Reliability of Physicians' Measurements when Manually Annotating Images of Microbial Keratitis

#2108

KELLOGG EYE CENTER
MICHIGAN MEDICINE

Matthias F. Kriegel, MD[1,2], Jessica Loo, MS[3], Venkatesh Prajna, FRCOphth[4], Sina Farsiu, PhD[3], Megan Tuohy, MD[1], Preethi M. Gompa, MD[4], Leslie M. Niziol, MS[1], Maria A. Woodward, MD, MS[1]

1.University of Michigan, Kellogg Eye Center 2. Augenzentrum am St. Franziskus Hospital Muenster, Germany. 3. Duke University 4. Aravind Eye Care System

INTRODUCTION

- Microbial keratitis is a sight threatening disease
- Automated algorithms may help in the future to identify and quantify the infiltrate size of a corneal ulcer on slit-lamp photographs [1]
- To train algorithms that are based on convolutional neural networks (CNN) large data sets are required [2]
- These data sets have to be annotated by experienced graders
- In order to assess how well the algorithm might perform, eventually, it's important to assess the inter-reader reliability by gold-standard physician annotation

PURPOSE

- To determine the reliability of physicians to annotate microbial keratitis on slit-lamp images

METHODS

- Prospectively-enrolled participants with microbial keratitis underwent slit-lamp imaging at their first clinical encounter at two academic eye hospitals (Aravind Eye Care Systems, AECS, and University of Michigan, UM)
- Two physicians outlined the borders of the stromal infiltrate on diffuse, white-light images using ImageJ software (National Institute of Health, Bethesda, MD) (Figure 1)
- Physicians were masked to participants' clinical information and to second grader annotation
- After annotation area of ulcer was calculated. Image pixels were used as the unit of analysis because white-to-white distance was not measured
- Physicians' annotations were evaluated for reliability using DICE similarity coefficient (DSC) and intraclass correlation coefficient (ICC)
- The DSC is the proportion of twice the number of pixels identified by both graders divided by the sum of pixels identified in each image. The DSC ranges from 0 (no overlap) to 1 (perfect overlap) and a coefficient >0.8 is considered good

Figure 1: Annotation of stromal infiltrate using the "Freehand selections" tool in Image J

RESULTS

Patient characteristics (n=120)			Pathogens of ulcers
	Mean ± Standard deviation	Median (min-max)	Fungal
			64 (53.3%)
Age	49.0 ± 14.2	49.5 (21-88)	Bacterial
			21 (17.5%)
	Female	Male	Bacterial+Fungal
			1 (0.8%)
Gender	50 (41.7%)	70 (58.3%)	Acanthamoeba
			3 (2.5%)
	Aravind	Kellogg	Unidentified
			31 (25.8%)
Eye hospital	101 (84.2%)	19 (15.8%)	

Table 1: patient characteristics and pathogens causing ulcer. 120 eyes of 120 patients

COMPARISON OF DIFFERENT GRADERS

Pixel area by different graders	Mean ± Standard deviation	Median	Minimum	Maximum
Grader 1	434,083 ± 506,254	251,307	9,431	2,459,555
Grader 2	564,725 ± 670,304	327,681	15,978	4,670,410

Table 2: pixel area measured by two different graders

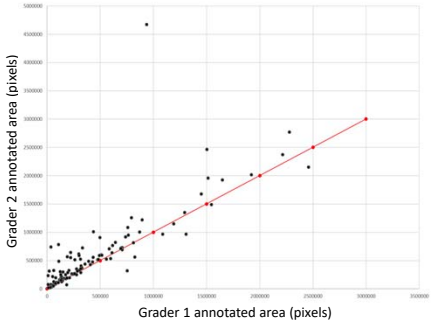

Figure 2: Comparison of pixel area between two different graders. X- and Y-axis displaying number of pixels

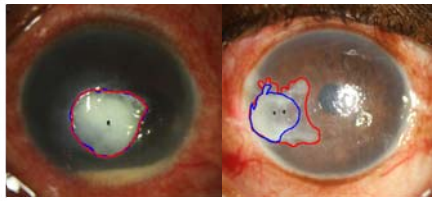

Figure 3: Corneas with infectious ulcer. Comparison of manual tracing of reader 1 (red area) and reader 2 (blue area); left: DSC=0.98; right: DSC=0.70

	ICC	95% CI
Stromal infiltrate pixel area	0.79	0.72-0.85

Table 3: ICC of 0.79 showing good agreement between graders

Overlap between graders (DSC)	Mean ± Standard deviation	Median	Minimum	Maximum
DSC	0.79 ± 0.21	0.88	0	0.98

Table 4: DICE similarity coefficient (DSC) showing the actual overlap of the annotations between graders

LIMITATIONS

- We only reported pixel area because White-to-white distance was not available for all eyes
- We could not asses the training effect during the course of the annotations of graders and its impact on variability in corneal ulcer measurement

CONCLUSIONS

- Physicians can reliably identify and trace the infiltrate for microbial keratitis from slit-lamp imaging; however, physicians do have some differences when manually tracing images
- Training and review of test images with graders will be important if manual tracing is to be used to design automated image analysis software

REFERENCES AND DISCLOSURES

1. Novel Image-Based Analysis for Reduction of Clinician-Dependent Variability in Measurement of the Corneal Ulcer Size
Patel et al. Cornea. 2018 Mar;37(3):331-339. doi: 10.1097/ICO.0000000000001488
2. Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search
Fang et al. Biomed Opt Express. 2017 Apr 27;8(5):2732-2744. doi: 10.1364/BOE.8.002732. eCollection 2017 May 1

Disclosures: Matthias F. Kriegel, None; Jessical Loo, None; Venkatesh Prajna, None; Sina Farsiu, None, Megan Tuohy, None; Preethi M. Gompa, None; Leslie M. Niziol, None; Maria A. Woodward, NIH (Grant: #K23EY023596)

ACKNOWLEDGMENTS

National Eye Institute
NATIONAL INSTITUTES OF HEALTH

CONTACT

Please direct questions or comments to
Maria Woodward (mariawoo@med.umich.edu) or
Matthias Kriegel (kriegel.matthias@gmail.com)