliper measurements in the electronic health record,¹ and measurements differ among ophthalmologists even in controlled settings.^{2,3}

In order to overcome these limitations, a standardized strategy to record and quantify corneal anatomic and pathologic features is needed. Images of the cornea can be used to improve our measurement of corneal anatomy and pathology.⁴ Slit lamp photography (SLP) is a low-cost method available in most eye clinics that provides an opportunity for a standardized image-based approach. In our prior work, SLP images were used to develop an image-based, semi-automated algorithm to quantify two MK morphologic features: the epithelial defect (ED) and stromal infiltrate (SI).² This quantification was more reliable than measurements between physicians. Beyond quantification, the use of images creates an opportunity to build anterior eye image databases that exist in posterior eye imaging.⁵ Diagnostic algorithms for automated SLP image analysis may also be possible.⁶ While anterior segment imaging is complex, the utility of SLP to build an annotated library of corneal anatomy and pathology has potential to improve our understanding of anterior eye diseases.

The purpose of this study was to evaluate the reliability of clinicians' annotations of both corneal anatomy and pathology on SLP images. SLP images can be manually annotated to quantify corneal anatomy and pathology on each patient, but annotated images are also a resource to develop image-analysis algorithms. The reliability of manual annotation techniques is critical when comparing the performance of image-analysis algorithms to an image-based gold standard.

Methods

The study adhered to the tenets of the Declaration of Helsinki. IRB approval from the University of Michigan (UM) and ethics committee approval from the Aravind Eye Care System (AECS) were obtained. Inclusion criteria were adult patients with MK. Patients were

excluded if they were from a vulnerable population (i.e., hospitalized, imprisoned, institutionalized), pregnant, or not able to provide consent. Eyes were excluded if the eye did not have an epithelial defect.

Prospectively enrolled participants with MK underwent SLP at their first clinical encounter at two academic eye hospitals (AECS or UM). SLP images were taken under diffuse white light and blue light illumination following a SLP protocol.² One white and one blue light SLP image were chosen for analysis. Two physicians, masked to the patients' clinical information, independently annotated features on the SLP images using ImageJ software (National Institute of Health, Bethesda, MD) on white light SLP (Figure 1A) and on blue light SLP (Figure 1B). One physician (MFK) repeated annotation of the same photos four months later for intragrader reliability assessment.

On white light SLP images, annotated corneal anatomy included the corneal limbus and pupil, and annotated MK pathology included the stromal infiltrate (SI), white blood cell infiltration (WBC), and hypopyon, if present. On blue light SLP images, annotated MK pathology was the epithelial defect (ED). Each annotated image was processed with ImageJ software (National Institutes of Health, Bethesda, MD; available at: <http://imagej.nih.gov/ij>) to obtain area measurements in squared pixels for each feature.

Areas of MK features were measured as squared pixels from a photograph. Pixel area was converted to area in squared millimeters (mm^2) by using the published average adult limbus diameter of 11.7 mm as a mean white-to-white distance to provide meaningful descriptive statistics.⁷ This distance was compared to the same diameter measured in pixels from the SLP image, and the ratio provided a mm/pixel conversion factor. This factor was calculated for each individual image given their unique resolutions. Areas of features in pixels (as calculated by ImageJ) were multiplied by the pixel conversion factor to get area in mm^2

Statistical Methods

Statistical analysis was carried out using R version 3.6.1 (Vienna, Austria). Patient demographics and areas of clinical features of the MK sample were summarized with descriptive statistics, including means, standard deviations (SD), frequencies, and percentages. The absolute difference in the area of the annotated features between graders and repeat annotation by the same grader were also presented with descriptive statistics. Annotations were evaluated for inter- and intragrader reliability using Dice similarity coefficient (DSC) and intraclass correlation coefficients (ICC).^{8,9} Results are displayed with 95% confidence intervals (CI). The DSC represents the overlap in area of annotation between physicians or repeat annotations of the same SLP image, and values above 0.7 represent a high score for intergrader and intragrader comparisons. DSC was only calculated when both physicians annotated the features. For intergrader reliability, ICCs were calculated using 2-way random effects models for absolute agreement; for intragrader reliability, ICCs were calculated using 2-way mixed effects models for absolute agreement.

Two sensitivity analyses were performed with respect to WBC. First, ICC and DSC scores for the SI were calculated only in the sample where both graders agreed about the presence/absence of WBCs. This was to account for the disagreements among graders concerning the

presence of a WBC border. Second, a subsequent analysis was performed where the area of both the SI and WBC infiltrate were combined and ICC and DSC were calculated. This was to account for the discrepancy of WBC being included or not included in the SI area annotation.

Results

A total of 75 images from 75 patients were analyzed. Participants were 33% female and were on average 50.4 years (SD=14.9). The sample included 19 patients (25.3%) from UM and 56 (74.7%) from AECS. MK was caused by fungal organisms in 39 cases (52.7%), bacteria in 14 cases (18.9%), both bacteria and fungi in 1 case (1.4%), and acanthamoeba in 3 cases (4.1%). Cultures yielded no growth in 17 cases (23.0%)

Table 1 summarizes the size of anatomic landmarks and MK features by grader and repeat annotation. Stromal infiltrates were on average 9.9 mm² (SD=9.2) when annotated by grader 1 and 12.8 mm² (SD=12.9) by grader 2. This resulted in an average absolute difference between graders of 4.1 mm² (SD= 9.9). The repeat annotation by grader 2 of the stromal infiltrate showed areas on average 11.6 mm² (SD=12.5), for an average absolute intragrader difference of 2.7 mm² (SD=3.6). The ED showed the smallest difference between measurements with average absolute intergrader differences of 0.7 mm² (SD=1.2) and absolute intragrader differences of 0.8 mm² (SD=1.3). Alternatively, SI and WBC combined area showed the largest average absolute intergrader difference of 6.4 mm² (SD=9.5) and limbus showed the largest difference between intragrader measurements with average absolute intragrader difference of 5.4 (SD=3.5). The comparison of the feature area (mm²) annotated between graders and within graders is shown in Figures 2 and 3, respectively. Measurement size showed good agreement between graders with ED and hypopyon showing less discrepancy than limbus, pupil, SI, and WBC infiltration features. However, each feature had some outliers showing large disagreement in measurement size, with SI annotations having the most variability. Similar results were observed for intragrader annotation comparisons, but with less discrepancy between measures.

DSC scores for anatomic landmarks and each MK feature were calculated and are displayed in Table 2. On average, all features showed good agreement between graders or on repeat annotation by the same grader, with DSCs above 0.7. The lowest DSC score was the pupil measurement with average intergrader overlap of 0.80 (95% CI: 0.73, 0.87) and intragrader overlap of 0.81 (95% CI: 0.75, 0.86). The highest DSC scores were observed for the limbus with average intergrader and intragrader overlap of 0.97 (95% CI: 0.96, 0.97). Graders agreed on the presence of a WBC border in 53% of the images (40 of 75). The DSC score for the SI, in the subset of cases where graders agreed on WBC infiltration presence or absence, showed average intergrader overlap of 0.86 (95% CI: 0.81, 0.91) and intragrader overlap of 0.87 (95% CI: 0.84, 0.90). The DSC score for combined SI and WBC areas was on average 0.76 (95% CI: 0.71, 0.81) for intergrader annotations and 0.84 (95% CI: 0.81, 0.87) for intragrader annotations.

ICCs for each of the anatomic landmarks and MK features are shown in Table 2. The lowest reliability for intergrader comparisons was for the SI with ICC of 0.67 (95% CI: 0.52, 0.78),

while the lowest reliability for intragrader comparisons was for WBC and pupil with ICC of 0.92 (95% CI: 0.87, 0.95) and 0.92 (95% CI: 0.87, 0.96), respectively. The highest reliability for intergrader comparisons was for the ED with ICC of 0.99 (95% CI: 0.99, 1.00), and the highest reliability for intragrader comparisons was for the limbus and ED with ICC of 0.99 (95% CI: 0.98, 0.99) and 0.99 (95% CI: 0.98, 0.99), respectively. Reliability of SI area increased, when the sample was restricted to cases where graders agreed on WBC infiltration presence or absence, with an intergrader ICC of 0.91 (95% CI: 0.80, 0.95) and intragrader ICC of 0.94 (95% CI: 0.90, 0.96). The combined SI and WBC area had an intergrader ICC of 0.88 (95% CI: 0.71, 0.94) and intragrader ICC of 0.94 (95% CI: 0.91, 0.96).

Discussion

This study addresses the important challenge of evaluating the reliability of a standardized approach to quantify corneal anatomic and pathologic features. Using SLP images, physicians had high reliability in detecting the boundaries of anatomic and pathologic features. The study was conducted with images of corneal pathology from two countries, the United States and India, improving generalizability. Intergrader ICC ranged from 0.67 (95% CI: 0.52, 0.78) to 0.99 (95% CI: 0.99, 1.00). ICC was highest for annotation of the limbus and ED and lowest for SI and pupil. DSCs showed high reliability for the anatomic and pathologic features annotated between graders and for repeat annotation by the same grader. The ED and limbus also had the highest average DSC scores at 0.94 (95% CI: 0.93, 0.95) and 0.97 (95% CI: 0.96, 0.97), respectively. There was also a high average intergrader DSC of 0.82 (95% CI: 0.77, 0.87) observed for SI and DSC of 0.83 (95% CI: 0.77, 0.88) observed for WBC.

For corneal specialists who routinely deal with ulcers, the epithelial defect and hypopyon are two easily identifiable features and thus we would not expect much variability with respect to demarcating these features on photographs. In fact, both intergrader and intragrader area measurements showed remarkable agreement (Figure 2 and Figure 3). This is reflected in the very high ICC values for these features, where the variability between repeat measurements of the same ulcer is very small compared to variability in size between ulcers. Even though the ICC statistics show near perfect reliability, that doesn't mean the annotations are exactly the same, just that the inconsistencies don't result in large differences in areas. The DSC better picks up on small inconsistencies between image annotations and thus shows lower (but still excellent) reliability compared to the ICC.

The differences in ICC scores between features may be due to several factors. First, the borders for corneal limbus and EDs are more easily discernable and more uniform than those for the SI and WBC border. Second, the WBC border is not easily distinguishable from the SI border. Graders annotated an SI border in 100% of the images; however, they only agreed on the presence of a WBC border in 53% of the images (40 of 75). In cases where the graders disagreed upon the presence of a WBC border, the grader who did not annotate a WBC border usually had an SI border that was broader than the grader who annotated both an SI and WBC border. In cases with WBC presence was in agreement by both graders, the intergrader ICC and DSC improved. As a secondary analysis, we analyze the SI and WBC as

a combined area. After aggregating WBC and SI borders into one measurement, the intergrader ICC was higher compared to SI and WBC measured independently. This highlights the need to explore the measurements of WBC versus SI borders in future work, as most clinicians recognize that WBC infiltrates are dynamic, so a distinct border does not appear until the keratitis begins to heal.

Improving the process to quantify MK anatomy and pathology has important clinical implications. The size, location, and depth of MK features have significant effects on visual prognosis after treatment.¹⁰⁻¹² Accurate identification and recording of morphological factors play an important role in influencing the treatment regimen initiated, as well as monitoring of clinical improvement throughout the course of treatment. Current methods to record eye findings include SLP, caliper measurements, and sketches of clinical findings. However, images are often not utilized, and caliper measurements are limited in their capacity to capture pathology and measures of area. The issue of consistent measurements is compounded when a patient will be managed by multiple cornea and comprehensive ophthalmologists, given inter-observer variability.² Recent advances in imaging, technology, and software may help clinicians overcome those challenges.

Use of images as part of clinical care are being increasingly adopted by ophthalmologists. With the entry of robust image-analysis algorithms, the reliability of image analysis becomes a critical area of study. Image-based manual annotation has been studied in images of choroidal thickness, retinopathy of prematurity (ROP), and retinal hemorrhages.¹³⁻¹⁵ Our intergrader ICC values (based on a continuous measure) range from 0.67 to 0.99 and are greater than the intergrader reliability results of other studies that mainly used kappa values (based on categorical measures). A 2014 multicenter study looking at the image-based diagnosis of ROP had moderate to substantial intergrader agreement, with weighted kappa values of 0.586 to 0.723.¹⁶ A major multicenter study examining diabetic retinopathy classification resulted in weighted kappa for intergrader reliability of 0.41 to 0.80.¹⁷

There are limitations with manual annotation in the quantification of clinical features. As previously mentioned, physicians can disagree on the presence of WBC which may subsequently affect an accurate measurement of stromal infiltrates. A potential solution for this may be to revise grader training on the clinical presentation of WBC borders with examples of WBC versus SI borders. Patients were only included if they had an ED present, thereby excluding some cases of deeper MK. Only two graders were involved in annotation, although consistent with other published studies, may limit generalizability to the broader clinical community. The third dimension is complex with photographic imaging. When physicians examine the eye, our brains create a three-dimensional picture from millions of nanosecond slices of the images seen with the slit beam. In future work, we will take multiple slit beam photographs of the ulcer to recreate an evaluation of depth and explore different imaging modalities better suited to evaluating depth.

In conclusion, manual annotation of corneal anatomic and pathologic features based on SLP images showed high reliability indicating usefulness in clinical quantification of features and in future applications including image analysis algorithms. SLP offers a method for features quantification that could even be explored in other pathologies beyond MK and potentially

used to diagnose disease. Having more images in clinical practice will increase their accessibility to clinicians and could help to improve overall coordination and quality of care.¹⁸ The next steps for our manual annotation technique is to apply manual annotations to create advanced automated image analysis algorithms.

Acknowledgements:

Personnel effort (MFK) for this research at the University of Michigan was supported by Ms. Susan Lane.

Disclosure of funding: This work was supported by the National Institutes of Health R01EY031033 (MAW) (Rockville, MD). The funding organization had no role in study design or conduct, data collection, management, analysis, interpretation of the data, decision to publish, or preparation of the manuscript. MAW had full access to the data and takes responsibility for the integrity and accuracy of the data analysis.

References

1. Maganti N, Tan H, Niziol LM, et al. Natural Language Processing to Quantify Microbial Keratitis Measurements. *Ophthalmology*. 2019; 126:1722–1724. [PubMed: 31307829]
2. Patel TP, Prajna NV, Farsiu S, et al. Novel Image-Based Analysis for Reduction of Clinician-Dependent Variability in Measurement of the Corneal Ulcer Size. *Cornea*. 2018;37:331–339. [PubMed: 29256985]
3. Parikh PC, Valikodath NG, Estopinal CB, et al. Precision of Epithelial Defect Measurements HHS Public Access. *Cornea*. 2017;36:419–424. [PubMed: 28129296]
4. Esaka Y, Kojima T, Dogru M, et al. Prediction of Best-Corrected Visual Acuity With Swept-Source Optical Coherence Tomography Parameters in Keratoconus. *Cornea*. 2019;38:1154–1160. [PubMed: 31394554]
5. Kauppi T, Kalesnykiene V, Kamarainen JK, et al. The DIARETDB1 diabetic retinopathy database and evaluation protocol. *BMVC – Proceedings of the British Machine Vision Conference*. 2007;1:1–18.
6. Almazroa AA, Alodhayb S, Osman E, et al. Retinal fundus images for glaucoma analysis: the RIGA dataset In: Zhang J, Chen P-H, eds. *Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications*. Vol 10579 SPIE; 2018:8.
7. Baumeister M, Terzi E, Ekici Y, et al. Comparison of manual and automated methods to determine horizontal corneal diameter. *J Cataract Refract Surg*. 2004;30:374–380. [PubMed: 15030827]
8. Zou KH, Warfield SK, Bharatha A, et al. Statistical validation of image segmentation quality based on a spatial overlap index. *Acad Radiol*. 2004;11:178–189. [PubMed: 14974593]
9. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*. 2016;15:155–163. [PubMed: 27330520]
10. Miedziak AI, Miller MR, Rapuano CJ, et al. Risk factors in microbial keratitis leading to penetrating keratoplasty. *Ophthalmology*. 1999;106:1166–1171. [PubMed: 10366087]
11. Schaefer F, Bruttin O, Zografos L, et al. Bacterial keratitis: A prospective clinical and microbiological study. *Br J Ophthalmol*. 2001;85:842–847. [PubMed: 11423460]
12. Efron N, Morgan PB, Hill EA, et al. The size, location, and clinical severity of corneal infiltrative events associated with contact lens wear. *Optom Vis Sci*. 2005;82:519–527. [PubMed: 15976590]
13. Philip AM, Gerendas BS, Zhang L, et al. Choroidal thickness maps from spectral domain and swept source optical coherence tomography: Algorithmic versus ground truth annotation. *Br J Ophthalmol*. 2016;100:1372–1376. [PubMed: 26769670]
14. Chiang MF, Keenan JD, Du YE, et al. Assessment of image-based technology: impact of referral cutoff on accuracy and reliability of remote retinopathy of prematurity diagnosis. *AMIA Annu Symp Proc*. 2005:126–130. [PubMed: 16779015]
15. Takashima Y, Sugimoto M, Kato K, et al. Method of Quantifying Size of Retinal Hemorrhages in Eyes with Branch Retinal Vein Occlusion Using 14-Square Grid: Interrater and Intrarater Reliability. *J Ophthalmol*. 2016; Epub.

16. Ryan MC, Ostmo S, Jonas K, et al. Development and Evaluation of Reference Standards for Image-based Telemedicine Diagnosis and Clinical Research Studies in Ophthalmology. *AMIA Annu Symp Proc.* 2014;2014:1902–1910. [PubMed: 25954463]
17. 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]
18. Chiang MF, Boland MV., Brewer A, et al. Special requirements for electronic health record systems in ophthalmology. *Ophthalmology.* 2011;118:1681–1687. [PubMed: 21680023]

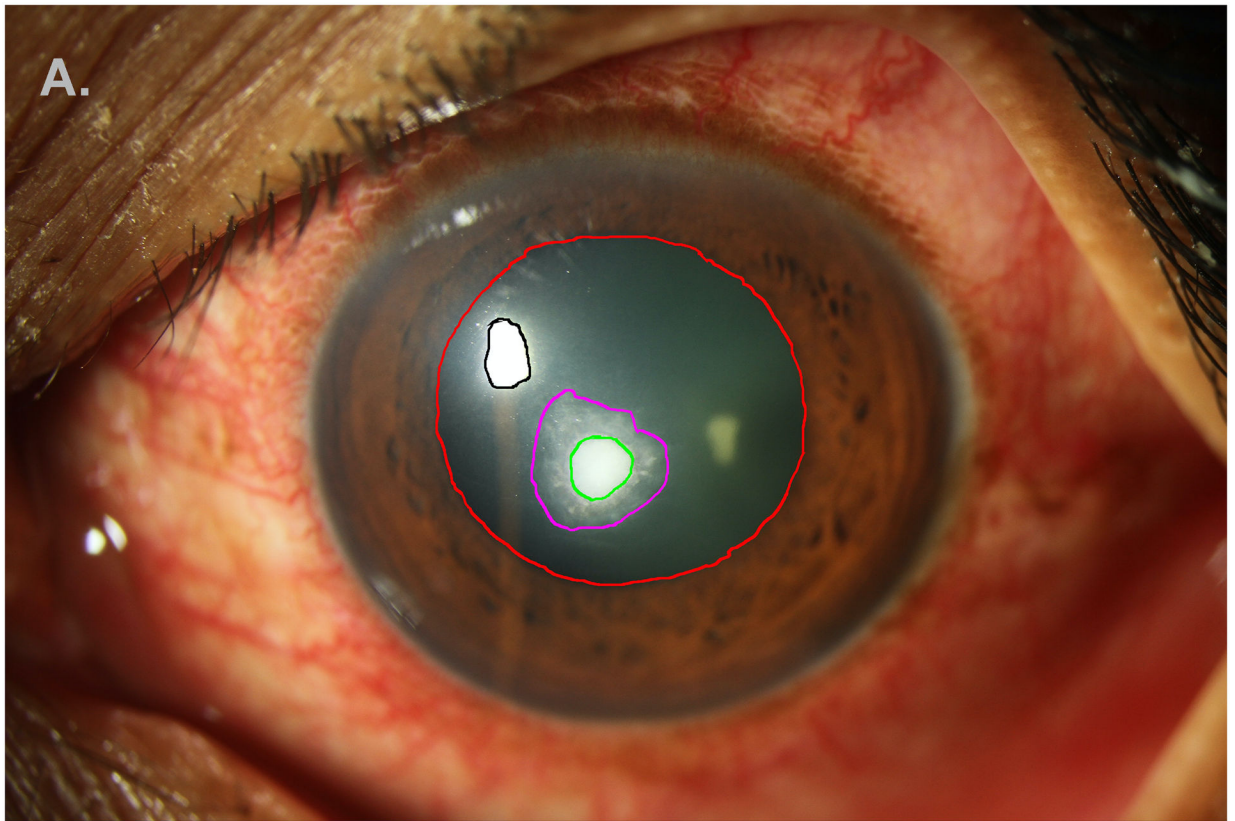


Figure 1A. Annotation of anatomic and pathologic features on white-light SLP: ulcer (green), white blood cells (magenta), light reflex (black), pupil (red). Limbus and hypopyon are not shown.

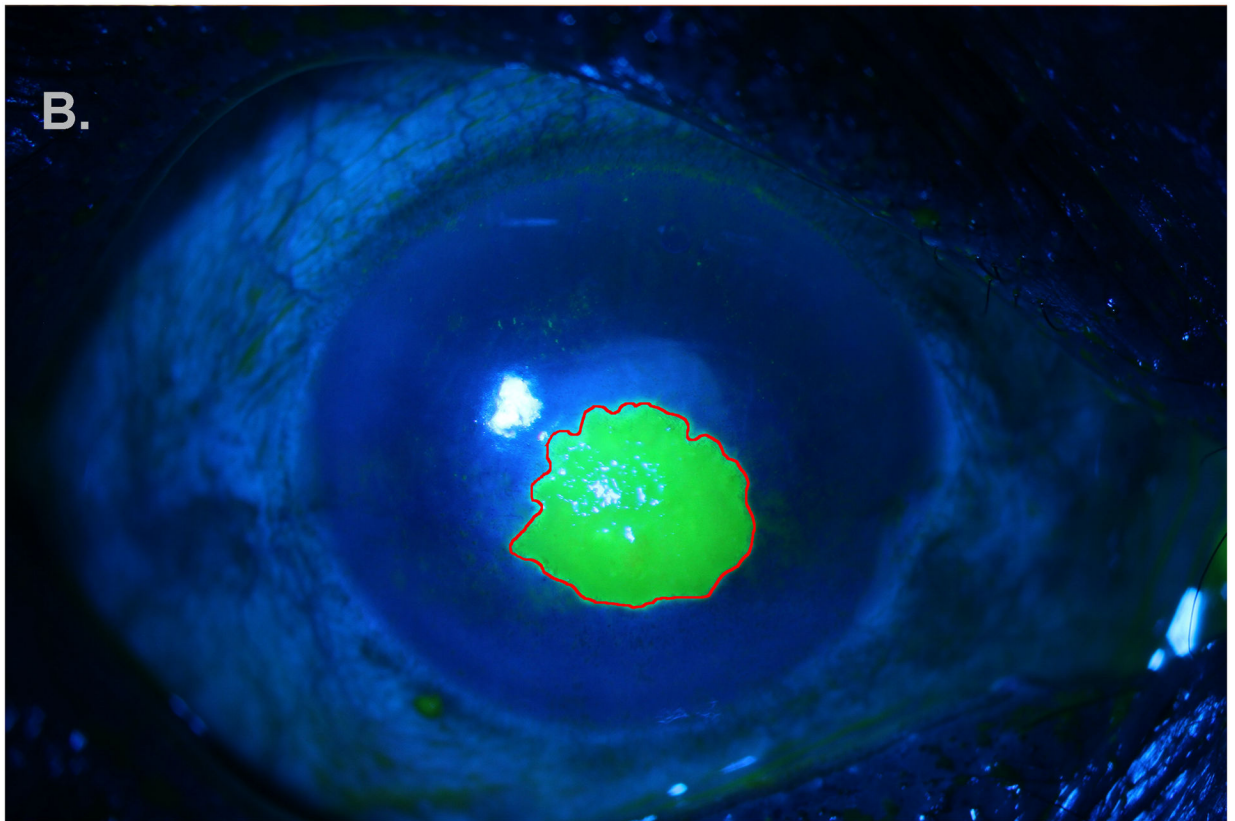


Figure 1B.
Annotation (red) of epithelial defect stained with fluorescein on blue-light SLP.

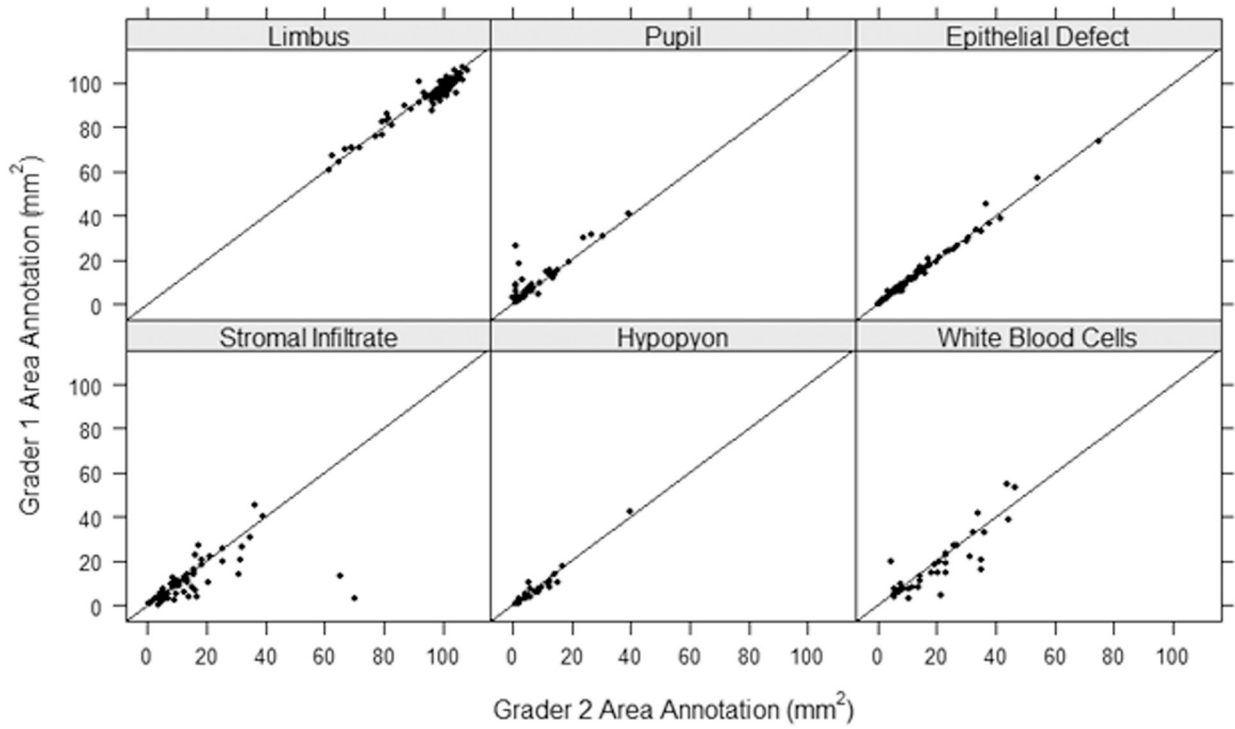


Figure 2. Intergrader (Grader 1 versus Grader 2) agreement scatter plot of the area annotated for each of the anatomic and pathologic features with the line $y=x$ as a reference for exact agreement.

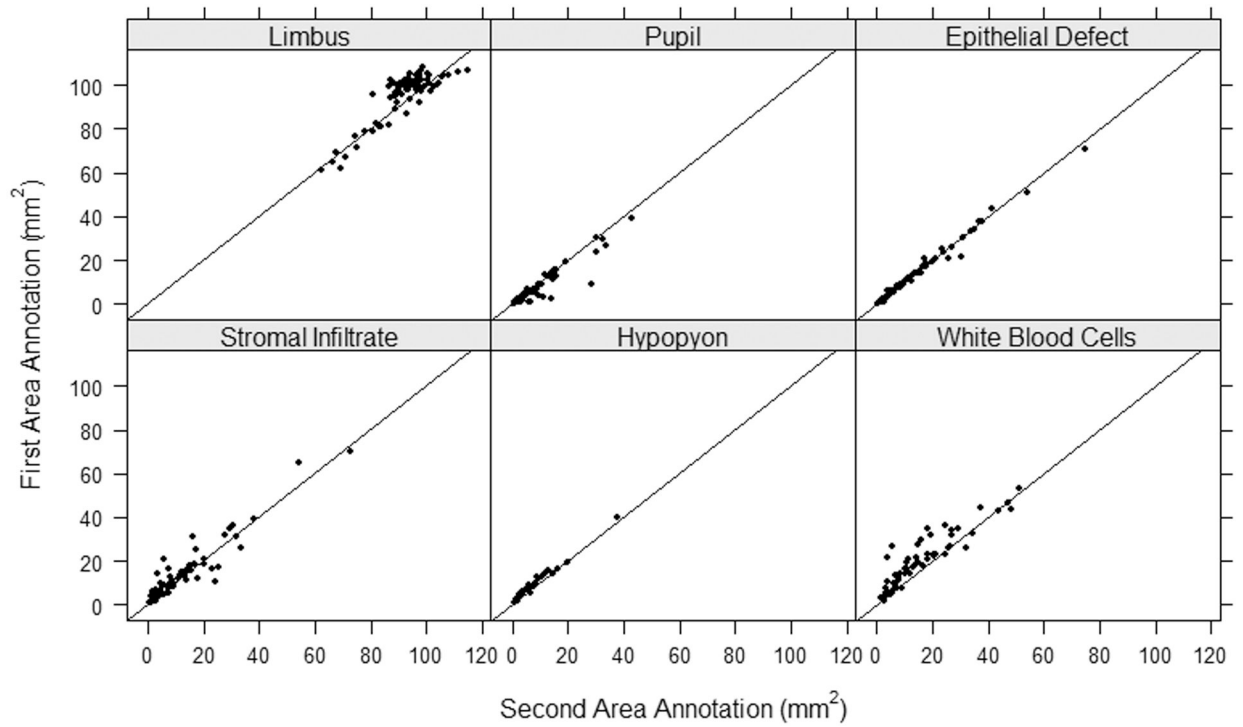


Figure 3.

Intragrader (Grader 1: Annotation 1 vs Annotation 2) agreement scatter plot of the area annotated for each of the anatomic and pathologic features with the line $y=x$ as a reference for exact agreement.

Table 2.

Intergrader and Intragrader Intraclass Correlation Coefficients (ICC) and Dice Similarity Coefficient (DSC)

Intergrader						
Feature	N	ICC	95% Confidence Interval	DSC	95% Confidence Interval	1- DSC
Limbus	75	0.98	0.88, 0.99	0.97	0.96, 0.97	0.03
Pupil	52	0.78	0.63, 0.87	0.80	0.73, 0.87	0.20
Epithelial Defect	75	0.99	0.99, 1.00	0.94	0.93, 0.95	0.06
Stromal Infiltrate	75	0.67	0.52, 0.78	0.82	0.77, 0.87	0.18
Stromal Infiltrate (in Sample with White Blood Cell Presence Agreement)	40	0.91	0.80, 0.95	0.86	0.81, 0.91	0.14
Stromal Infiltrate and White Blood Cells	75	0.88	0.71, 0.94	0.76	0.71, 0.81	0.24
Hypopyon	28	0.97	0.84, 0.99	0.82	0.78, 0.86	0.18
White Blood Cells	36	0.86	0.72, 0.93	0.83	0.77, 0.88	0.17
Intragrader						
Limbus	75	0.99	0.98, 0.99	0.97	0.96, 0.97	0.03
Pupil	55	0.92	0.87, 0.96	0.81	0.75, 0.86	0.19
Epithelial Defect	75	0.99	0.99, 1.00	0.94	0.92, 0.96	0.06
Stromal Infiltrate	75	0.94	0.91, 0.96	0.85	0.82, 0.88	0.15
Stromal Infiltrate (in Sample with White Blood Cell Presence Agreement)	60	0.94	0.90, 0.96	0.87	0.84, 0.90	0.13
Stromal Infiltrate and White Blood Cells	75	0.94	0.91, 0.96	0.84	0.81, 0.87	0.16
Hypopyon	31	0.99	0.98, 0.99	0.84	0.81, 0.87	0.16
White Blood Cells	59	0.92	0.87, 0.95	0.82	0.78, 0.86	0.18