A Deep Learning Model for Detection of Ocular Surface Squamous Neoplasia in Slit Lamp Photos

Osmel P. Alvarez, BA; Jason Greenfield, BA; Rafael Scherer, MD, PhD; Douglas Rodrigues da Costa, MD; Diego Alba, MD; Michael Antonietti, BS; William Herskowitz, BA; Sofia de Arrigunaga, MD; Anat Galor, MD, MSPH; Swarup S. Swaminathan, MD; Felipe Medeiros, MD, PhD; Carol L. Karp, MD

Funding: Singerman and Schwartz Research Scholarship (Bascom Palmer Eye Institute, University of Miami Miller School of Medicine)

Background: Ocular surface squamous neoplasia (OSSN) is a potentially vision-threatening malignancy and may be difficult to diagnose as it often appears similar to other lesions such as pterygium and pingueculae. Ocular surface images are easy to obtain with a slit lamp or even a smart phone. The purpose of this study is to develop a deep learning model to differentiate OSSN from benign ocular surface lesions in slit lamp pictures (SLPs).

Methods: In this retrospective study, SLPs of patients seen at Bascom Palmer Eye Institute from 2012 to 2023 with biopsy proven ocular surface lesions were extracted from a database and labeled with biopsy results. 48,781 SLPs were used to train a masked autoencoder. 3,772 additional SLPs (from 241 eyes of 186 patients) with biopsy-proven OSSN or pterygia/pinguecula diagnoses were split at patient level into training, validation and testing sets, and used to train models with a supervised vision transformer. Performance was assessed on testing set.
To compare the model to human graders, one image from each eye in the testing set was randomly chosen. This subset was shown to four eye care professionals not specialized in ocular surface oncology (two general ophthalmologists, two general optometrists); each was asked to label each image as either containing OSSN or pterygium/pinguecula. This labeling task was repeated with the same images in a different order one week after the first instance. The four graders' performance (accuracy, sensitivity, specificity) and interrater agreement was calculated.

Results: The best-performing model had an accuracy of 89.9% (95% confidence interval [CI]: 87.7-91.8%), sensitivity of 95.1% (95% CI: 93.3-96.5%) and specificity of 47.4% (95% CI: 37.2-57.8%). The area under the ROC curve was 0.85.
The best performance by a human grader on the comparison subset showed an accuracy of 60.6%, sensitivity of 58.3%, and specificity of 73.3%. The AI model's performance on this same comparison subset showed an accuracy of 88.9%, sensitivity of 94%, and specificity of 60%, similar to its performance on the whole testing set.

Conclusion: Our best-performing model was able to diagnose OSSN in SLPs with high accuracy (89.9%) and sensitivity (95.1%). Compared to primary eye care providers, the AI model showed higher accuracy and sensitivity. This model could be used for screening by primary care providers, specialists, other providers, or in low resource settings to screen for OSSN.

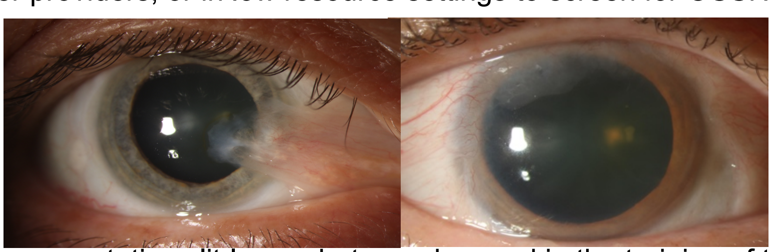

Figure 1: Two representative slit-lamp photographs used in the training of the model. A: eye with pterygium. B: eye with OSSN.