JAMA Ophthalmology | Original Investigation

A Competition for the Diagnosis of Myopic Maculopathy by Artificial Intelligence Algorithms

Bo Qian, MS; Bin Sheng, PhD; Hao Chen, PhD; Xiangning Wang, MD; Tingyao Li, BS; Yixiao Jin, MMed; Zhouyu Guan, MBBS; Zehua Jiang, MMed; Yilan Wu, BS; Jinyuan Wang, MD; Tingli Chen, MD; Zhengrui Guo, BS; Xiang Chen, BS; Dawei Yang, PhD; Junlin Hou, PhD; Rui Feng, PhD; Fan Xiao, MS; Yihao Li, PhD; Mostafa El Habib Daho, PhD; Li Lu, MS; Ye Ding, PhD; Di Liu, PhD; Bo Yang, PhD; Wenhui Zhu, MS; Yalin Wang, PhD; Hyeonmin Kim, BS; Hyeonseob Nam, MS; Huayu Li, MS; Wei-Chi Wu, MD, PhD; Qiang Wu, MD, PhD; Rongping Dai, MD; Huating Li, PhD; Marcus Ang, PhD; Daniel Shu Wei Ting, MD, PhD; Carol Y. Cheung, PhD; Xiaofei Wang, PhD; Ching-Yu Cheng, MD, PhD; Gavin Siew Wei Tan, MD, PhD; Kyoko Ohno-Matsui, MD, PhD; Jost B. Jonas, MD; Yingfeng Zheng, PhD; Yih-Chung Tham, PhD; Tien Yin Wong, MD, PhD; Ya Xing Wang, MD

IMPORTANCE Myopic maculopathy (MM) is a major cause of vision impairment globally. Artificial intelligence (AI) and deep learning (DL) algorithms for detecting MM from fundus images could potentially improve diagnosis and assist screening in a variety of health care settings.

OBJECTIVES To evaluate DL algorithms for MM classification and segmentation and compare their performance with that of ophthalmologists.

DESIGN, SETTING, AND PARTICIPANTS The Myopic Maculopathy Analysis Challenge (MMAC) was an international competition to develop automated solutions for 3 tasks: (1) MM classification, (2) segmentation of MM plus lesions, and (3) spherical equivalent (SE) prediction. Participants were provided 3 subdatasets containing 2306, 294, and 2003 fundus images, respectively, with which to build algorithms. A group of 5 ophthalmologists evaluated the same test sets for tasks 1 and 2 to ascertain performance. Results from model ensembles, which combined outcomes from multiple algorithms submitted by MMAC participants, were compared with each individual submitted algorithm. This study was conducted from March 1, 2023, to March 30, 2024, and data were analyzed from January 15, 2024, to March 30, 2024.

EXPOSURE DL algorithms submitted as part of the MMAC competition or ophthalmologist interpretation.

MAIN OUTCOMES AND MEASURES MM classification was evaluated by quadratic-weighted κ (QWK), F1 score, sensitivity, and specificity. MM plus lesions segmentation was evaluated by dice similarity coefficient (DSC), and SE prediction was evaluated by R^2 and mean absolute error (MAE).

RESULTS The 3 tasks were completed by 7, 4, and 4 teams, respectively. MM classification algorithms achieved a QWK range of 0.866 to 0.901, an F1 score range of 0.675 to 0.781, a sensitivity range of 0.667 to 0.778, and a specificity range of 0.931 to 0.945. MM plus lesions segmentation algorithms achieved a DSC range of 0.664 to 0.687 for lacquer cracks (LC), 0.579 to 0.673 for choroidal neovascularization, and 0.768 to 0.841 for Fuchs spot (FS). SE prediction algorithms achieved an R^2 range of 0.791 to 0.874 and an MAE range of 0.708 to 0.943. Model ensemble results achieved the best performance compared to each submitted algorithms, and the model ensemble outperformed ophthalmologists at MM classification in sensitivity (0.801; 95% CI, 0.764-0.840 vs 0.727; 95% CI, 0.684-0.768; $P = .006$) and specificity (0.946; 95% CI, 0.939-0.954 vs 0.933; 95% CI, 0.925-0.941; $P = .009$), LC segmentation (DSC, 0.698; 95% CI, 0.649-0.745 vs DSC, 0.570; 95% CI, 0.515-0.625; $P < .001$), and FS segmentation (DSC, 0.863; 95% CI, 0.831-0.888 vs DSC, 0.790; 95% CI, 0.742-0.830; $P < .001$).

CONCLUSIONS AND RELEVANCE In this diagnostic study, 15 AI models for MM classification and segmentation on a public dataset made available for the MMAC competition were validated and evaluated, with some models achieving better diagnostic performance than ophthalmologists.

JAMA Ophthalmol. doi:10.1001/jamaophthalmol.2024.3707
Published online September 26, 2024.

Invited Commentary
Supplemental content

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Authors: Tien Yin Wong, MD, PhD, School of Clinical Medicine, Tsinghua Medicine, Tsinghua University, 30 Shuangqing Rd, Haidian District, Beijing 100084, China (wongtienyin@tsinghua.edu.cn); Ya Xing Wang, MD, Beijing Institute of Ophthalmology, No. 17 Hougou Alley, Dongcheng District, Beijing 100005, China (yaxingw@gmail.com).

Myopia is one of the most common causes for irreversible vision impairment and blindness worldwide,[1] with estimates projecting that nearly 5 billion individuals will be affected by myopia and 1 billion individuals will be affected by high myopia by 2050.[2,3] Patients with high myopia face a heightened risk of sight-threatening complications, particularly myopic maculopathy (MM).[4-6] Accurate identification of MM grades is a major clinical and public health need, as it enables the timely screening, identification, and close monitoring of individuals with high myopia and MM risk, addressing issues of underdiagnosis and misclassification.[7-11] Moreover, precise MM grading by ophthalmologists facilitates targeted health guidance, which aims to prevent or manage concurrent complications effectively and mitigate the risk of blindness caused by MM.

Color fundus photography (CFP) serves as a primary diagnostic tool for MM, offering rapid, accurate identification and a cost-effective solution for initial diagnosis and large-scale screening of MM[12] across Asia,[13-15] Europe,[16,17] and other regions.[18,19] Given the imperative for timely MM detection, the integration of artificial intelligence (AI), especially deep learning (DL), into medical images analysis has emerged as a promising avenue to enhance screening efficiency and accuracy.[20] Recent DL applications have shown potential in diagnosing and screening multiple ophthalmic diseases, including retinopathy of prematurity[21-23] and diabetic retinopathy.[24-30] Several DL algorithms have been developed based on CFP for high myopia and its complications,[31-35] and these algorithms may assist accurate and time-efficient screening and diagnoses of MM in community and primary care settings among individuals with high myopia.[36,37]

However, major challenges remain in developing robust DL systems for MM diagnosis, including the scarcity of annotated datasets and limited access to preexisting algorithms for comparative evaluation.[38,39] There are more than 120 publicly available CFP-based datasets worldwide, primarily featuring diabetic retinopathy, glaucoma, and age-related macular degeneration. However, only 5 datasets include myopia data, and none are specifically dedicated to MM.[40,41] Organizing competitions and releasing public datasets is an effective way to attract more attention and research to specific ophthalmic diseases. For instance, the volume of competitions focused on diabetic retinopathy[42-44] has led to the development of numerous diagnostic algorithms using these datasets. In contrast, the absence of competitions and public datasets for MM has resulted in less attention and research dedicated to this condition, impeding the development and evaluation of DL algorithms.[45,46]

To address these challenges, we present a public dataset for MM diagnosis used in a competition named the Myopic Maculopathy Analysis Challenge (MMAC). The dataset was annotated using the criteria proposed by the META-PM Study Group.[12] Three clinical tasks with corresponding image labels were provided, including MM classification, MM plus lesions segmentation, and myopic spherical equivalent (SE) prediction.[47] The MMAC datasets, algorithms, codes, and model weights have been made available. Instructions for downloading these resources are provided in eFigure 1 in Supplement 1. Researchers can use these resources to further develop and evaluate different DL algorithms.

Key Points

Question Can a competition using a public dataset be used to develop accurate deep learning algorithms for the diagnosis of myopic maculopathy (MM)?

Findings In this diagnostic study, the ensemble of 15 algorithms submitted in the Myopic Maculopathy Analysis Challenge (MMAC) showed better performance than the mean performance of a group of 5 ophthalmologists in sensitivity and specificity.

Meaning The dataset, deep learning algorithms, and codes submitted in the MMAC competition facilitate the development and translation of automated algorithms for diagnosing MM.

Methods

This study arose as part of the 26th International Conference on Medical Image Computing and Computer-Assisted Intervention, which was held from October 8 through 12, 2023, in Vancouver, British Columbia, Canada. Three tasks were defined in MMAC competition: (1) classification of MM, (2) segmentation of MM plus lesions, and (3) prediction of SE. For each task, a training set and a validation set were provided for competition participants to develop algorithms. The final evaluation and ranking of the algorithms was performed on an independent test set, which was not available to participants during competition. The detailed competition setup and organization are provided in eAppendix 1 in Supplement 1. Between May 25, 2023, and August 25, 2023, approximately 120 participants from more than 12 countries registered for the competition. At the end of the registration and development period, 9 teams submitted a total of 15 algorithms: 7 for task 1, 4 for task 2, and 4 for task 3. These algorithms were documented in 11 papers published in the competition proceedings.[48] A summary of the top 3 algorithms is presented in eAppendix 2 in Supplement 1. This study was approved by the ethics committee of Shanghai Sixth People's Hospital and was conducted in accordance with the Declaration of Helsinki. Informed consent for inclusion of data was waived due to the data's retrospective nature and prior deidentification of the CFP images. This study followed the Standards for Reporting of Diagnostic Accuracy (STARD) reporting guidelines.

Dataset

To enable the development of DL algorithms for MM analysis, we collected 2306, 294, and 2003 images for the 3 tasks, respectively. These images were from 2 hospitals: Shanghai Health and Medical Center and Shanghai Sixth People's Hospital in China. All the images in the MMAC dataset were acquired with the Topcon TRC-NW400 nonmydriatic retinal camera (Topcon Corporation). The annotation for MM classification followed the criteria of the META-PM Study Group,[12] which divided images into 5 categories: (1) no macular lesions, (2) tessellated fundus, (3) diffuse chorioretinal atro-

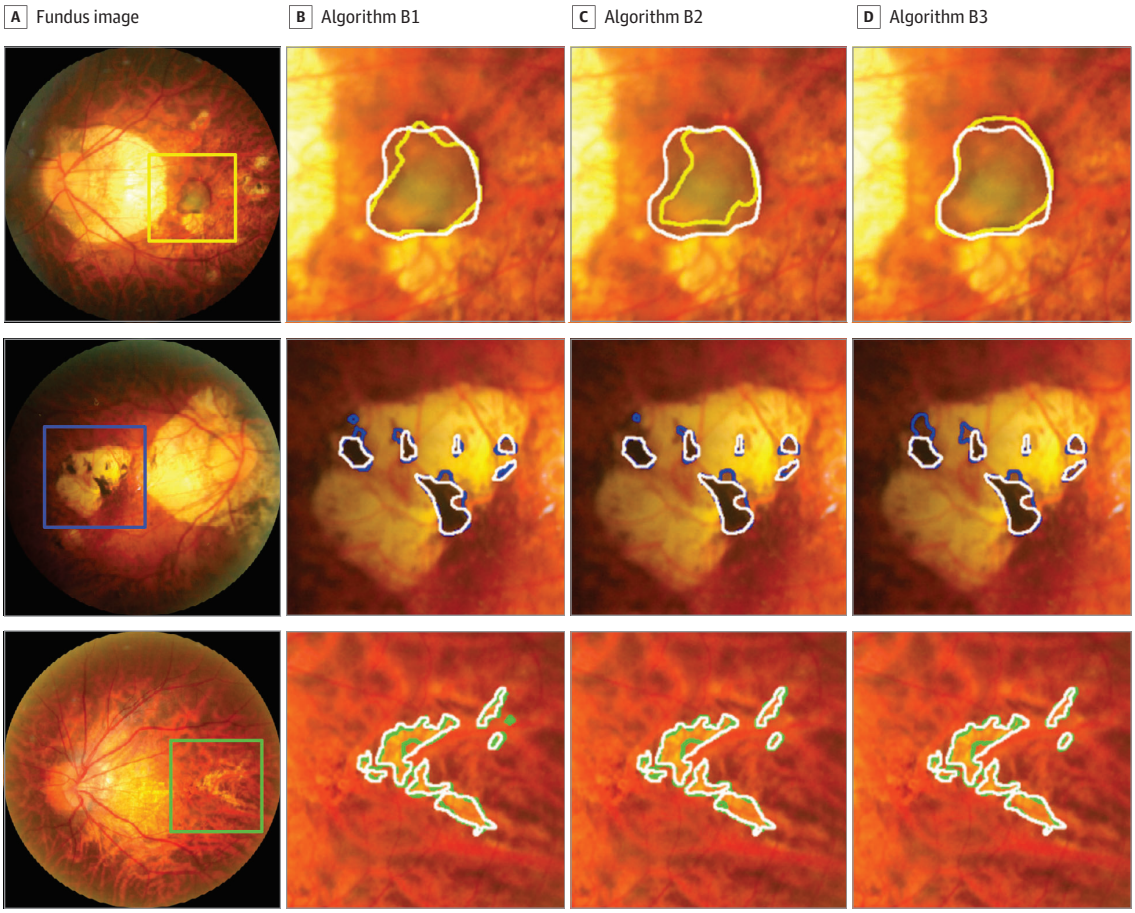
Table 1. Characteristics of the Datasets Used for the MMAC Competition

Task	No.			Patient age, mean (SD), y	Sex	
	Images	Patients	Eyes		% Male	% Female
Task 1: classification of myopic maculopathy						
Training set						
SHMC	990	429	603	52.22 (9.16)	70.4	29.6
SSPH	153	100	133	64.40 (12.38)	NA	NA
Validation set						
SHMC	215	118	160	49.90 (11.42)	70.2	29.8
SSPH	33	22	30	60.42 (12.32)	NA	NA
Test set						
SHMC	783	385	530	46.61 (11.47)	67.0	33.0
SSPH	132	90	118	60.15 (15.55)	NA	NA
Task 2: segmentation of myopic maculopathy plus lesions						
LC						
Training set						
SHMC	21	19	20	52.81 (11.24)	71.4	28.6
SSPH	42	40	42	60.50 (11.16)	NA	NA
Validation set						
SHMC	5	4	5	43.20 (8.58)	60.0	40.0
SSPH	7	7	7	57.29 (10.63)	NA	NA
Test set						
SHMC	21	19	19	53.48 (7.00)	71.4	28.6
SSPH	25	24	24	61.04 (12.63)	NA	NA
CNV						
Training set						
SHMC	5	5	5	48.20 (8.04)	60.0	40.0
SSPH	27	26	27	62.26 (15.58)	NA	NA
Validation set						
SHMC	2	2	2	38.00 (19.80)	50.0	50.0
SSPH	5	5	5	71.20 (9.50)	NA	NA
Test set						
SHMC	4	4	4	50.50 (9.18)	100	0
SSPH	18	17	18	62.67 (20.41)	NA	NA
FS						
Training set						
SHMC	21	20	21	57.00 (9.92)	61.9	38.1
SSPH	33	31	33	61.82 (13.11)	NA	NA
Validation set						
SHMC	6	6	6	57.00 (9.88)	83.3	16.7
SSPH	7	7	7	56.86 (6.87)	NA	NA
Test set						
SHMC	20	18	19	50.75 (7.20)	70.0	30.0
SSPH	25	24	25	63.88 (9.77)	NA	NA
Task 3: prediction of spherical equivalent						
Training set						
SHMC	992	560	992	53.03 (9.05)	49.5	50.5
Validation set						
SHMC	205	118	205	52.37 (8.57)	49.3	50.7
Test set						
SHMC	806	448	806	51.99 (8.91)	49.4	50.6

Abbreviations: CNV, choroidal neovascularization; FS, Fuchs spot; LC, lacquer cracks; MMAC, Myopic Maculopathy Analysis Challenge; NA, not available; SHMC, Shanghai Health and Medical Center; SSPH, Shanghai Sixth People's Hospital.

phy, (4) patchy chorioretinal atrophy, and (5) macular atrophy. Two junior ophthalmologists, each with more than 5 years of experience, independently graded the images, achieving a κ agreement value of 0.910. A senior ophthalmologist with more than 10 years of experience adjudicated any discrepancies between the 2 junior graders. The annotation for MM plus

Figure 1. Visualization of Segmentation Results From the Top 3 Algorithms in the Myopic Maculopathy Analysis Challenge (MMAC) Competition



Color fundus images (A) and segmentation results from the first-ranked algorithm (B), second-ranked algorithm (C), and third-ranked algorithm (D). The 3 horizontal rows are the segmentation results for choroidal neovascularization (yellow lines), Fuchs spot (blue lines), and lacquer cracks (green lines), respectively. White lines represent the ground truth.

lesions segmentation also followed the META-PM Study Group criteria and included 3 types of lesions: (1) lacquer cracks (LC), (2) choroidal neovascularization (CNV), and (3) Fuchs spots (FS). A junior ophthalmologist initially annotated the lesions and then refined them in consultation with another junior ophthalmologist. A senior ophthalmologist reviewed and refined the final annotations. The ground truth for the SE prediction task was obtained with the Topcon KR-8900 corneal curvature computer refractometer (Topcon Corporation). The characteristics of the MMAC dataset are outlined in **Table 1**.

Performance of Ophthalmologists

To assess the comparative performance of the algorithms in this competition against ophthalmologists in MM assessment, a comparison experiment was designed and executed. Five ophthalmologists, each with a minimum of 5 years of grading experience, were recruited to independently annotate the images in the test sets of the MM classification task (915 images) and the MM plus lesions segmentation task (113 images). These test sets were not disclosed to the participating teams during the development of their algorithms. All participating ophthalmologists were thoroughly informed about the objectives and rationale of the study and voluntarily consented to their involvement. The Shanghai Sixth People's Hospital research ethics committee, having reviewed the study design and procedures, determined that the ophthalmologists were exempt from providing written informed consent.

Statistical Analysis

The 3 tasks in the MMAC competition had different evaluation metrics. For the MM classification task, the algorithms' performance was evaluated using quadratic-weighted κ (QWK), macro F1 score, and macro specificity, where macro averaging meant calculating the metric for each label and then finding their unweighted average.[49,50] In addition to these 3 metrics used in the competition, we also used macro sensitivity in this study to evaluate the classification performance of each algorithm. For segmentation of MM plus lesions, the algorithms' performance was measured using the dice similarity coefficient (DSC). For the prediction of SE, the performance of a regression algorithm was evaluated using the coefficient of determination (R^2) and the mean absolute error (MAE). The

ranking method in each task of the competition is shown in eAppendix 3 in Supplement 1. Calculation methods for each metric are provided in eAppendix 4 in Supplement 1.

Sensitivity and specificity were used to compare the performance of DL algorithms and ophthalmologists in the MM classification task. For the MM plus lesions segmentation task, the comparison metric was DSC score. The statistical tests were 2-sided 1-sample *t* tests for MM classification and 2-sided Mann-Whitney *U* tests for MM plus lesions segmentation. A *P* value less than .05 was considered statistically significant, and *P* values were adjusted for multiple comparisons using the Bonferroni correction. Data were analyzed using Python version 3.6 and SciPy version 1.5.4 (Python Software Foundation).

Results

Competition Results

Task 1: MM Classification

The quantitative results of 7 submitted algorithms in the MM classification task are presented in eTable 1 in Supplement 1. The confusion matrix of each algorithm is shown in eFigure 2 in Supplement 1. The first-ranked algorithm achieved a QWK of 0.901 (95% CI, 0.878-0.919), an F1 score of 0.781 (95% CI, 0.739-0.819), a sensitivity of 0.778 (95% CI, 0.736-0.818), and a specificity of 0.945 (95% CI, 0.937-0.953).

Task 2: Segmentation of MM Plus Lesions

The quantitative results for MM plus lesions segmentation are presented in eTable 2 in Supplement 1. The first-ranked algorithm achieved a DSC of 0.665 (95% CI, 0.619-0.709) for LC segmentation, 0.673 (95% CI, 0.584-0.757) for CNV segmentation, and 0.841 (95% CI, 0.808-0.868) for FS segmentation, where the DSC scores for CNV and FS were the best among all teams. The best performance for LC segmentation was achieved by the third-ranked team, with a DSC score of 0.687 (95% CI, 0.630-0.741). Visualization of the segmentation results from the top 3 teams is shown in **Figure 1**.

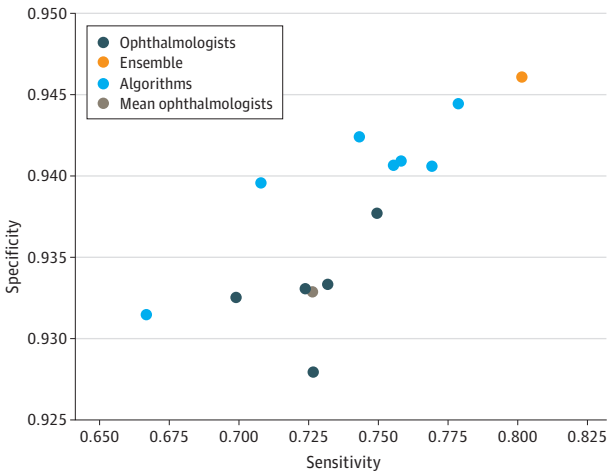
Task 3: Prediction of SE

The quantitative results of SE prediction are presented in eTable 3 in Supplement 1. The first-ranked algorithm achieved an R^2 of 0.874 (95% CI, 0.854-0.889) and an MAE of 0.708 diopters (D) (95% CI, 0.662-0.755). The distribution of the predicted vs actual values in the test set is visualized in eFigure 3 in Supplement 1. The predicted values of the 4 submitted algorithms were within 1 D of the actual values 74.9% (604 of 806), 76.9% (620), 71.6% (577), and 63.8% (514) of the time, respectively.

Combining Predictions

Ensemble learning, which integrates the outputs of multiple algorithms, is a powerful approach that combines the strengths of various models to reduce biases that might arise from a single model.[51,52] The most commonly used strategies include majority voting, logical OR, and logical AND. The calculation methods for these model ensemble strategies are described in eAppendix 5 in Supplement 1.

Figure 2. Comparison of Algorithms vs Ophthalmologists for the Myopic Maculopathy Classification Task



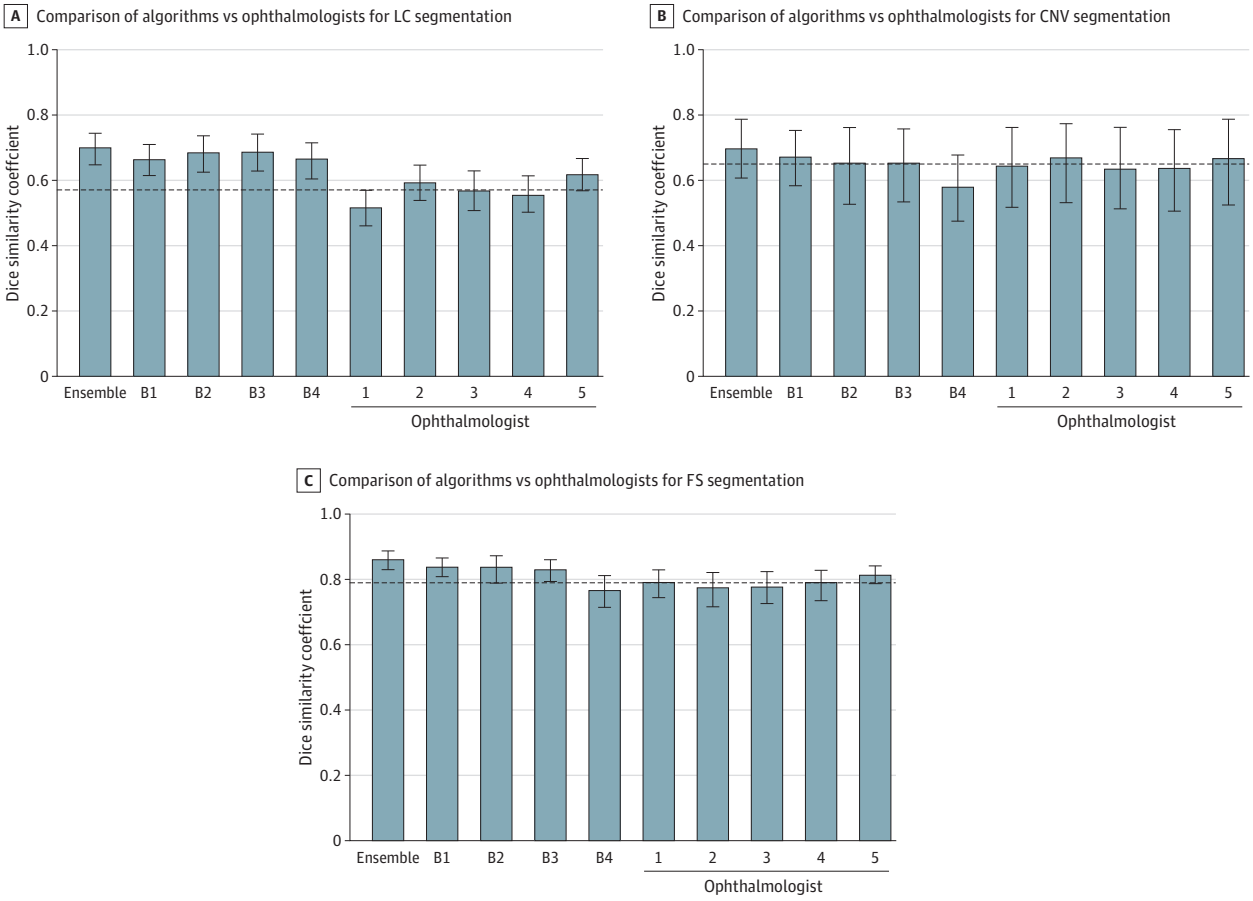
The model ensemble results achieved higher sensitivity and specificity than individual algorithms and ophthalmologists.

For the MM classification task, the model ensemble results achieved a QWK of 0.913 (95% CI, 0.897-0.927), an F1 score of 0.804 (95% CI, 0.766-0.841), a sensitivity of 0.801 (95% CI, 0.764-0.840), and a specificity of 0.946 (95% CI, 0.939-0.954), as shown in eTable 1 in Supplement 1. These values are 1.2%, 2.3%, 2.3%, and 0.1% higher, respectively, than the best submitted algorithm. For the MM plus lesions segmentation task, the model ensemble results achieved a DSC score of 0.698 (95% CI, 0.649-0.745) for LC segmentation, 0.699 (95% CI, 0.609-0.788) for CNV segmentation, and 0.863 (95% CI, 0.831-0.888) for FS segmentation, as shown in eTable 2 in Supplement 1. For the SE prediction, the model ensemble results, as shown in eTable 3 in the Supplement, achieved an R^2 of 0.883 (95% CI, 0.867-0.897) and an MAE of 0.685 (95% CI, 0.637-0.729). Compared with the first-ranked algorithm, the ensemble results showed an improvement of 0.9% for R^2 and 0.023 D for MAE.

Comparison of DL Algorithms and Ophthalmologists

The quantitative classification and segmentation results of the 5 ophthalmologists are shown in eTables 4 and 5 in Supplement 1. Visual comparison of the MM classification task for the algorithms and the ophthalmologists is shown in **Figure 2**. The significant test results comparing each algorithm with the mean performance of 5 ophthalmologists in MM classification are shown in eTable 6 in Supplement 1. For the MM classification task, the model ensemble results outperformed the ophthalmologists in sensitivity (0.801; 95% CI, 0.764-0.840 vs 0.727; 95% CI, 0.684-0.768; P = .006) and specificity (0.946; 95% CI, 0.939-0.954 vs 0.933; 95% CI, 0.925-0.941; P = .009). For the MM plus lesions segmentation task, the visual comparison of the algorithms and the ophthalmologists is shown in **Figure 3**, and the results of the significant test are shown in eTable 7 in Supplement 1. The model ensemble results outperformed the ophthalmologists in DSC scores for LC segmentation (0.698;

Figure 3. Comparison of Algorithms vs Ophthalmologists for the Myopic Maculopathy Plus Lesions Segmentation Task



Comparison of algorithms vs ophthalmologists for LC segmentation (A), CNV segmentation (B), and FS segmentation (C). The dashed line represents the average dice score of the 5 ophthalmologists. Error bars indicate 95% CIs, calculated using the bootstrap method with 1000 bootstraps. CNV indicates choroidal neovascularization; FS, Fuchs spot; LC, lacquer cracks.

95% CI, 0.649-0.745 vs 0.570; 95% CI, 0.515-0.625; $P < .001$) and FS segmentation (0.863; 95% CI, 0.831-0.888 vs 0.790; 95% CI, 0.742-0.830; $P < .001$). For the CNV segmentation task, the model ensemble achieved performance comparable to that of the ophthalmologists. In MM classification and LC and FS segmentation, some participants' algorithms also outperformed the ophthalmologists.

Discussion

Complications of pathologic myopia, including MM and refractive errors, pose significant global health challenges. Accurate grading of MM aids in the timely screening, identification, and close monitoring of patients at high risk of myopia, which helps to address diagnostic inadequacies and provide targeted health guidance.[12,53,54] Focusing on these diseases, we devised an AI competition that integrated the tasks of MM classification, MM plus lesions segmentation, and prediction of SE as key refractive metrics. The purpose of this competition was to catalyze the advancement and refinement of state-of-the-art DL algorithms for MM analysis, with the dataset, evaluation pipeline, algorithms, and codes made publicly available. Researchers can use this dataset to develop DL algorithms, for external validation, and to compare new algorithms with existing ones.

Several studies on myopia and MM have previously been reported. Tang and colleagues[55] validated a DL algorithm for MM classification using 1395 fundus images, achieving a κ value of 0.932. They also evaluated the DL algorithm's performance in the segmentation of MM plus lesions, achieving a DSC score of 0.238 for LC segmentation, but failing to segment CNV and Fuchs spots. Zheng and colleages[56] and Sun and colleagues[57] achieved κ values of 0.836 and 0.938 for MM classification on 4642 and 714 images, respectively. In this study, the κ value for MM classification was 0.913 on 2306 fundus images, and our segmentation of MM plus lesions achieved DSC scores of more than 0.65. These results are comparable to or outperform reported results. Varadarajan and colleagues[58] trained and validated a DL algorithm for SE prediction using more than 200 000 images, achieving R^2 scores of 0.9 and 0.69 on 2 datasets. In this study, comparable performance for SE

Table 2. List and Details of the Top 3 Algorithms in 3 Tasks of the MMAC Competition

Team name	Algorithm[a]	Score[b]	Pretraining weights	Architecture	Loss function	Data augmentation	Postprocecssing
fdvts_mm	A1	0.8752	ImageNet,[59] EyePACS[60]	ResNet,[61] EfficientNet,[62] ViT,[63] MobileNet,[64] LANet[64]	CE, WCE, FL	C, F, R, CJ, GBL, GS, SP	Ensemble by averaging
DGUT_luli	A2	0.8660	ImageNet	Swin Transformer[65]	AL, CE	C, F, autoaugment	Ensemble by averaging
Taco Friday	A3	0.8638	ImageNet	MobileNetV2[66]	CE	C, F, R, CJ, CT, SP, RE, MU, CM, posterize, solarize, shear, brightness	Ensemble by averaging
fdvts_mm	B1	0.7264	ImageNet	UNet,[67] UNet++,[68] DeeplabV3+[69]	DL, FL	C, F, R, CJ, GN	Ensemble by averaging and logical OR
latim	B2	0.7248	ImageNet	MANet[70]	DL, CE	F, SSR, RBC, RG, SP, BL, GD, CD, downscale	TTA
hyeonminkim0625	B3	0.7224	ImageNet	Convnext-small[71]	DL, CE	C, F, R, CJ, BL	TTA, ensemble by averaging
latim	C1	0.8735	ImageNet	EfficientNet[72]	Smooth L1	F, SSR, RBC, RG, SP, BL, CD, downscale	Ensemble by averaging
AIFuture	C2	0.8636	DINO[73]	ResNet50[61]	MSE	F, R	TTA
Taco Friday	C3	0.8433	None	ResNet50[61]	MSE	F, R, CJ, SP, RE, MU	None

Abbreviations: AL, ArcFace loss; BL, blur; C, cropping; CD, coarse dropout; CE, cross entropy; CJ, color jittering; CM, CutMix; CT, contrast; DL, dice loss; DSC, dice similarity coefficient; F, flipping; FL, focal loss; GBL, Gaussian blur; GD, grid distortion; GN, Gaussian noise; GS, gray scale; MMAC, Myopic Maculopathy Analysis Challenge; MSE, mean square error; MU, MixUp; R, rotation; RBC, random brightness contrast; RE, random erasing; RG, random gamma; SP, sharpen; SSR, shift scale rotate; TTA, test-time augmentation; WCE, weighted cross entropy; WDL, weighted dice loss.

[a] A, B, and C were used to represent the algorithms in tasks 1, 2, and 3, respectively. The numbers following each letter indicate the ranking.

[b] The score for task 1 is the average of the quadratic-weighted κ, sensitivity, and specificity scores. The score for task 2 is the average DSC, and the score for task 3 is the R^2.

prediction was achieved with an R^2 of 0.883, despite using a smaller dataset. Our DL algorithms achieved comparable or better performance across tasks. Previous methods were primarily trained and validated on private datasets, as researchers lacked public datasets for evaluation. Our released dataset can bridge this gap. Additionally, compared to these studies, we also explored model ensemble methods and compared the performance of ophthalmologists with DL algorithms.

The summarized characteristics of the top-performing algorithms in each task of the MMAC competition are listed in **Table 2**.[59-73] All the participants used DL algorithms as the solution in these tasks. The strategies summarized from participating algorithms to improve the model performance are presented in eAppendix 6 in Supplement 1. In addition, we used a model ensemble approach to integrate the outputs of different algorithms submitted by participants and demonstrate the effectiveness of the model ensemble in improving algorithm performance. In clinical practice, multiple MM classification models, each trained locally on data from different centers, can be integrated into a clinical software platform used by ophthalmologists. By using model ensemble methods and sharing locally trained models, health care centers can enhance prediction reliability and confidence, thus overcoming barriers to data sharing.

To evaluate the performance of ophthalmologists in MM classification and MM plus lesions segmentation, 5 ophthalmologists were recruited to annotate images within the test sets of these 2 tasks. The performance of the DL algorithms was compared with the ophthalmologists' mean performance. Comparison showed that in the MM classification task, the sensitivity and specificity of some algorithms were better than those of the ophthalmologists. In the MM plus lesions segmentation task, some algorithms for LC and FS segmentation outperformed ophthalmologists.

The integration of DL algorithms with cost-effective, easily accessible CFP holds promise as a powerful tool for risk stratification and the detection of MM and high myopia. Given the substantial economic burden imposed by uncorrected refractive errors and the heightened risk of vision-threatening complications, such as myopic CNV, glaucoma, retinal detachment, and macular holes, the application of these algorithms in MM and high myopia screening could represent a strategic approach to the early identification of this vulnerable population. Subsequently, individuals identified with high myopia or MM should be promptly referred to tertiary eye care facilities for meticulous examination by specialists, where more advanced diagnostic modalities like optical coherence tomography or fundus fluorescein angiography could be used to conduct a thorough evaluation and ensure timely intervention in the future. The algorithms presented in the MMAC competition are pertinent for large-scale clinical screening and epidemiological research using retrospective datasets. The development of highly accurate automated classification and refractive error precision systems can enhance research leveraging extensive retrospective datasets, particularly when dealing with preexisting datasets that lack refractive error information.

Limitations

This diagnostic study has several limitations. First, the dataset was derived exclusively from Chinese patients and was captured using a single imaging device. This homogeneity in both patient ethnicity and imaging technology may limit the generalizability to other populations and settings, where varia-

tions in genetic predispositions, environmental factors, and imaging equipment characteristics could influence the presentation and detection of MM. Second, while the algorithms emerging from this competition exhibited performance comparable to that of human experts, their practical utility in real-world clinical environments and epidemiological research settings remains to be rigorously evaluated. Third, it is essential to acknowledge that these SE prediction algorithms cannot provide the separate cylinder and sphere powers needed for prescribing spectacles. Conventional methods, such as subjective refraction using trial lenses or automated refractometry, are superior and more cost-effective. However, these algorithms can be useful in large-scale screening, opportunistic screening, and epidemiological research, particularly when dealing with retrospective datasets where refractive error information was not originally collected. Finally, the number of algorithm submissions did not meet initial expectations. Besides insufficient global promotion before the competition, the types of tasks was also a contributing factor. The MM classification task received the most submissions due to its extensive research and direct clinical significance in large-scale disease screening. In contrast, the segmentation and regression tasks had fewer submissions, as they are used less frequently in clinical practice.

Conclusions

In this diagnostic study of AI for MM classification and segmentation, algorithms submitted to the MMAC competition showed promising diagnostic performance compared with ophthalmologists, with some DL algorithms outperforming ophthalmologists. In this study, we presented an evaluation platform for assessing algorithms, a publicly available MMAC dataset with annotation, and codes for the submitted algorithms. These publicly available resources may lay the groundwork for the development of computer-assisted automatic diagnostic systems for MM.

ARTICLE INFORMATION

Accepted for Publication: July 11, 2024.

Published Online: September 26, 2024. doi:10.1001/jamaophthalmol.2024.3707

Author Affiliations: Shanghai Belt and Road International Joint Laboratory of Intelligent Prevention and Treatment for Metabolic Diseases, Department of Computer Science and Engineering, School of Electronic, Information, and Electrical Engineering, Shanghai Jiao Tong University, Department of Endocrinology and Metabolism, Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai Diabetes Institute, Shanghai Clinical Center for Diabetes, Shanghai, China (Qian, Sheng, T. Li, Guan, X. Chen, Huating Li); Ministry of Education Key Laboratory of Artificial Intelligence, School of Electronic, Information, and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China (Qian, Sheng, T. Li, X. Chen); Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Hong Kong, China (H. Chen, Guo); Department of Chemical and Biological Engineering, The Hong Kong University of Science and Technology, Hong Kong, China (H. Chen); Department of Ophthalmology, Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China (Xiangning Wang, Q. Wu); School of Clinical Medicine, Tsinghua Medicine, Tsinghua University, Beijing, China (Jin, Jiang, Y. Wu, J. Wang, Wong); School of Clinical Medicine, Beijing Tsinghua Changgung Hospital, Beijing, China (Jin, Jiang, J. Wang, Wong); Department of Ophthalmology, Shanghai Health and Medical Center, Wuxi, China (T. Chen); Department of Ophthalmology and Visual Sciences, The Chinese University of Hong Kong, Hong Kong, China (D. Yang, Cheung); Shanghai Key Laboratory of Intelligent Information Processing, School of Computer Science, Fudan University, Shanghai, China (Hou, Feng); Academy for Engineering and Technology, Fudan University, Shanghai, China (Feng, Xiao); Laboratoire de Traitement de l'Information Médicale UMR 1101, Inserm, Brest, France (Y. Li, El Habib Daho); Université de Bretagne Occidentale, Brest, France (Y. Li, El Habib Daho); School of Computer Science and Technology, Dongguan University of Technology, Dongguan, China (Lu, Ding); AIFUTURE Laboratory, Beijing, China (Liu, B. Yang); National Digital Health Center of China Top Think Tanks, Beijing Normal University, Beijing, China (Liu); School of Journalism and Communication, Beijing Normal University, Beijing, China (Liu); School of Computing and Augmented Intelligence, Arizona State University, Tempe (Zhu, Y. Wang); Mediwhale, Seoul, South Korea (Kim, Nam); Pohang University of Science and Technology, Pohang, South Korea (Kim); Department of Electrical and Computer Engineering, University of Arizona, Tucson (Huayu Li); Department of Ophthalmology, Linkou Chang Gung Memorial Hospital, College of Medicine, Chang Gung University, Taoyuan, Taiwan (W.-C. Wu); Department of Ophthalmology, Peking Union Medical College Hospital, Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing, China (Dai); Singapore Eye Research Institute, Singapore National Eye Centre, Singapore (Ang, Ting, Cheng, Tan, Jonas, Wong); Key Laboratory for Biomechanics and Mechanobiology of Ministry of Education, Beijing Advanced Innovation Center for Biomedical Engineering, School of Biological Science and Medical Engineering, Beihang University, Beijing, China (Xiaofei Wang); Department of Ophthalmology and Visual Science, Tokyo Medical and Dental University, Tokyo, Japan (Ohno-Matsui); Institut Français de Myopie, Rothschild Foundation Hospital, Paris, France (Jonas); Zhongshan Ophthalmic Center, Guangzhou, China (Zheng, Wong); Center for Innovation and Precision Eye Health, Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore (Tham); Ophthalmology and Visual Science Academic Clinical Program, Duke–National University of Singapore Medical School, Singapore (Tham); Beijing Institute of Ophthalmology, Beijing Ophthalmology and Visual Sciences Key Laboratory, Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, Beijing, China (Y. X. Wang).

Author Contributions: B. Qian and Dr Xiangning Wang had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. B. Qian, Dr Sheng, Dr H. Chen, Dr Xiangning Wang, T. Li, Y. Jin, Z. Guan, and Z. Jiang contributed equally.
Concept and design: Qian, Sheng, H. Chen, T. Li, Jin, Guan, T. Chen, Guo, Hou, Feng, Xiao, Y. Li, Lu, Liu, Zhu, Q. Wu, Ting, Jonas, Tham, Wong, Ya Xing Wang.
Acquisition, analysis, or interpretation of data: Qian, Sheng, H. Chen, Xiangning Wang, T. Li, Jiang, Y. Wu, J. Wang, X. Chen, D. Yang, Hou, Y. Li, El Habib Daho, Lu, Ding, Liu, B. Yang, Yalin Wang, Kim, Huayu Li, W. Wu, Dai, Huating Li, Ang, Cheung, Xiaofei Wang, Cheng, Tan, Ohno-Matsui, Jonas, Zheng.
Drafting of the manuscript: Qian, H. Chen, Jin, J. Wang, X. Chen, Lu, Liu, Zhu, Tan, Wong.
Critical review of the manuscript for important intellectual content: Qian, Sheng, H. Chen, Xiangning Wang, T. Li, Jin, Guan, Jiang, Y. Wu, J. Wang, T. Chen, Guo, D. Yang, Hou, Feng, Xiao, Y. Li, El Habib Daho, Lu, Ding, Liu, B. Yang, Zhu, Yalin Wang, Kim, Huayu Li, W. Wu, Q. Wu, Dai, Huating Li, Ang, Ting, Cheung, Xiaofei Wang, Cheng, Ohno-Matsui, Jonas, Zheng, Tham, Wong, Ya Xing Wang.
Statistical analysis: Qian, Jin, J. Wang, X. Chen, Hou, Y. Li, Lu, Liu, B. Yang, Zhu, Yalin Wang, Kim.
Obtained funding: Sheng, El Habib Daho, Liu, Yalin Wang.
Administrative, technical, or material support: Qian, Sheng, H. Chen, Xiangning Wang, T. Li, Guan, Jiang, Y. Wu, Guo, X. Chen, D. Yang, Feng, Xiao, B. Yang, Zhu, Yalin Wang, Q. Wu, Dai, Cheng, Zheng, Wong, Ya Xing Wang.
Supervision: Sheng, H. Chen, T. Chen, Feng, Xiao, Liu, Yalin Wang, W. Wu, Huating Li, Ang, Ting, Cheung, Xiaofei Wang, Tan, Ohno-Matsui, Wong.
Other: Jonas.

Conflict of Interest Disclosures: None reported.

Funding/Support: This study was supported by funding from the National Key Research & Development Program of China (2022YFC2502800), the National Natural Science Foundation of China (NSFC) (82388101), and the Beijing Natural Science Foundation (IS23096)

awarded to Dr T. Wong, funding from the NSFC (62272298) and Shanghai Municipal Science and Technology Major Project (2021SHZDZX0102) awarded to Dr Sheng, and funding from the College-level Project Fund of Shanghai Sixth People's Hospital (YNLC201909) and the Interdisciplinary Program of Shanghai Jiao Tong University (YG2022QN089) awarded to Dr Xiangning Wang.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Data Sharing Statement: See Supplement 2.

Additional Contributions: We thank the organizing committee of the 26th International Conference on Medical Image Computing and Computer Assisted Intervention for hosting the workshop held as part of the study reported in this article, the collaborators, and the funding agencies.

REFERENCES

1. Holden B, Sankaridurg P, Smith E, Aller T, Jong M, He M. Myopia, an underrated global challenge to vision: where the current data takes us on myopia control. *Eye*. 2014;28(2):142-146. doi:10.1038/eye.2013.256

2. Holden BA, Fricke TR, Wilson DA, et al. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. *Ophthalmology*. 2016;123(5):1036-1042. doi:10.1016/j.ophtha.2016.01.006

3. Rudnicka AR, Kapetanakis VV, Wathern AK, et al. Global variations and time trends in the prevalence of childhood myopia, a systematic review and quantitative meta-analysis: implications for aetiology and early prevention. *Br J Ophthalmol*. 2016;100(7):882-890. doi:10.1136/bjophthalmol-2015-307724

4. Ohno-Matsui K, Wu PC, Yamashiro K, et al. IMI pathologic myopia. *Investigative Ophthalmol Vis Sci*. 2021;62(5):5. doi:10.1167/iovs.62.5.5

5. Yokoi T, Ohno-Matsui K. Diagnosis and treatment of myopic maculopathy. *Asia-Pacific J Ophthalmol*. 2018;7(6):415-421. doi:10.22608/APO.2018290

6. Wong TY, Ferreira A, Hughes R, Carter G, Mitchell P. Epidemiology and disease burden of pathologic myopia and myopic choroidal neovascularization: an evidence-based systematic review. *Am J Ophthalmol*. 2014;157(1):9-25.e12. doi:10.1016/j.ajo.2013.08.010

7. He M, Xiang F, Zeng Y, et al. Effect of time spent outdoors at school on the development of myopia among children in China: a randomized clinical trial. *JAMA*. 2015;314(11):1142-1148. doi:10.1001/jama.2015.10803

8. Zhu Z, Chen Y, Tan Z, Xiong R, McGuinness MB, Müller A. Interventions recommended for myopia prevention and control among children and adolescents in China: a systematic review. *Br J Ophthalmol*. 2023;107(2):160-166. doi:10.1136/bjophthalmol-2021-319306

9. Li R, Zhang K, Li SM, et al. Implementing a digital comprehensive myopia prevention and control strategy for children and adolescents in China: a cost-effectiveness analysis. *Lancet Reg Health*. 2023;38:100837. doi:10.1016/j.lanwpc.2023.100837

10. Ma Y, Wen Y, Zhong H, et al. Healthcare utilization and economic burden of myopia in urban China: a nationwide cost-of-illness study. *J Global Health*. 2022;12:11003. doi:10.7189/jogh.12.11003

11. Ang M, Flanagan JL, Wong CW, et al. Review: myopia control strategies recommendations from the 2018 WHO/IAPB/BHVI Meeting on Myopia. *Br J Ophthalmol*. 2020;104(11):1482-1487. doi:10.1136/bjophthalmol-2019-315575

12. Ohno-Matsui K, Kawasaki R, Jonas JB, et al; META-analysis for Pathologic Myopia (META-PM) Study Group. International photographic classification and grading system for myopic maculopathy. *Am J Ophthalmol*. 2015;159(5):877-83.e7. doi:10.1016/j.ajo.2015.01.022

13. Ueda E, Yasuda M, Fujiwara K, et al. Trends in the prevalence of myopia and myopic maculopathy in a Japanese population: the Hisayama study. *Investigative Ophthalmol Vis Sci*. 2019;60(8):2781-2786. doi:10.1167/iovs.19-26580

14. Asakuma T, Yasuda M, Ninomiya T, et al. Prevalence and risk factors for myopic retinopathy in a Japanese population: the Hisayama study. *Ophthalmology*. 2012;119(9):1760-1765. doi:10.1016/j.ophtha.2012.02.034

15. Yan YN, Wang YX, Yang Y, et al. Ten-year progression of myopic maculopathy: the Beijing Eye Study 2001-2011. *Ophthalmology*. 2018;125(8):1253-1263. doi:10.1016/j.ophtha.2018.01.035

16. Haarman AEG, Tedja MS, Brussee C, et al. Prevalence of myopic macular features in Dutch individuals of European ancestry with high myopia. *JAMA Ophthalmol*. 2022;140(2):115-123. doi:10.1001/jamaophthalmol.2021.5346

17. Bikbov MM, Gilmanshin TR, Kazakbaeva GM, et al. Prevalence of myopic maculopathy among adults in a Russian population. *JAMA Netw Open*. 2020;3(3):e200567. doi:10.1001/jamanetworkopen.2020.0567

18. Shi H, Guo N, Zhao Z, He X, Li J, Duan J. Global prevalence of myopic macular degeneration in general population and patients with high myopia: a systematic review and meta-analysis. *Eur J Ophthalmol*. 2024;34(3):631-640. doi:10.1177/11206721231185816

19. Zou M, Wang S, Chen A, et al. Prevalence of myopic macular degeneration worldwide: a systematic review and meta-analysis. *Br J Ophthalmol*. 2020;104(12):1748-1754. doi:10.1136/bjophthalmol-2019-315298

20. Zhou Y, Chia MA, Wagner SK, et al; UK Biobank Eye & Vision Consortium. A foundation model for generalizable disease detection from retinal images. *Nature*. 2023;622(7981):156-163. doi:10.1038/s41586-023-06555-x

21. Coyner AS, Murickan T, Oh MA, et al. Multinational external validation of autonomous retinopathy of prematurity screening. *JAMA Ophthalmol*. 2024;142(4):327-335. doi:10.1001/jamaophthalmol.2024.0045

22. Young BK, Cole ED, Shah PK, et al. Efficacy of smartphone-based telescreening for retinopathy of prematurity with and without artificial intelligence in India. *JAMA Ophthalmol*. 2023;141(6):582-588. doi:10.1001/jamaophthalmol.2023.1466

23. Coyner AS, Oh MA, Shah PK, et al. External validation of a retinopathy of prematurity screening model using artificial intelligence in 3 low- and middle-income populations. *JAMA Ophthalmol*. 2022;140(8):791-798. doi:10.1001/jamaophthalmol.2022.2135

24. Korot E, Gonçalves MB, Huemer J, et al. Clinician-driven AI: code-free self-training on public data for diabetic retinopathy referral. *JAMA Ophthalmol*. 2023;141(11):1029-1036. doi:10.1001/jamaophthalmol.2023.4508

25. Ipp E, Liljenquist D, Bode B, et al; EyeArt Study Group. Pivotal evaluation of an artificial intelligence system for autonomous detection of referrable and vision-threatening diabetic retinopathy. *JAMA Netw Open*. 2021;4(11):e2134254. doi:10.1001/jamanetworkopen.2021.34254

26. Natarajan S, Jain A, Krishnan R, Rogye A, Sivaprasad S. Diagnostic accuracy of community-based diabetic retinopathy screening with an offline artificial intelligence system on a smartphone. *JAMA Ophthalmol*. 2019;137(10):1182-1188. doi:10.1001/jamaophthalmol.2019.2923

27. Gomez Rossi J, Rojas-Perilla N, Krois J, Schwendicke F. Cost-effectiveness of artificial intelligence as a decision-support system applied to the detection and grading of melanoma, dental caries, and diabetic retinopathy. *JAMA Netw Open*. 2022;5(3):e220269. doi:10.1001/jamanetworkopen.2022.0269

28. 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*. 2017;318(22):2211-2223. doi:10.1001/jama.2017.18152

29. 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*. 2016;316(22):2402-2410. doi:10.1001/jama.2016.17216

30. Guan Z, Li H, Liu R, et al. Artificial intelligence in diabetes management: advancements, opportunities, and challenges. *Cell Rep Med*. 2023;4(10):101213. doi:10.1016/j.xcrm.2023.101213

31. Lu L, Ren P, Tang X, et al. AI-model for identifying pathologic myopia based on deep learning algorithms of myopic maculopathy classification and "plus" lesion detection in fundus images. *Frontiers Cell Dev Biol*. 2021;9:719262. doi:10.3389/fcell.2021.719262

32. Du Y, Chen Q, Fan Y, et al. Automatic identification of myopic maculopathy related imaging features in optic disc region via machine learning methods. *J Translational Med*. 2021;19(1):167. doi:10.1186/s12967-021-02818-1

33. Du R, Xie S, Fang Y, et al. Deep learning approach for automated detection of myopic maculopathy and pathologic myopia in fundus images. *Ophthalmol Retina*. 2021;5(12):1235-1244. doi:10.1016/j.oret.2021.02.006

34. Tan TE, Anees A, Chen C, et al. Retinal photograph-based deep learning algorithms for myopia and a blockchain platform to facilitate artificial intelligence medical research: a retrospective multicohort study. *Lancet Digital Health*. 2021;3(5):e317-e329. doi:10.1016/S2589-7500(21)00055-8

35. Foo LL, Lim GYS, Lanca C, et al. Deep learning system to predict the 5-year risk of high myopia using fundus imaging in children. *NPJ Digital Med*. 2023;6(1):10. doi:10.1038/s41746-023-00752-8

36. Foo LL, Ng WY, Lim GYS, Tan TE, Ang M, Ting DSW. Artificial intelligence in myopia: current and future trends. *Curr Opinion Ophthalmol*. 2021;32(5):413-424. doi:10.1097/ICU.0000000000000791

37. Ang M, Wong CW, Hoang QV, et al. Imaging in myopia: potential biomarkers, current challenges and future developments. *Br J Ophthalmol*. 2019;103(6):855-862. doi:10.1136/bjophthalmol-2018-312866

38. Langlotz CP, Allen B, Erickson BJ, et al. A roadmap for foundational research on artificial intelligence in medical imaging: from the 2018 NIH/RSNA/ACR/The Academy Workshop. *Radiology*. 2019;291(3):781-791. doi:10.1148/radiol.2019190613

39. Liu X, Faes L, Kale AU, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *Lancet Digital Health*. 2019;1(6):e271-e297. doi:10.1016/S2589-7500(19)30123-2

40. Krzywicki T, Brona P, Zbrzezny AM, Grzybowski AE. A global review of publicly available datasets containing fundus images: characteristics, barriers to access, usability, and generalizability. *J Clin Med*. 2023;12(10):3587. doi:10.3390/jcm12103587

41. Khan SM, Liu X, Nath S, et al. A global review of publicly available datasets for ophthalmological imaging: barriers to access, usability, and generalisability. *Lancet Digital Health*. 2021;3(1):e51-e66. doi:10.1016/S2589-7500(20)30240-5

42. Niemeijer M, van Ginneken B, Cree MJ, et al. Retinopathy online challenge: automatic detection of microaneurysms in digital color fundus photographs. *IEEE Trans Med Imaging*. 2010;29(1):185-195. doi:10.1109/TMI.2009.2033909

43. Porwal P, Pachade S, Kokare M, et al. IDRiD: diabetic retinopathy - segmentation and grading challenge. *Med Image Analysis*. 2020;59:101561. doi:10.1016/j.media.2019.101561

44. Liu R, Wang X, Wu Q, et al. DeepDRiD: diabetic retinopathy-grading and image quality estimation challenge. *Patterns*. 2022;3(6):100512. doi:10.1016/j.patter.2022.100512

45. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Analysis*. 2017;42:60-88. doi:10.1016/j.media.2017.07.005

46. Prashar J, Tay N. Performance of artificial intelligence for the detection of pathological myopia from colour fundus images: a systematic review and meta-analysis. *Eye*. 2024;38(2):303-314. doi:10.1038/s41433-023-02680-z

47. Wong YL, Zhu X, Tham YC, et al; Asian Eye Epidemiology Consortium (AEEC). Prevalence and predictors of myopic macular degeneration among Asian adults: pooled analysis from the Asian Eye Epidemiology Consortium. *Br J Ophthalmol*. 2021;105(8):1140-1148. doi:10.1136/bjophthalmol-2020-316648

48. Sheng B, Chen H, Wong TY, eds. *Myopic Maculopathy Analysis*. Springer; 2024. Lecture Notes in Computer Science; vol 14563. doi:10.1007/978-3-031-54857-4.

49. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res*. 2011;12(85):2825-2830. https://www.jmlr.org/papers/v12/pedregosa11a.html

50. Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. *Inf Processing Manage*. 2009;45(4):427-437. doi:10.1016/j.ipm.2009.03.002

51. Sagi O, Rokach L. Ensemble learning: a survey. *WIREs Data Mining Knowledge Discovery*. 2018;8(4):e1249. doi:10.1002/widm.1249

52. Muller D, Soto-Rey I, Kramer F. An analysis on ensemble learning optimized medical image classification with deep convolutional neural networks. *IEEE Access*. 2022;10:66467-66480. doi:10.1109/ACCESS.2022.3182399

53. GBD 2019 Blindness and Vision Impairment Collaborators; Vision Loss Expert Group of the Global Burden of Disease Study. Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the Right to Sight: an analysis for the Global Burden of Disease Study. *Lancet Global Health*. 2021;9(2):e144-e160. doi:10.1016/S2214-109X(20)30489-7

54. Naidoo KS, Fricke TR, Frick KD, et al. Potential lost productivity resulting from the global burden of myopia: systematic review, meta-analysis, and modeling. *Ophthalmology*. 2019;126(3):338-346. doi:10.1016/j.ophtha.2018.10.029

55. Tang J, Yuan M, Tian K, et al. An artificial-intelligence-based automated grading and lesions segmentation system for myopic maculopathy based on color fundus photographs. *Translational Vis Sci Technol*. 2022;11(6):16. doi:10.1167/tvst.11.6.16

56. Zheng B, Zhang M, Zhu S, et al. Research on an artificial intelligence-based myopic maculopathy grading method using EfficientNet. *Indian J Ophthalmol*. 2024;72(suppl 1):S53-S59. doi:10.4103/IJO.IJO_48_23

57. Sun Y, Li Y, Zhang F, et al. A deep network using coarse clinical prior for myopic maculopathy grading. *Comput Biol Med*. 2023;154:106556. doi:10.1016/j.compbiomed.2023.106556

58. Varadarajan AV, Poplin R, Blumer K, et al. Deep learning for predicting refractive error from retinal fundus images. *Investigative Ophthalmol Vis Sci*. 2018;59(7):2861-2868. doi:10.1167/iovs.18-23887

59. Deng J, Dong W, Socher R, Li LJ, Li K, Li FF. ImageNet: a large-scale hierarchical image database. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE; 2009:248-255, doi:10.1109/CVPR.2009.5206848.

60. Cuadros J, Bresnick G. EyePACS: an adaptable telemedicine system for diabetic retinopathy screening. *J Diabetes Sci Technol*. 2009;3(3):509-516. doi:10.1177/193229680900300315

61. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE; 2016:770-778, doi:10.1109/CVPR.2016.90.

62. Tan M, Le Q. EfficientNet: rethinking model scaling for convolutional neural networks. *Proc Machine Learning Res*. 2019;97:6105-6114. https://proceedings.mlr.press/v97/tan19a.html

63. Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: transformers for image recognition at scale. *arXiv*. Preprint posted online June 3, 2021. doi:10.48550/arXiv.2010.11929

64. Howard AG, Zhu M, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications. *arXiv*. Preprint posed online April 17, 2017. doi:10.48550/arXiv.1704.04861

65. Liu Z, Lin Y, Cao Y, et al. Swin transformer: hierarchical vision transformer using shifted windows. In: *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE; 2021:9992-100002, doi:10.1109/ICCV48922.2021.00986.

66. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. MobileNetV2: inverted residuals and linear bottlenecks. In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE; 2018:4510-4520, doi:10.1109/CVPR.2018.00474.

67. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, eds. *Medical Image Computing and Computer-Assisted Intervention*. Springer; 2015:234-241. Lecture Notes in Computer Science; vol 14563. doi:10.1007/978-3-319-24574-4_28

68. Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, Liang J. UNet++: a nested U-net architecture for medical image segmentation. In: Stoyanov D, Taylor Z, Carneiro G, et al, eds. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer; 2018:3-11. Lecture Notes in Computer Science; vol 11045. doi:10.1007/978-3-030-00889-5_1

69. Chen LC, Zhu Y, Papandreou G, Schroff F, Adam H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y, eds. *Computer Vision – ECCV 2018*. Springer; 2018:833-851. Lecture Notes in Computer Science; vol 11211. doi:10.1007/978-3-030-01234-2_49

70. Fan T, Wang G, Li Y, Wang H. MA-Net: a multi-scale attention network for liver and tumor segmentation. *IEEE Access*. 2020;8:179656-179665. doi:10.1109/ACCESS.2020.3025372

71. Liu Z, Mao H, Wu CY, Feichtenhofer C, Darrell T, Xie S. A ConvNet for the 2020s. *arXiv*. Preprint posted online March 2, 2022. doi:10.48550/arXiv.2201.03545

72. Tan M, Le Q. EfficientNetV2: smaller models and faster training. *Proc Machine Learning Res*. 2021;139:10096-10106. https://proceedings.mlr.press/v139/tan21a.html

73. Caron M, Touvron H, Misra I, et al. Emerging properties in self-supervised vision transformers. In: *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE; 2021:9630-9640, doi:10.1109/ICCV48922.2021.00951.