



Published in final edited form as:

*Br J Ophthalmol.* 2023 March ; 107(3): 349–354. doi:10.1136/bjophthalmol-2021-319058.

## Anterior Segment Biometric Measurements Explain Misclassifications by a Deep Learning Classifier for Detecting Gonioscopic Angle Closure

Alice Shen, MD<sup>1</sup>, Michael Chiang, PhD<sup>1</sup>, Anmol A. Pardeshi, MS<sup>1</sup>, Roberta McKean-Cowdin, PhD<sup>1,2</sup>, Rohit Varma, MD, MPH<sup>3</sup>, Benjamin Y. Xu, MD, PhD<sup>1</sup>

<sup>1</sup>Roski Eye Institute, Department of Ophthalmology, Keck School of Medicine at the University of Southern California, Los Angeles, California

<sup>2</sup>Department of Preventive Medicine, Keck School of Medicine at the University of Southern California, Los Angeles, California

<sup>3</sup>Southern California Eye Institute, CHA Hollywood Presbyterian Medical Center, Los Angeles, California

### Abstract

**Background/Aims:** To identify biometric parameters that explain misclassifications by a deep learning classifier for detecting gonioscopic angle closure in anterior segment OCT (AS-OCT) images.

**Methods:** Chinese American Eye Study (CHES) participants underwent gonioscopy and AS-OCT of each angle quadrant. A subset of CHES AS-OCT images were analyzed using a deep learning classifier to detect positive angle closure based on manual gonioscopy by a reference human examiner. Parameter measurements were compared between four prediction classes: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Logistic regression models were developed to differentiate between true and false predictions. Performance was assessed using area under the receiver operating curve (AUC) and classifier accuracy metrics.

**Results:** 584 images from 127 participants were analyzed, yielding 271 TP, 224 TN, 77 FP, 12 FN. Parameter measurements differed ( $p < 0.001$ ) between prediction classes among anterior segment parameters, including iris curvature (IC) and lens vault (LV), and angle parameters, including angle opening distance (AOD). FP resembled TP more than FN and TN in terms of anterior segment parameters (steeper IC and higher LV), but resembled TN more than TP and FN in terms of angle parameters (wider AOD). Models for detecting FP (AUC=0.752) and FN (AUC=0.838) improved classifier accuracy from 84.8 to 89.0%.

---

**Corresponding Author:** Benjamin Xu, Department of Ophthalmology, Keck School of Medicine at the University of Southern California, 1450 San Pablo Street, 4th Floor, Suite 4700, Los Angeles, CA 90033, Phone number: 323-442-6780; Fax number: 323-442-6412, benjamin.xu@med.usc.edu.

Contributorship

AS, RMC, RV, and BYX conceived and designed this study. AS, MC, AAP, RMC, RV and BYX contributed to the drafting of manuscript. AS, MC, AAP and BYX contributed to the data analysis and interpretation. All authors contributed to the critical appraisal and final approval of the manuscript. BYX provided the overall supervision of this work.

**Competing Interests:** None

**Conclusions:** Misclassifications by an OCT-based deep learning classifier for detecting gonioscopic angle closure are explained by disagreement between anterior segment and angle parameters. This finding could be used to improve classifier performance and highlights differences between gonioscopic and AS-OCT definitions of angle closure.

## PRÉCIS

Biometric measurements help explain misclassifications by an OCT-based deep learning classifier for detecting gonioscopic angle closure and can be used to develop statistical models that adjust classifier predictions and improve overall performance.

---

## INTRODUCTION

Primary angle closure glaucoma (PACG) is the most severe form of primary angle closure disease (PACD), a spectrum of disease that is characterized by apposition between the iris and trabecular meshwork (TM).<sup>1</sup> Although PACG comprises only a quarter of cases of primary glaucoma worldwide, it is responsible for approximately half of all cases of glaucoma-related blindness.<sup>2</sup> The risk of progression from early PACD to PACG and vision loss can be decreased by treatments, such as laser peripheral iridotomy (LPI) and lens extraction; however, PACD must first be detected.<sup>3–5</sup> The current clinical standard for diagnosing PACD is identifying angle closure, defined as inability to visualize the pigmented TM, on gonioscopy. However, manual gonioscopy is subjective and dependent on examiner expertise, leading to only fair intra- and inter-observer reproducibility even among experienced examiners.<sup>6,7</sup>

In recent years, anterior segment optical coherence tomography (AS-OCT) has played an increasing role in clinical evaluations of the anterior chamber angle. AS-OCT is a non-contact, in vivo imaging method that acquires cross-sectional images of anterior segment anatomical structures by measuring their optical reflections.<sup>8</sup> A previously described deep learning classifier automates analysis of AS-OCT images to detect eyes with gonioscopic angle closure and PACD.<sup>9</sup> This method of “automated gonioscopy” could serve as a clinical screening tool to detect patients at risk for PACG. However, factors that reduce the accuracy of the classifier, and potentially that of other deep learning algorithms for detecting angle closure, have not been studied.<sup>10</sup>

In order to optimize the accuracy and clinical utility of this deep learning classifier, there is a need to understand why the angles in some AS-OCT images are misclassified as open or closed relative to reference diagnoses by a human examiner performing manual gonioscopy. Quantitative analysis of AS-OCT images can be performed using highly reproducible measurements of biometric parameters that represent the dimensions and configurations of anterior segment anatomical structures.<sup>11,12</sup> In this study, we perform quantitative analysis of AS-OCT measurements to identify biometric differences between correctly and incorrectly classified images. We also develop statistical models using biometric measurements to assess if misclassifications can be reclassified to improve the accuracy of the original deep learning classifier for detecting gonioscopic angle closure.<sup>9</sup>

## METHODS

Participants were recruited from the Chinese American Eye Study (CHES), a population-based, cross-sectional study that included 4,572 Chinese participants aged 50 years and older residing in the city of Monterey Park, California.<sup>13</sup> Ethics committee approval was previously obtained from the University of Southern California Medical Center Institutional Review Board. All study procedures adhered to the recommendations of the Declaration of Helsinki. All study participants provided informed consent.

Inclusion criteria for the study included CHES participants whose gonioscopy grades and AS-OCT images were previously used in the test dataset to validate a deep learning classifier to detect gonioscopic angle closure based on AS-OCT images.<sup>14</sup> Exclusion criteria included images that were missing AS-OCT measurements for one or more parameters.

### Clinical Assessment

Each CHES participant received a complete eye examination by a trained ophthalmologist including gonioscopy and AS-OCT imaging.<sup>13</sup> Gonioscopy was performed in the seated position under dark ambient lighting ( $0.1 \text{ cd/m}^2$ ) with a 1-mm light beam and a Posner-type 4-mirror lens (Model ODPSG; Ocular Instruments, Inc., Bellevue, WA, USA) by one of two trained ophthalmologists (D.W., C.L.G.) masked to other examination findings. One ophthalmologist (D.W.) performed the majority (over 90%) of gonioscopic examinations. Care was taken to avoid light falling on the pupil and inadvertent indentation of the globe. The gonioscopy lens could be tilted up to 10 degrees. The angle was graded in each quadrant according to the modified Shaffer classification system: grade 0, no structures visible; grade 1, non-pigmented TM visible; grade 2; pigmented TM visible; grade 3, scleral spur visible; grade 4, ciliary body visible. Average gonioscopy score was calculated by averaging the numerical values of the Shaffer grades. Gonioscopy grades were dichotomized into two categories: gonioscopic angle closure (grade 0 or 1), in which pigmented TM could not be visualized, and gonioscopic open angle (grades 2 to 4).

### AS-OCT Imaging and Data Analysis

AS-OCT imaging in CHES was performed in the seated position under dark ambient lighting ( $0.1 \text{ cd/m}^2$ ) after gonioscopy and prior to pupillary dilation by a single trained ophthalmologist (D.W.) with the Tomey CASIA SS-1000 swept-source Fourier-domain device (Tomey Corporation, Nagoya, Japan). 128 two-dimensional cross-sectional AS-OCT images were acquired per eye. During imaging, the eyelids were gently retracted taking care to avoid inadvertent pressure on the globe.

AS-OCT data was imported into the Tomey SS OCT Viewer software (version 3.0, Tomey Corporation, Nagoya, Japan) which automatically segmented anatomical structures and produced biometric measurements once the scleral spurs were marked. One observer (A.A.P.) masked to the identities and examination results of the participants confirmed the segmentation and marked the scleral spurs in each image.<sup>15</sup> Two images oriented along the horizontal (temporal-nasal) and vertical meridians were analyzed per eye to measure

biometric parameters describing the anterior chamber angle and anterior segment in each quadrant.

The following biometric parameters describing the anterior chamber angle were analyzed in each image: angle opening distance (AOD500/750), trabecular iris space area (TISA500/750), and scleral spur angle (SSA500/750) measured at 500  $\mu\text{m}$  and 750  $\mu\text{m}$  from the scleral spur.<sup>16</sup> The following biometric parameters describing the anterior segment and its anatomical structures were also analyzed in each image: iris area (IA), iris curvature (IC), anterior chamber depth (ACD), lens vault (LV), anterior chamber width (ACW), pupillary diameter (PD), and anterior chamber area (ACA). IA was defined as the cross-sectional area of the full length of the iris. ACD was defined as the distance from the apex of the anterior lens surface to the apex of the corneal endothelium. IC was defined as the distance from the apex of the iris convexity to a line extending from the peripheral to central iris pigment epithelium. ACW was defined as the distance between scleral spurs. ACA was defined as the space bounded by the corneal endothelium, anterior iris surface, and anterior lens capsule. Iris thickness measured 750 and 2000  $\mu\text{m}$  (IT750/2000) from the iris root were not included as biometric parameters due to large numbers of missing values.

Intra-observer repeatability of averaged parameter measurements was calculated in the form of intraclass correlation coefficients (ICCs) based on images from 20 open angle and 20 angle closure eyes graded three months apart. ICC values reflected excellent measurement repeatability for all parameters, ranging between 0.89 (TISA750) to 0.98 (ACA). This analysis was performed using MATLAB (Mathworks, Natick, MA).

### Prediction Classes

Each AS-OCT image was categorized into one of four prediction classes based on the relationship between angle status assigned by an ophthalmologist performing manual gonioscopy and angle status predicted by the previously described deep learning classifier. True positive (TP) was defined as an image with predicted angle closure in an angle quadrant with angle closure (grades 0 and 1) based on manual gonioscopy. False positive (FP), or an overcall, was defined as an image with predicted angle closure in an angle quadrant with open angle (grades 2 to 4) based on manual gonioscopy. True negative (TN) was defined as an image with predicted open angle in an angle quadrant with open angle based on manual gonioscopy. False negative (FN), or an undercall, was defined as an image with predicted open angle in an angle quadrant with angle closure based on manual gonioscopy.

### Statistical Analysis

Normality testing was performed on biometric measurements using the Kolmogorov-Smirnov test. Median biometric measurements and their interquartile ranges (IQRs) were calculated for all parameters due to non-normal distributions. Median biometric measurements were compared between prediction classes using the Kruskal-Wallis test while adjusting for age and sex. Pairwise comparisons of median biometric measurements between prediction classes were performed using the post-hoc Dunn's test, adjusted for multiple comparisons.

A multivariable logistic regression model classifying between TP and FP was developed using the best subsets variable selection method. The final model was selected using a combination of adjusted  $R^2$ , Mallows' CP, and BIC statistics. A univariable logistic regression model classifying between TN and FN was developed by selecting the biometric parameter with the highest adjusted  $R^2$  statistic. The number of independent variables in this model was limited by the small number of FN in the dataset ( $N = 12$ ). Area under the receiver operating curve (AUC) was used to assess model performance. The models were used to reclassify all positive and negative predictions as TP or FP and TN or FN, respectively, using a default probability threshold of 0.5. All analyses were performed using the R programming interface (version 4.0.2). Analyses were conducted using a significance level of 0.05.

## RESULTS

The independent test dataset used to test the original deep learning classifier was comprised of 640 AS-OCT images from 127 CHES participants. 56 images (8.8%) from 12 participants were excluded due to missing AS-OCT measurements for one or more biometric parameters, with 38 images (5.9%) missing measurements of angle width parameters due to poorly defined scleral spurs. 584 images from 115 participants with corresponding gonioscopy grades and AS-OCT measurements were included in this study. 283 images corresponded with gonioscopic angle closure and 301 corresponded with gonioscopic open angle. The mean gonioscopy grade was  $1.5 \pm 1.1$  on the modified Shaffer grading scale. The mean age of all participants was  $62.1 \pm 9.0$  years (range, 50–86 years). 75 participants (65.2%) were female, and 40 participants (34.8%) were male. The mean IOP was  $15.75 \pm 3.4$  mmHg, mean refractive error was  $-0.3 \pm 3.4$  D spherical equivalent, and mean axial length was  $23.5 \pm 1.4$  mm.

Among 584 total images, 271 images (46.4%) with gonioscopic angle closure and 224 images (38.4%) with gonioscopic open angle were correctly predicted by the deep learning classifier. 77 images (13.2%) with gonioscopic open angle were overcalled as closed (FP) by the deep learning classifier (Figure 1A–B). 12 images (2.1%) with gonioscopic angle closure were undercalled as open (FN) by the deep learning classifier (Figure 1C–D). The deep learning classifier had an overall accuracy of 84.8%, with a sensitivity of 95.8% and specificity of 74.4%.

There were significant differences ( $p < 0.001$ ) in median parameter measurements between prediction classes among all anterior chamber angle parameters (AOD, TISA, SSA). For these parameters, the order of median parameter measurements from smallest to largest was consistent: TP, FN, FP, TN (Table 1). Significance of pairwise comparisons varied, but there was a consistent significant difference ( $p < 0.008$ ) between TP and FN (Table 1).

There were significant differences ( $p < 0.036$ ) in median parameter measurements between prediction classes among all anterior segment parameters except PD ( $p = 0.88$ ). For some anterior segment parameters (IA, ACD, ACW, ACA), the order of median parameter measurements from smallest to largest was consistent, but differed from the angle parameters: TP, FP, FN, TN. This order was the same as for other anterior segment

parameters (IC, LV) when median parameter measurements were ordered from largest to smallest. Significance of pairwise comparisons varied, but there was a consistent significant difference ( $p < 0.008$ ) between TP and FN (Table 1).

A univariable logistic regression model with AOD500 produced the best performance in discriminating between TN and FN (AUC = 0.837, Figure 2A). There was a significant association ( $p < 0.001$ ) between these prediction classes and AOD500.

A multivariable logistic regression model with AOD500, ACD, LV, and IA produced the best performance in discriminating between TP and FP (AUC = 0.752, Figure 2B). The association with prediction classes was significant ( $p < 0.03$ ) for AOD500, ACD, and IA and borderline significant ( $p = 0.07$ ) for LV.

Positive and negative predictions were re-classified as TP or FP and TN or FN, respectively, using the logistic regression models and a probability threshold of 0.5. Among the 348 images predicted to closed by the deep learning classifier, 264 were correctly identified as TP, 25 were correctly identified as FP, 52 were incorrectly identified as TP (remaining as FP), and 7 were incorrectly identified as FP (becoming FN) by the logistic regression model. Among the 236 images predicted as open by the deep learning classifier, all were identified as TN by the logistic regression model, which preserved the 224 TN and 12 FN by the deep learning classifier. Factoring in the revised predictions, the classifier and logistic regression models produced 296 TP, 45 FP, 19 FN, and 224 TN, which corresponded with an overall accuracy of 89.0%, sensitivity of 94.0%, and specificity of 83.3% (Figure 3).

## Discussion

Deep learning algorithms provide powerful tools for analyzing ocular images and detecting a wide range of ophthalmologic disease, including diabetic retinopathy, age-related macular degeneration, and glaucomatous optic neuropathy.<sup>17–19</sup> However, these algorithms also often function as black boxes, providing little information about why images are correctly or incorrectly classified. This study takes advantage of established methods for quantitative analysis of AS-OCT images to obtain insights into how specific biometric parameters contribute to diagnostic disagreement between manual gonioscopy and a deep learning classifier for detecting gonioscopic angle closure in AS-OCT images. We identified consistent differences and patterns when comparing biometric measurements across positive and negative prediction classes. We then used these differences and patterns to develop statistical models that improved the accuracy of the deep learning classifier. We believe this work advances our understanding of disagreements between gonioscopic and AS-OCT assessments of the anterior chamber angle and enhances the utility of deep learning classifiers for detecting eyes at risk for PACG.

Angle closure is detected on manual gonioscopy when an examiner cannot visualize the pigmented TM. This is more likely to occur in eyes with steep IC and higher LV, which have been identified as primary causes of diagnostic disagreement between angle closure detected on gonioscopy but not AS-OCT.<sup>20</sup> Our quantitative analyses demonstrated that AS-OCT measurements from FP (overall) images were more similar to TP images than



FN and TN images in terms of anterior segment parameters, including steeper IC, higher LV, shallower ACD, and narrower ACW. However, FP images were more similar to TN than TP and FN images among angle parameters, including wider AOD750. These findings support the relative importance of iris convexity and lens position over angle width for predicting gonioscopic angle closure by the classifier. In addition, these overcalls are consistent with anatomical configurations typically seen in eyes that require examiners to perform dynamic gonioscopy techniques, such as tilting of the goniolens, to assess the angle over the convexity of the lens and iris. These configurations represent a small minority of open angles in the original training dataset used to develop the classifier, which could explain why the classifier does not appear to simulate lens tilting.

Our quantitative analyses demonstrated that FN (undercall) images tended to more closely resemble TN images than TP and FP images in terms of anterior segment parameters, including shallower IC, deeper ACD, wider ACW, and lower LV. However, FN were more similar to TP than TN and FP images in terms of angle parameters, including narrower AOD750. These incorrect predictions of gonioscopic open angle are consistent with anatomical variants of angle closure, such as plateau iris, that have deeper anterior chambers and flatter iris configurations more commonly found in eyes with open angles. These misclassifications also suggest that the classifier does not function by detecting the pigmented trabecular meshwork, which is not discernable in most AS-OCT images. Finally, decreased angle width and iridotrabecular contact (ITC) detected on AS-OCT are more commonly missed on gonioscopy in eyes with deeper ACD, which again supports the relative importance of anterior segment over angle parameters in detecting angle closure on gonioscopy.<sup>20</sup>

Our findings highlight discrepancies between gonioscopy and AS-OCT in the definitions of angle closure. While angle closure is broadly defined as apposition between the iris and trabecular meshwork, angle closure on gonioscopy is specifically defined as inability to visualize the pigmented TM whereas angle closure on AS-OCT is specifically defined as ITC anterior to the scleral spur. Our classifier, developed to simulate a human examiner performing gonioscopy using AS-OCT images, appears to make predictions by prioritizing anterior chamber dimensions and iris configurations that impair visualization of the pigmented TM over direct measurements of angle width. This observation is consistent with previous findings that gonioscopy grades are only weakly associated with AS-OCT measurements of angle width in eyes with gonioscopic angle closure.<sup>21,22</sup> While outside the scope of the current study, it would be beneficial to investigate if misclassifications are more common when angle and iris configurations are consistent with specific subtypes of ITC.<sup>23</sup>

We developed logistic regression models to identify anatomical factors associated with misclassifications and assess if these factors could be used to re-classify predictions and improve overall classifier accuracy. Models differentiating between TP and FP or TN and FN were both based on measurements of angle width (AOD500), although the model differentiating between TP and FP also included ACD, LV, and IA. These models improved classifier accuracy from 84.8% to 89.0% and specificity from 74.4% to 83.3% with a small decrease in sensitivity from 95.8% to 94.0%. We speculate that performing this secondary quantitative analysis may simulate dynamic tilting of the goniolens in eyes with apparent

angle closure and prominent iris convexity. This approach could enhance the viability of the deep learning classifier as a diagnostic tool by substantially decreasing the number of FP while only modestly increasing the number of FN. However, significant time and effort must be expended to perform quantitative analysis of AS-OCT images. Therefore, there is a need for fully automated methods to obtain biometric measurements from AS-OCT images before secondary quantitative analysis to optimize deep learning classifier performance is practical.<sup>14</sup>

Although gonioscopy is the current clinical standard for evaluating the anterior chamber angle and is used as the reference standard for training and evaluating the performance of the deep learning classifier, it has its own limitations. First, gonioscopy alone is poorly predictive of which primary angle closure suspects (PACS) will progress to PAC or PACG.<sup>5</sup> Therefore, devising a strategy to detect all patients with gonioscopic angle closure, no matter how convenient or accurate, will still yield a majority of patients who will not develop PACG. However, automated angle assessment methods may still be valuable as gonioscopy tends to be under-performed and undetected PACG can lead to significant ocular morbidity.<sup>24</sup> Second, AS-OCT measurements of angle width are more strongly associated with IOP in subsets of eyes with PACD than gonioscopy grades.<sup>14,16</sup> AS-OCT also better predicts development of gonioscopic angle closure after 4 years and response to interventions like LPI.<sup>25,26</sup> Therefore, there is a need for longitudinal studies to determine whether gonioscopy or AS-OCT is more predictive of clinical outcomes in patients with angle closure.

Our study has several limitations. First, there was a small sample size of FN due to the excellent sensitivity of the deep learning classifier in detecting gonioscopic angle closure. While we detected consistent differences between FN and other prediction classes, we may have missed others due to the small sample size of FN. Second, the classifier is based on gonioscopy performed primarily by one human examiner during CHES. Although gonioscopy is the current clinical standard for evaluating the anterior chamber angle, it is intrinsically limited by human subjectivity and inter-examiner differences in technique. Therefore, the generalizability of ground truth labels and classifier performance should be tested using independent AS-OCT and gonioscopy data. Third, the classifier predicts angle status based on analysis of only one AS-OCT image per quadrant, which could miss localized areas of angle closure within a quadrant. Therefore, future iterations of the classifier may benefit from analyzing multiple images per quadrant to improve its sensitivity. Finally, our logistic regression models were developed and tested using the same dataset, due to its relatively small size. This may limit the generalizability of these models, and improvements in classifier performance after prediction reclassification should be validated in a larger independent dataset.

In conclusion, biometric measurements help explain misclassifications by a deep learning classifier that detects gonioscopic angle closure in AS-OCT images. Misclassifications occur due to disagreement between biometric measurements of anterior chamber angle and anterior segment parameters, and secondary quantitative analysis of these measurements may be beneficial to re-classify predictions and improve overall performance. Future studies that confirm the generalizability of these findings and compare the clinical significance of



different definitions of angle closure are needed to define the role of automated OCT-based methods for detecting patients with angle closure.

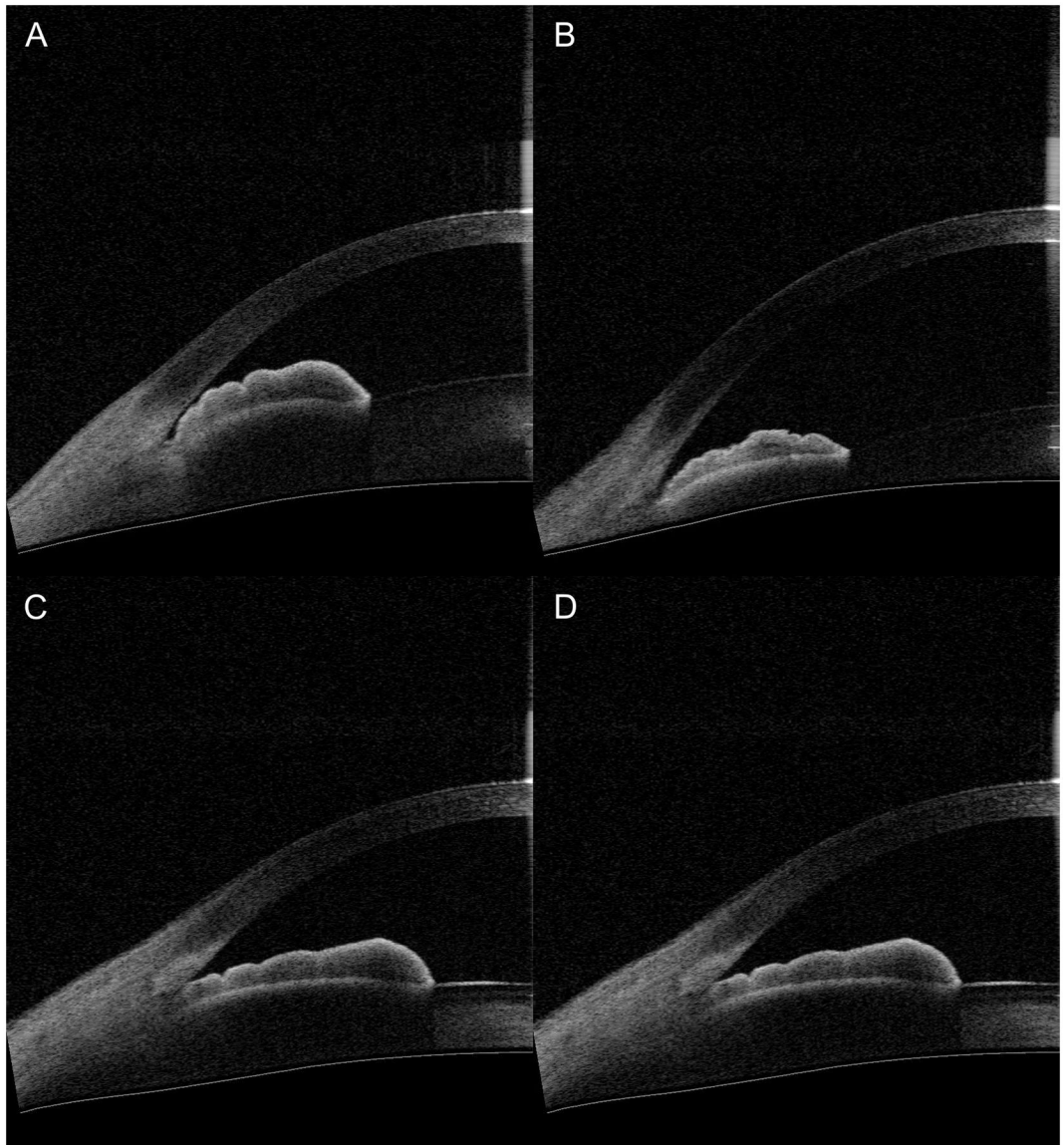
## Funding

This work was supported by grants U10 EY017337, P30 EY029220, and K23 EY029763 from the National Eye Institute, National Institute of Health, Bethesda, Maryland; a Young Clinician Scientist Research Award from the American Glaucoma Society (no grant number), San Francisco, CA; a Grant-in-Aid Research Award from Fight for Sight (no grant number), New York, NY; a SC-CTSI Clinical and Community Research Award from the Southern California Clinical and Translational Science Institute (no grant number), Los Angeles, CA; and an unrestricted grant to the Department of Ophthalmology from Research to Prevent Blindness (no grant number), New York, NY.

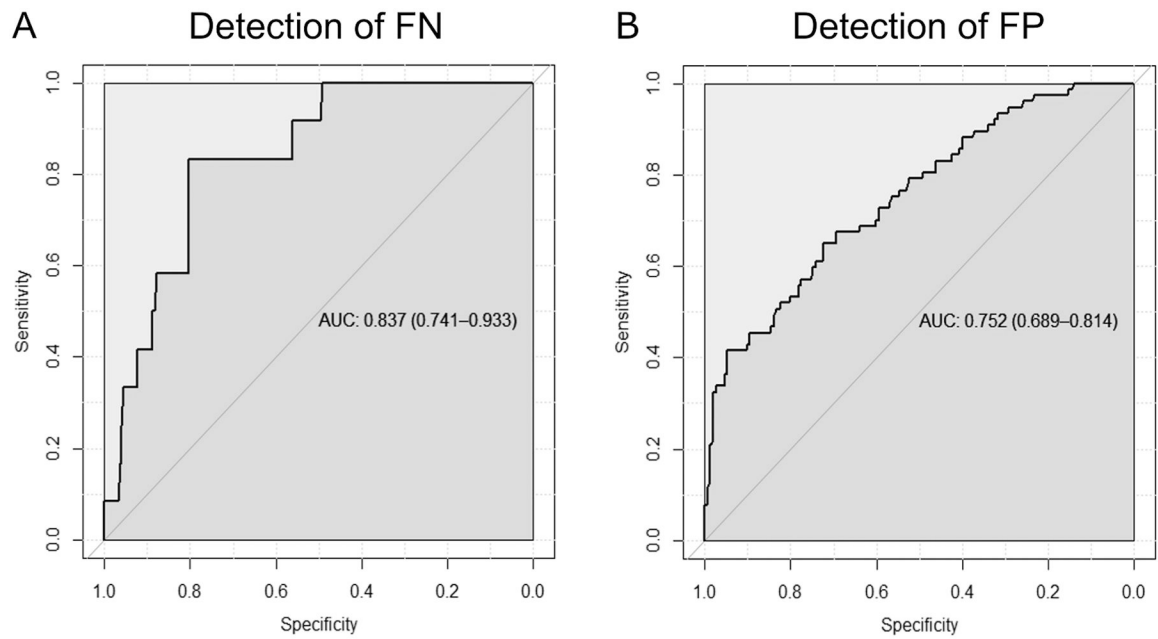
## REFERENCES

1. Wright C, Tawfik MA, Waisbourd M, Katz LJ. Primary angle-closure glaucoma: An update. *Acta Ophthalmol.* 2016;94(3):217–225. doi:10.1111/aos.12784 [PubMed: 26119516]
2. Quigley H, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. *Br J Ophthalmol.* 2006;90(3):262–267. doi:10.1136/bjo.2005.081224 [PubMed: 16488940]
3. He M, Friedman DS, Ge J, et al. Laser Peripheral Iridotomy in Primary Angle-Closure Suspects: Biometric and Gonioscopic Outcomes. The Liwan Eye Study. *Ophthalmology.* 2007;114(3):494–500. doi:10.1016/j.ophtha.2006.06.053 [PubMed: 17123610]
4. Azuara-Blanco A, Burr J, Ramsay C, et al. Effectiveness of early lens extraction for the treatment of primary angle-closure glaucoma (EAGLE): a randomised controlled trial. *Lancet.* 2016;388(10052):1389–1397. doi:10.1016/S0140-6736(16)30956-4 [PubMed: 27707497]
5. He M, Jiang Y, Huang S, et al. Laser peripheral iridotomy for the prevention of angle closure: a single-centre, randomised controlled trial. *Lancet.* 2019;393(10181):1609–1618. doi:10.1016/S0140-6736(18)32607-2 [PubMed: 30878226]
6. Rigi M, Bell NP, Lee DA, et al. Agreement between Gonioscopic Examination and Swept Source Fourier Domain Anterior Segment Optical Coherence Tomography Imaging. *J Ophthalmol.* 2016;2016:1–8. doi:10.1155/2016/1727039
7. Campbell P, Redmond T, Agarwal R, Marshall LR, Evans BJW. Repeatability and comparison of clinical techniques for anterior chamber angle assessment. *Ophthalmic Physiol Opt.* 2015;35(2):170–178. doi:10.1111/opo.12200 [PubMed: 25761580]
8. Shan J, DeBoer C, Xu BYBY. Anterior segment optical coherence tomography: Applications for clinical care and scientific research. *Asia-Pacific J Ophthalmol.* 2019;8(2):146–157. doi:10.22608/APO.201910
9. Xu BYBY, Chiang M, Chaudhary S, Kulkarni S, Pardeshi AAAA, Varma R. Deep Learning Classifiers for Automated Detection of Gonioscopic Angle Closure Based on Anterior Segment OCT Images. *Am J Ophthalmol.* 2019;208:273–280. doi:10.1016/j.ajo.2019.08.004 [PubMed: 31445003]
10. Fu H, Baskaran M, Xu Y, et al. A Deep Learning System for Automated Angle-Closure Detection in Anterior Segment Optical Coherence Tomography Images. *Am J Ophthalmol.* 2019;203(0):37–45. doi:10.1016/j.ajo.2019.02.028 [PubMed: 30849350]
11. Li H, Leung KKS, Cheung CYL, et al. Repeatability and reproducibility of anterior chamber angle measurement with anterior segment optical coherence tomography. *Br J Ophthalmol.* 2007;91(11):1490–1492. doi:10.1136/bjo.2007.118901 [PubMed: 17475709]
12. Müller M, Dahmen G, Pörksen E, et al. Anterior chamber angle measurement with optical coherence tomography: Intraobserver and interobserver variability. *J Cataract Refract Surg.* 2006;32(11):1803–1808. doi:10.1016/j.jcrs.2006.07.014. [PubMed: 17081861]
13. Varma R, Hsu C, Wang D, Torres M, Azen SP. The Chinese American eye study: Design and methods. *Ophthalmic Epidemiol.* 2013;20(6):335–347. doi:10.3109/09286586.2013.823505 [PubMed: 24044409]

14. Xu BY, Chiang M, Pardeshi AA, Moghimi S, Varma R. Deep Neural Network for Scleral Spur Detection in Anterior Segment OCT Images: The Chinese American Eye Study. *Transl Vis Sci Technol.* 2020;9(2):18. doi:10.1167/tvst.9.2.18
15. Ho SW, Baskaran M, Zheng C, et al. Swept source optical coherence tomography measurement of the iris-trabecular contact (ITC) index: A new parameter for angle closure. *Graefe's Arch Clin Exp Ophthalmol.* 2013;251(4):1205–1211. doi:10.1007/s00417-012-2158-6 [PubMed: 23001586]
16. Xu BY, Burkemper B, Lewinger JP, et al. Correlation between Intraocular Pressure and Angle Configuration Measured by OCT. *Ophthalmol Glaucoma.* 2018;1(3):158–166. doi:10.1016/j.ogla.2018.09.001 [PubMed: 31025032]
17. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA - J Am Med Assoc.* 2016;316(22):2402–2410. doi:10.1001/jama.2016.17216
18. Ting DSW, Cheung CYL, Lim G, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA - J Am Med Assoc.* 2017;318(22):2211–2223. doi:10.1001/jama.2017.18152
19. Li Z, He Y, Keel S, Meng W, Chang RT, He M. Efficacy of a Deep Learning System for Detecting Glaucomatous Optic Neuropathy Based on Color Fundus Photographs. *Ophthalmology.* 2018;125(8):1199–1206. doi:10.1016/j.ophtha.2018.01.023 [PubMed: 29506863]
20. Porporato N, Baskaran M, Tun TA, et al. Understanding diagnostic disagreement in angle closure assessment between anterior segment optical coherence tomography and gonioscopy. *Br J Ophthalmol.* 2020;104(6):795–799. doi:10.1136/bjophthalmol-2019-314672 [PubMed: 31492674]
21. Xu BY, Pardeshi AA, Burkemper B, et al. Differences in anterior chamber angle assessments between gonioscopy, eyecam, and anterior segment OCT: The Chinese American eye study. *Transl Vis Sci Technol.* 2019;8(2):5. doi:10.1167/tvst.8.2.5
22. Xu BY, Pardeshi AA, Burkemper B, et al. Quantitative Evaluation of Gonioscopic and EyeCam Assessments of Angle Dimensions Using Anterior Segment Optical Coherence Tomography. *Transl Vis Sci Technol.* 2018;7(6):33. doi:10.1167/tvst.7.6.33
23. Kong X, Foster PJ, Huang Q, et al. Appositional closure identified by ultrasound biomicroscopy in population-based primary angle-closure glaucoma suspects: The Liwan eye study. *Investig Ophthalmol Vis Sci.* 2011;52(7):3970–3975. doi:10.1167/iovs.10-6412 [PubMed: 21357394]
24. Coleman AL, Yu F, Evans SJ. Use of gonioscopy in medicare beneficiaries before glaucoma surgery. *J Glaucoma.* 2006;15(6):486–493. doi:10.1097/01.ijg.0000212287.62798.8f [PubMed: 17106360]
25. Baskaran M, Iyer JV., Narayanaswamy AK, et al. Anterior segment imaging predicts incident gonioscopic angle closure. *Ophthalmology.* 2015;122(12):2380–2384. doi:10.1016/j.ophtha.2015.07.030 [PubMed: 26359189]
26. Kwon J, Sung KR, Han S. Long-term Changes in Anterior Segment Characteristics of Eyes With Different Primary Angle-Closure Mechanisms. *Am J Ophthalmol.* 2018;191:54–63. doi:10.1016/j.ajo.2018.04.005 [PubMed: 29655644]



**Figure 1:**  
Representative misclassified anterior segment OCT images. (A-B) Representative false positive (overcall) images incorrectly classified as gonioscopic angle closure. (C-D) Representative false negative (undercall) images incorrectly classified as gonioscopic open angle.



**Figure 2:** Receiver operating characteristic (ROC) curves of (A) a univariable logistic regression model based on AOD500 for differentiating between true and false negatives (AUC = 0.837) and (B) a multivariable logistic regression model based on AOD500, ACD, LV, and IA for differentiating between true and false positives (AUC = 0.752).

**A**

Original Predictions

	Closed	Open
Closed	271	12
Open	77	224

Reference Label

Predicted Label

**B**

Adjusted Predictions

	Closed	Open
Closed	296	19
Open	45	224

Reference Label

Predicted Label

**Figure 3:** Confusion matrices of (A) original deep learning classifier predictions and (B) deep learning classifier predictions adjusted using logistic regression models.

**Table 1:**

Analysis of variance test on ranks comparing median biometric measurements of anterior chamber angle and anterior segment parameters.

Parameter	MEDIAN (INTERQUARTILE RANGE)				P-value <sup>a</sup>	Pairwise comparisons <sup>b</sup>
	TP N = 271	FP N = 77	FN N = 12	TN N = 224		
<b>Anterior Chamber Angle</b>						
AOD500	0.14 (0.09)	0.19 (0.13)	0.17 (0.05)	0.28 (0.12)	<b>&lt;0.001</b>	TP<FN<FP<TN
AOD750	0.18 (0.11)	0.26 (0.19)	0.21 (0.15)	0.41 (0.21)	<b>&lt;0.001</b>	TP<FN<FP<TN
TISA500	0.07 (0.05)	0.09 (0.04)	0.09 (0.09)	0.12 (0.06)	<b>&lt;0.001</b>	TP<FN=FP<TN
TISA750	0.11 (0.07)	0.15 (0.08)	0.14 (0.10)	0.21 (0.10)	<b>&lt;0.001</b>	TP<FN=FP<TN
SSA500	15.23 (9.01)	19.47 (11.62)	15.66 (7.72)	28.29 (10.94)	<b>&lt;0.001</b>	TP=FN<FP<TN
SSA750	13.39 (8.50)	18.46 (13.00)	14.01 (10.82)	26.48 (12.00)	<b>&lt;0.001</b>	TP=FN<FP<TN
<b>Anterior Segment</b>						
IA	1.48 (0.31)	1.49 (0.29)	1.53 (0.21)	1.53 (0.28)	<b>0.036</b>	TP=FP<FN=TN
IC	0.30 (0.12)	0.25 (0.13)	0.25 (0.13)	0.20 (0.14)	<b>&lt;0.001</b>	TP>FP=FN>TN
ACD	2.14 (0.30)	2.33 (0.35)	2.47 (0.37)	2.69 (0.41)	<b>&lt;0.001</b>	TP<FP=FN<TN
LV	0.77 (0.31)	0.67 (0.30)	0.53 (0.40)	0.41 (0.29)	<b>&lt;0.001</b>	TP>FP=FN>TN
ACW	11.36 (0.48)	11.39 (0.55)	11.45 (0.42)	11.68 (0.63)	<b>&lt;0.001</b>	TP=FP=FN<TN
PD	3.79 (1.06)	3.68 (1.14)	3.83 (1.20)	3.68 (1.03)	0.880	TP=FP=FN=TN
ACA	14.5 (2.94)	16.66 (3.02)	17.55 (2.27)	19.96 (4.80)	<b>&lt;0.001</b>	TP<FP=FN<TN

**Abbreviations:** AOD500/750: Angle opening distance 500/750 um from the scleral spur. TISA500/750: Trabecular-iris space area 500/750 um from the scleral spur. IT750/2000: Iris Thickness 750/2000 um from the scleral spur. IA: Iris Area. IC: Iris Curvature. ACD: Anterior Chamber Depth. ACW: Anterior Chamber Width. ACA: Anterior Chamber Area. LV: Lens Vault. PD: Pupillary Diameter.

<sup>a</sup>P-values calculated using Kruskal-Wallis test.

<sup>b</sup>Pairwise comparisons performed using Dunn's test adjusting for multiple comparisons. Boldface indicates significant at P < 0.05.