Evaluation of Aurora Fundus Camera Performance in Diabetic Retinopathy Screening by Ophthalmologist Grader and Visual Artificial Intelligence

Shang Ruan, Weiting Hu, Huixun Jia, ShanShan Wang, Tao Ye, Fenghua Wang

Purpose: To evaluate the images quality as well as specificity and sensitivity of diabetic retinopathy screening by Optomed Aurora fundus camera in comparison to standard screening procedure by traditional table top fundus camera with ophthalmologist's assessment. We also explored the possibility of using Aurora combined with Phoebus Medical DR algorithm system in diabetic retinopathy screening.

Key Words: Diabetic retinopathy, Image quality, Screening, Handheld Camera, DR algorithm

Background

Diabetic retinopathy (DR) is the most common cause of vision loss worldwide and the leading cause of vision impairment and blindness among working-age adults. Anyone with diabetes is at risk of developing diabetic retinopathy, or damage to the retina. More than 110 billion Chinese now live with diabetes, and approximately 24.7–37.5 percent of those with diabetes have some form of DR. Continued damage can lead to blindness.[1] Fortunately, DR can be detected and early intervention can prevent or reduce vision loss. For service providers, DR screening is not only cost-effective but, in the long term, it can even save costs. Traditional table top fundus camera has been widely used in DR screening. However, they have limitations in large-scale screening in communities due to their size and weight. Therefore, the application of handheld fundus camera is a suitable choice in diabetic retinopathy screening. Previous studies

have assessed the feasibility of handheld fundus cameras produced by Optomed[2-4]. This study focused on a new handheld nonmydriatic Aurora fundus camera used for screening of vision-threatening diabetic retinopathy. Aurora camera is the latest camera product by Optomed, a manufacturer of portable fundus cameras. The portable camera has 50 degrees field of view (FOV), non-mydriatic operation, 9 internal fixation targets and Wifi for transmitting images to the PC. Compared with retinal photography with manual interpretation, which is widely accepted for screening of diabetic retinopathy, automated grading of diabetic retinopathy with deep learning algorithm has potential benefits such as reduced barriers to access, increased efficiency, reproducibility, coverage of screening programs and so on.[5,6] In this study, we assessed the performance and feasibility of Aurora fundus camera and Phoebus Medical DR algorithm system in DR screening.

Methods

This was a multi-centered double-blinded observational clinical study. Patients were enrolled from three hospitals (Shanghai General Hospital, Shanghai, China; West Nanjing Road Community Health Center, Shanghai, China; and Zhaoqing Gaoyao People's Hospital, Guangdong, China). Fundus images were transferred to the grading center, Shanghai General Hospital through the INSIGHT real world patient registry platform www.chinadr.com.cn (Phoebusmed, Shanghai, China), for remote digital retinal imaging grading. This study was approved by Shanghai General Hospital Institutional Review Board at Shanghai Jiaotong University. Informed consent was obtained from all participants. This study was performed according to the ICH-GCP (International Conference on Harmonisation-Good Clinical Practice) guidelines and fulfilled the tenets of the Declaration of Helsinki. Phoebus Medical has developed an algorithm to detect changes related to diabetic retinopathy including retinal hemorrhage, hard exudate and macular edema so as to screen out referable DR (moderate or worse DR) for some further treatment.

Enrollment

Participants over 18 years of age with diagnosis of diabetes over 5 years were recruited from ophthalmology clinics at 3 centers. Patients were excluded if the patients had unvisualizable fundus or were pregnant. 47 patients were recruited at Shanghai General Hospital, 7 were excluded; 154 patients were recruited at West Nanjing Road Community Health Center, 5 were excluded; and 143 patients were recruited at Zhaoqing Gaoyao People's Hospital, 17 were excluded. All participants were divided into 4 groups for image quality cooperation.

Imaging

Ophthalmic photographers of three centers were trained with Aurora before a retina specialist deemed him/her able to obtain gradable images. The photographer aimed to take 4 images per eye, one macular centered image and one optic disc centered image from Aurora and table top fundus camera respectively. As a matter of fact, in the real world, images able to be acquired were less than expected.

Group 1: 40 participants from center 1, all fundus images were captured mydriatic by Aurora (Optomed, Oulu, Finland) and Visucam 200 (Zeiss, Oberkochen, Germany). Group 2: 149 participants from center 2, all fundus images were captured non-mydriatic by Aurora and CR-2 (Canon, Tokyo, Japan). Group 3: 81 participants from center 3, all fundus images were captured mydriatic by Aurora and TRC-50DX (Topcon, Tyoko, Japan). Group 4: 45 participants from center 3, all fundus images were captured non-mydriatic by Aurora and Reticam 3100 (Newvision, Chongqing, China). (Table 1)

Table 1. Enrollment and Fundus cameras used in three centers

Centers	Groups	Handheld camera	Table top camera	Participants	Pupil
Center 1: Shanghai General Hospital	Group 1	Aurora	Visucam 200	40	Mydriatic
Center 2: West Nanjing Road Community Health Center	Group 2	Aurora	CR-2	149	Non-Mydriatic
Center 3: Zhaoqing Gaoyao People's Hospital	Group 3	Aurora	TRC-50DX	81	Mydriatic
	Group 4	Aurora	Reticam 3100	45	Non-Mydriatic

Clinical Examination and Statistical Analysis

As an observational study, all participants received the standard clinical examination and treatment for the diabetes in the three centers. Data acquisition and grading in this study would not affect the diagnosis and treatment of the participants.

Performance of Aurora and tabletop fundus cameras in image quality

The fundus image was graded by three masked retina specialists according to "centration", "sharpness" and "visible range", each criterion was graded as "excellent", "acceptable", and "ungradable" respectively (Appendix 1). For "centration", "excellent" represented macula/optic disc within 1 PD range of the image center, "acceptable" represented macula/optic disc within 2 PD range of the image center and macula/optic disc out of 2 PD range of the image center was "ungradable". As for sharpness, three grades represented clear focused, recognizable and unrecognizable. For "visible range", 100% visible, more than 80% visible and less than 80% visible of the whole image were graded as "excellent", "acceptable", and "ungradable". The image quality was defined by aggregating scores of three criteria.

Performance of Aurora and tabletop fundus cameras DR screening result

In this study, we set the DR screening result graded by the same ophthalmologist

with all the fundus images taken by Aurora and tabletop fundus camera from the same patient as the gold standard. Then compare the DR screening result of this patient graded by the same ophthalmologist with Aurora fundus images and tabletop fundus image to evaluate their sensitivity and specificity separately.

For Aurora camera has 50 degrees field of view on non-mydriatic condition, which is wider than 45(mydriatic) or 30(non-mydriatic) degrees of table top camera. Combination of Aurora and table top cameras are expected to minimalize errors and false negative caused by photograph deficiency, difference in image quality, sharpness or visible range. The screening results included microaneurysm, retinal hemorrhage, venous beading, hard exudation, macular edema, drusen, cotton-wool spots, neovascularization, vitreous hemorrhage and so on. Microaneurysm, retinal hemorrhage, hard exudation and macular edema were critical to differentiate screening results of DR. In addition, considering that the low incidence of other clinical signs may increase statistic error, this study assayed mainly four signs of DR screening: microaneurysm, retinal hemorrhage, hard exudation and macular edema. Patients with no retina changes were graded as no DR. Patients with microaneurysms only were graded as mild DR. Referable DR were graded when retinal hemorrhage, hard exudation or macular edema were visible. In addition, referral suggestion was given according to the ophthalmologist's judgement. (Figure 1)

Figure 1. Grading Interface of INSIGHT Real World Platform

Results

Image Quality

Overall 630 eyes from 315 subjects were included in this study. Group 1: 40 participants were enrolled, 158 images taken by Aurora and 155 by Visucam 200. Mean score of centration (standard deviation, Sd) by Aurora is 1.65 (0.18) and Visucam 200 is 1.62 (0.13) (P=0.369). Mean score of sharpness of Aurora is 1.79 (0.20) and Visucam 200 is 1.79 (0.17) (P=0.95). Mean score of visible range of Aurora is 1.73 (0.23) and Visucam 200 is 1.59 (0.22) (P=0.007). Group 2: 149 participants were enrolled, 565 images taken by Aurora and 552 by CR-2. Mean score of centration by Aurora is 1.36 (0.42) and CR-2 is 1.15 (0.47) (P<0.001). Mean score of sharpness of Aurora is 1.40 (0.34) and CR-2 is 1.17 (0.41) (P<0.001). Mean score of visible range of Aurora is 1.26 (0.37) and CR-2 is 1.01 (0.46) (P<0.001). Group 3: 81 participants were enrolled, 324 images taken by Aurora and 316 by TRC-50DX. Mean score of centration by Aurora is 1.79 (0.26) and TRC-50DX is 1.79 (0.25) (P=0.912). Mean score of sharpness of Aurora is 1.70 (0.20) and TRC-50DX is 1.76 (0.25) (P=0.093). Mean score of visible

range of Aurora is 1.73 (0.26) and TRC-50DX is 1.80 (0.27) (P=0.077). Group 4: 45 participants were enrolled, 180 images taken by Aurora and 178 by Reticam 3100. Mean score of centration by Aurora is 1.74 (0.27) and Reticam 3100 is 1.71 (0.27) (P=0.684). Mean score of sharpness of Aurora is 1.73 (0.13) and Reticam 3100 is 1.65 (0.16) (P=0.007). Mean score of visible range of Aurora is 1.76 (0.18) and Reticam 3100 is 1.68 (0.22) (P=0.061). (Table 2 and Figure 2)

In mydriatic condition, there was no significant difference between the centration and sharpness score in group 1 (P>0.05), while there was statistical difference of visible range (P=0.007) in group 1. It shows that the image quality of Aurora is equal to Visucam 200 in centration and sharpness and Aurora gets better visible range in mydriatic condition. No significant differences were found in group 3 (P>0.05), which indicates that the image quality of Aurora is equal to TRC-50DX. In non-mydriatic condition, there were significant differences of the centration, sharpness and visible range in group 2 (P<0.001), as well as the sharpness in group 4 (P=0.007). It shows better performance of Aurora in centration, sharpness and visible range than CR-2, and better sharpness than Reticam 3100. Aurora may behave better than table top cameras in non-mydriatic condition in DR screening, especially for patients not suitable for pupil dilation. While in mydriatic condition Aurora's image quality is as good as standard table top cameras.

Table 2. Image quality score of Aurora and tabletop fundus cameras in three centers

Center 1				
		Group 1 (Mydriatic)		
	level	Aurora	Visucam 200	P
N (%)		40	40	
Score of Centration (Mean (Sd))		1.65 (0.18)	1.62 (0.13)	0.369
Score of Sharpness (Mean (Sd))		1.79 (0.20)	1.79 (0.17)	0.95
Score of Visible Range (Mean (Sd))		1.73 (0.23)	1.59 (0.22)	0.007
Score of Centration with levels (%)	0~1.5	6 (15.0)	3 (7.5)	0.015

	1.5~	24 (60.0)	35 (87.5)	
	1.8~	10 (25.0)	2 (5.0)	
Score of Sharpness with levels (%)	0~1.5	4 (10.0)	3 (7.5)	0.736
	1.5~	10 (25.0)	13 (32.5)	
	1.8~	26 (65.0)	24 (60.0)	
Score of Visible Range with levels (%)	0~1.5	7 (17.5)	7 (17.5)	0.001
	1.5~	12 (30.0)	27 (67.5)	
	1.8~	21 (52.5)	6 (15.0)	

Center 2

		Group 2 (Non-Mydriatic)		
	level	Aurora	CR-2	P
N (%)		149	149	
Score of Centration (Mean (Sd))		1.36 (0.42)	1.15 (0.47)	<0.001
Score of Sharpness (Mean (Sd))		1.40 (0.34)	1.17 (0.41)	<0.001
Score of Visible Range (Mean (Sd))		1.26 (0.37)	1.01 (0.46)	<0.001
Score of Centration with levels (%)	0~1.0	22 (14.8)	48 (32.2)	<0.001
	1.0~	54 (36.2)	59 (39.6)	
	1.5~	73 (49.0)	42 (28.2)	
Score of Sharpness with levels (%)	0~1.0	18 (12.1)	46 (30.9)	<0.001
	1.0~	60 (40.3)	61 (40.9)	
	1.5~	71 (47.7)	42 (28.2)	
Score of Visible Range with levels (%)	0~1.0	35 (23.5)	71 (47.7)	<0.001
	1.0~	65 (43.6)	51 (34.2)	
	1.5~	49 (32.9)	27 (18.1)	

Center 3

		Group 3 (Mydriatic)		
	level	Aurora	TRC-50DX	P
N (%)		81	81	
Score of Centration (Mean (Sd))		1.79 (0.26)	1.79 (0.25)	0.912

Score of Sharpness (Mean (Sd))		1.70 (0.20)	1.76 (0.25)	0.093
Score of Visible Range (Mean (Sd))		1.73 (0.26)	1.80 (0.27)	0.077
Score of Centration with levels (%)	0~1.5	12 (14.8)	4 (4.9)	0.015
	1.5~	14 (17.3)	27 (33.3)	
	1.8~	55 (67.9)	50 (61.7)	
Score of Sharpness with levels (%)	0~1.5	8 (9.9)	7 (8.6)	0.001
	1.5~	46 (56.8)	24 (29.6)	
	1.8~	27 (33.3)	50 (61.7)	
Score of Visible Range with levels (%)	0~1.5	12 (14.8)	11 (13.6)	0.046
	1.5~	27 (33.3)	14 (17.3)	
	1.8~	42 (51.9)	56 (69.1)	

Center 3

		Group 4 (Non-Mydriatic)		
	level	Aurora	Reticam 3100	P
N (%)		45	45	
Score of Centration (Mean (Sd))		1.74 (0.27)	1.71 (0.27)	0.684
Score of Sharpness (Mean (Sd))		1.73 (0.13)	1.65 (0.16)	0.007
Score of Visible Range (Mean (Sd))		1.76 (0.18)	1.68 (0.22)	0.061
Score of Centration with levels (%)	0~1.5	9 (20.0)	7 (15.6)	0.19
	1.5~	10 (22.2)	18 (40.0)	
	1.8~	26 (57.8)	20 (44.4)	
Score of Sharpness with levels (%)	0~1.5	1 (2.2)	6 (13.3)	0.036
	1.5~	29 (64.4)	32 (71.1)	
	1.8~	15 (33.3)	7 (15.6)	
Score of Visible Range with levels (%)	0~1.5	2 (4.4)	7 (15.6)	0.181
	1.5~	21 (46.7)	21 (46.7)	
	1.8~	22 (48.9)	17 (37.8)	

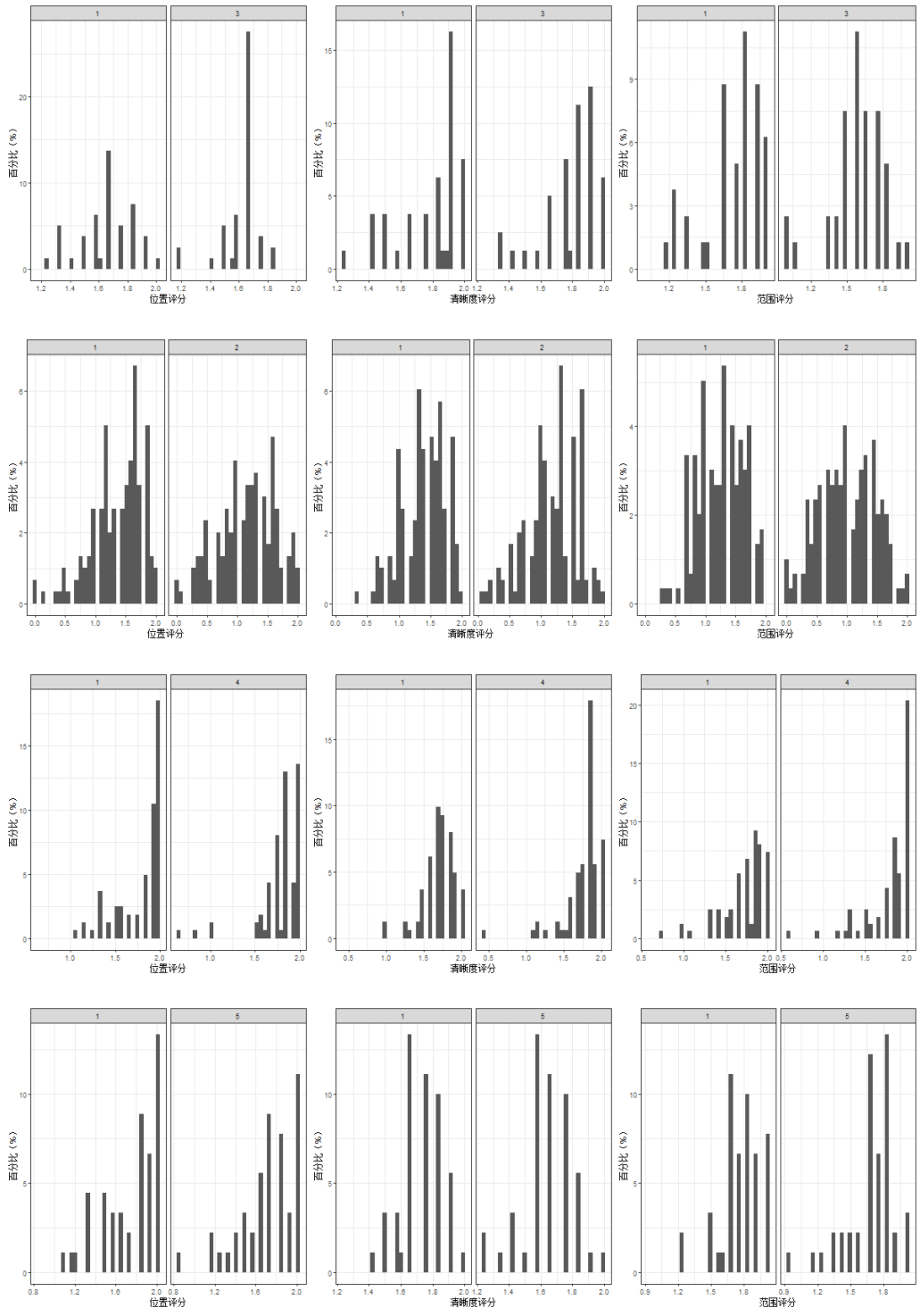

Figure 2. Distribution of centration, sharpness and visible range scores in three centers. (Top) Distribution of scores in center 1, Shanghai General Hospital, Aurora vs. Visucam 200, mydriatic; (Upper) Distribution of centration score in center 2, West Nanjing Road Community Health Center, Aurora vs. CR-2, non-mydriatic; (Lower) Distribution of centration score in center 3, Zhaoqing Gaoyao People's Hospital, Aurora vs. TRC-50DX, mydriatic; (Bottom) Distribution of centration score in center 3, Zhaoqing Gaoyao People's Hospital, Aurora vs. Reticam 3100, non-mydriatic; (Left) Distribution of centration score; (Middle) Distribution of sharpness score; (Right) Distribution

of vision range score. Group 1, Aurora; Group 2, CR-2; Group 3, Visucam 200; Group 4, TRC-50DX; Group 5, Reticam 3100.

Performance of Aurora camera and table top cameras in DR screening

Area Under the Curve (AUC) of Aurora screening of microaneurysm, retinal hemorrhage, hard exudate and macular edema were 0.964 (0.921~1.000), 0.893 (0.830~0.956), 0.946 (0.900~0.992) and 0.933 (0.874~0.993). AUC of table top cameras screening of microaneurysm, retinal hemorrhage, hard exudate and macular edema were 0.942 (0.893~0.990), 0.923 (0.874~0.972), 0.968 (0.935~1.000) and 0.942 (0.888~0.995). The sensitivity of Aurora screening microaneurysm, retinal hemorrhage, hard exudate and macular edema were 0.944 (0.870~1.000), 0.804 (0.690~0.919), 0.917 (0.838~0.995) and 0.882 (0.774~0.991) respectively. While the sensitivity of table top screening microaneurysm, retinal hemorrhage, hard exudate and macular edema were 0.897 (0.802~0.993), 0.860 (0.764~0.956), 0.940 (0.874~1.000) and 0.886 (0.780~0.991). The specificity of Aurora screening microaneurysm, retinal hemorrhage, hard exudate and macular edema were 0.984 (0.973~0.995), 0.982 (0.970~0.994), 0.975 (0.962~0.989) and 0.984 (0.973~0.993) respectively. While the specificity of table top screening microaneurysm, retinal hemorrhage, hard exudate and macular edema were 0.986 (0.976~0.996), 0.986 (0.976~0.996), 0.996 (0.990~1.000) and 0.998 (0.994~1.000). The sensitivity and specificity of referral requirement is 0.821, 0.927 and 0.975, 0.959 of Aurora and table top camera. (Table 3). No significant differences in sensitivity and specificity were found between Aurora and table top cameras. We assume that the performance of Aurora and table top cameras in DR screening results were comparable.

Table. 3 Performance of Aurora and Table Top cameras in DR Screening Result

Aurora Screening Result	AUC	CI of AUC	Sensitivity	CI of Sensitivity	Specificity	CI of Specificity
Microaneurysm	0.964	0.921 ~ 1.000	0.944	0.870 ~ 1.000	0.984	0.973 ~ 0.995
Retinal Hemorrhage	0.893	0.830 ~ 0.956	0.804	0.690 ~ 0.919	0.982	0.970 ~ 0.994

Hard Exudation	0.946	0.900 ~ 0.992	0.917	0.838 ~ 0.995	0.975	0.962 ~ 0.989
Macular Edema	0.933	0.874 ~ 0.993	0.882	0.774 ~ 0.991	0.984	0.973 ~ 0.995
Referral Requirement	0.898	0.846 ~ 0.949	0.821	0.721 ~ 0.922	0.974	0.954 ~ 0.995
Table Top Screening Result	**AUC**	**Cl of AUC**	**Sensitivity**	**Cl of Sensitivity**	**Specificity**	**Cl of Specificity**
Microaneurysm	0.942	0.893 ~ 0.990	0.897	0.802 ~ 0.993	0.986	0.976 ~ 0.996
Retinal Hemorrhage	0.923	0.874 ~ 0.972	0.860	0.764 ~ 0.956	0.986	0.976 ~ 0.996
Hard Exudation	0.968	0.935 ~ 1.000	0.940	0.874 ~ 1.000	0.996	0.990 ~ 1.000
Macular Edema	0.942	0.888 ~ 0.995	0.886	0.780 ~ 0.991	0.998	0.994 ~ 1.000
Referral Requirement	0.943	0.906 ~ 0.980	0.927	0.859 ~ 0.996	0.959	0.932 ~ 0.985

AUC, Area Under the Curve; Cl, Confidence interval.

Performance of DR Algorithm in DR Screening

Compared with gold standard, AUC of DR algorithm screening of retinal hemorrhage, hard exudate and macular edema were 0.777 (0.710~0.843), 0.765 (0.713~0.816) and 0.671 (0.585~0.757). The sensitivity of DR algorithm for detection of retinal hemorrhage, hard exudate and macular edema were 0.739 (0.612~0.866), 0.875 (0.781~0.969) and 0.441 (0.274~0.608). The specificity of retinal hemorrhage, hard exudate and macular edema were 0.815 (0.780~0.849), 0.654 (0.612~0.697) and 0.901 (0.874~0.927). The sensitivity and specificity of referral requirement is 0.882 and 0.407. (Table 4)

Table 4. Performance of DR Algorithm in DR Screening Result

Screening Result	**AUC**	**Cl of AUC**	**Sensitivity**	**Cl of Sensitivity**	**Specificity**	**Cl of Specificity**
Retinal Hemorrhage	0.777	0.710 ~ 0.843	0.739	0.612 ~ 0.866	0.815	0.780 ~ 0.849
Hard Exudation	0.765	0.713 ~ 0.816	0.875	0.781 ~ 0.969	0.654	0.612 ~ 0.697
Macular Edema	0.671	0.585 ~ 0.757	0.441	0.274 ~ 0.608	0.901	0.874 ~ 0.927
Referral Requirement	0.644	0.589 ~ 0.700	0.882	0.794 ~ 0.971	0.407	0.341 ~ 0.472

AUC, Area Under the Curve; Cl, Confidence interval.

Discussion

Diabetic retinopathy is one of the most common causes of visual loss, however, patients usually remain asymptomatic until severe complications occurs. DR screening is aimed to detect potential referable patients for comprehensive examination and standard treatment.

In this study, we evaluated the quality of image gathered from the Aurora handheld fundus camera. It may behave better than table top cameras in non-mydriatic condition in DR screening, especially for patients not suitable for mydriatic examination such as those with glaucoma. While in mydriatic condition Aurora's image quality is as good as standard table top cameras. We also tested the sensitivity and specificity of DR screening, no significant differences in sensitivity and specificity were found between Aurora and table top cameras. And for DR screening using Aurora+AI, the sensitivity and specificity of referral requirement is 0.882 and 0.407, The sensitivity and specificity of this system to detect retinal hemorrhage is >70% and 80% respectively. However, its detection for hard exudation and macular edema still required further learning and development. It suggests that, AI could be used for pre-screening to find out the referable DR patients before ophthalmologists grading as an auxiliary diagnostic tool.

Conclusion

Aurora fundus camera is able to acquire fundus images with comparable quality with traditional table top cameras with higher applicability and convenience in non-mydriatic situation. It could be a competitive technique used in community and hospital DR screening. DR algorithm-based automatic image grading to Aurora fundus images offers potential for even further improvements to the cost-effectiveness and convenience of screening programs

Appendix 1

Quality Criteria	Excellent (2')	Acceptable (1')	Ungradable (0')
Centration	Macula/optic disc is within 1	Macula/optic disc is	Macula/optic disc is out of 2

	PD range of the image center	within 2 PD range of the image center	PD range of the image center
sharpness	Clear focused	recognizable	unrecognizable
Visible Range	Visible range is the whole image	Visible range> 80% of the image	Visible range< 80% of the image

References:

[1] The National Committee for Prevention Blindness. The guideline for hierarchical diagnosis and treatment of diabetic retinopathy[J]. Chin J Gen Pract, 2017, 16(8): 589-593.

[2] Davila J R, Sengupta S S, Niziol L M, et al. Predictors of Photographic Quality with a Handheld Nonmydriatic Fundus Camera Used for Screening of Vision-Threatening Diabetic Retinopathy[J]. Ophthalmologica, 2017, 238(1-2): 89-99.

[3] Sengupta S, Sindal M D, Besirli C G, et al. Screening for vision-threatening diabetic retinopathy in South India: comparing portable non-mydriatic and standard fundus cameras and clinical exam[J]. Eye (Lond), 2018, 32(2): 375-383.

[4] Zhang W, Nicholas P, Schuman S G, et al. Screening for Diabetic Retinopathy Using a Portable, Noncontact, Nonmydriatic Handheld Retinal Camera[J]. J Diabetes Sci Technol, 2017, 11(1): 128-134.

[5] Ting D S W, Cheung C Y, 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[J]. JAMA, 2017, 318(22): 2211-2223.

[6] Gargeya R, Leng T. Automated Identification of Diabetic Retinopathy Using Deep Learning[J]. Ophthalmology, 2017, 124(7): 962-969.