ORIGINAL ARTICLE

Generative artificial intelligence for diabetic retinopathy screening in Hong Kong

Tsz Ho Shum[1], *MB BS (HK), FHKAM (Medicine)*, *Sunny Chi Lik Au*[2,3], *FCOphthHK, FHKAM (Ophthalmology)*, *Steffi Shing Yee Chong*[3], *MB BS (HK)*, *Yuen Fun Emmy Lau*[1], *MB ChB (CUHK), FHKAM (Medicine)*

[1] *Department of Medicine, Pamela Youde Nethersole Eastern Hospital, Hong Kong SAR, China*

[2] *Department of Ophthalmology, Tung Wah Eastern Hospital, Hong Kong SAR, China*

[3] *Department of Ophthalmology, Pamela Youde Nethersole Eastern Hospital, Hong Kong SAR, China*

Correspondence and reprint requests:
Dr Tsz Ho Shum, Department of Medicine, Pamela Youde Nethersole Eastern Hospital, Hong Kong SAR, China.
Email: drgthshum@gmail.com

Abstract

Objectives: To compare the performances of EyeArt and endocrinologists in diagnosing more-than-mild diabetic retinopathy (mtmDR) among patients with diabetes in a Hong Kong hospital.

Methods: Medical records of consecutive patients with diabetes mellitus aged ≥18 years who were screened for diabetic retinopathy (DR) at Pamela Youde Nethersole Eastern Hospital in February 2021 were retrospectively reviewed. Two fundus photographs (one macula-centered and one optic nerve head-centered) were taken from each eye and saved electronically. The photographs were independently graded by endocrinologists, and then by the EyeArt system and by two ophthalmologists specializing in vitreoretinal diseases (the gold standard). mtmDR is defined as moderate nonproliferative DR, severe nonproliferative DR, or proliferative DR, with or without the presence of clinically significant macular edema. Performances of EyeArt and the endocrinologists in diagnosing mtmDR were determined, as were factors associated with EyeArt's diagnostic accuracy and the variability between EyeArt and the endocrinologists.

Results: In total, 107 men and 57 women (328 eyes) were included in the analysis. Diagnostic performances of EyeArt and the endocrinologists were comparable in terms of sensitivity (88.0% vs 81.5%), specificity (90.9% vs 93.1%), positive predictive value (46.8% vs 53.7%), and negative predictive value (98.8% vs 98.1%). Gwet's first-order agreement coefficient was 0.832 ($p<0.001$), demonstrating almost perfect agreement between EyeArt and the endocrinologists. Inaccurate diagnosis by EyeArt was associated with the presence of hypertensive retinopathy($\chi^2=17.44$, $p<0.001$).

Conclusion: EyeArt demonstrated high sensitivity and specificity in diagnosing mtmDR among patients with diabetes, with almost perfect agreement with the endocrinologists. EyeArt can be used to screen for DR.

Key words: Diabetic retinopathy; Generative artificial intelligence; Mass screening; Sensitivity and specificity

Introduction

The global prevalence of diabetes mellitus was estimated to be 10.5% in 2021 and predicted to increase to 11.3% in 2030 and 12.2% in 2045.[1] Diabetic retinopathy (DR), a complication of diabetes mellitus, is a leading cause of new-onset blindness among adults aged 20 to 74 years in developed countries.[2] Globally, the number of adults with DR was estimated to be 103.12 million in 2020 and predicted to increase to 160.50 million by 2045.[3] Early detection and intervention can prevent vision loss. Screening for DR is recommended at the time of diagnosis of type 2 diabetes and within 5 years of diagnosis of type 1 diabetes.[2] However, 21% of patients with diabetes worldwide have not been screened, primarily owing to limited access to specialists.[4] In Hong Kong, approximately 20% of patients with diabetes have never been screened for DR.[5]

© 2025 Hong Kong Journal of Ophthalmology. CC BY-NC-ND 4.0

Artificial intelligence (AI) with machine learning algorithms has been used to screen for DR. Three AI systems—LumineticsCore (Digital Diagnostics, Coralville [IA], USA), EyeArt (Eyenuk, Woodland Hills [CA], USA), and AEYE Diagnostic Screening (AEYE Health, New York [NY], USA)—have demonstrated high sensitivity and specificity and have been approved by the US Food and Drug Administration (FDA).[6,7] However, these systems were developed predominantly using data from Caucasian populations and may have dataset bias.[8] Fundus photographs vary across ethnicities and refraction values. Diagnostic performance of AI screening for referable DR differs between lighter-and darker-skinned individuals if dataset bias is not addressed during development of the AI system.[9]

In Hong Kong, only EyeArt is available. EyeArt has shown promising results in screening high-risk patients, such as those with a history of mild DR, gestational diabetes, or newly diagnosed diabetes requiring hospitalization.[10] The current study aimed to compare the performance of EyeArt and endocrinologists in diagnosing more-than-mild DR (mtmDR) among patients with diabetes in a Hong Kong hospital.

Methods

Medical records of consecutive patients with diabetes aged ≥18 years who were screened for DR (in accordance with the American Diabetes Association's recommendations) at the Department of Medicine of Pamela Youde Nethersole Eastern Hospital in February 2021 were retrospectively reviewed. Patients who regularly visited an ophthalmologist or optometrist for DR or any other ophthalmological disease were excluded.

Two fundus photographs (one macula-centered and one optic nerve head-centered) were taken from each eye using the non-mydriatic auto fundus camera 330 (Nidek, Gamagori, Aichi, Japan) and saved electronically. The photographs were independently graded by endocrinologists (using the International Clinical Diabetic Retinopathy [ICDR] severity scale),[11] and then by the EyeArt system (version v2.1.0) and by two ophthalmologists specializing in vitreoretinal diseases (the gold standard).

The EyeArt system has been trained on >375 000 images and uses multiple image analysis algorithms.[12] It grades DR based on either the ICDR (International Clinical Diabetic Retinopathy) severity scale[11] or the UK NDESP (National Diabetic Eye Screening Programme) scale[13] (**Table 1**). mtmDR is defined as moderate nonproliferative DR, severe nonproliferative DR, or proliferative DR (equivalent to Early Treatment Diabetic Retinopathy Study [ETDRS] levels 35 to 85), with or without the presence of clinically significant macular edema. Vision-threatening DR (vtDR) is defined as severe nonproliferative DR or proliferative DR (ETDRS levels 35 to 85), with or without the presence of clinically significant macular edema. Ungradable eyes or eyes graded with mtmDR (including vtDR) would be referred to an ophthalmologist, as recommended by the International Council of Ophthalmology Guidelines for Diabetic Eye Care 2017.[14] A report was generated, showing gradings of each eye and summary of advised action.[15]

The fundus photographs were then graded by two ophthalmologists specializing in vitreoretinal diseases based on the ICDR severity scale. Additionally, myopic maculopathy was classified according to the International Photographic Classification and Grading System for Myopic Maculopathy.[16] Other incidental pathologies were also recorded.

Sensitivity, specificity, and positive and negative predictive values of EyeArt and the endocrinologists in diagnosing mtmDR were determined, as were factors associated with EyeArt's diagnostic accuracy and the variability between

Table 1. Gradings of diabetic retinopathy (DR) according to the International Clinical Diabetic Retinopathy (ICDR), the Early Treatment Diabetic Retinopathy Study (ETDRS), and the National Diabetic Eye Screening Programme (NDESP).

ICDR	Findings observable on dilated ophthalmoscopy	ETDRS	NDESP
No apparent retinopathy	No abnormalities	Level 10	R0: no apparent retinopathy
Mild nonproliferative DR	Microaneurysms only	Level 20	R1: background DR
Moderate nonproliferative DR	More than just microaneurysms but less than severe nonproliferative diabetic retinopathy	Level 35	
		Level 43	R2: pre-proliferative DR
		Level 47	
Severe nonproliferative DR	Any of the following and no signs of proliferative retinopathy: 20 intraretinal hemorrhages in each of the 4 quadrants, definite venous beading in ≥2 quadrants, prominent intraretinal microvascular abnormalities in ≥1 quadrant	Level 53	
Proliferative DR	One or more of the following: neovascularization, vitreous/preretinal hemorrhage	Levels 61, 65, 71, 75, 81, 85	R3: proliferative DR
No macular edema	No apparent retinal thickening or hard exudates in the posterior pole		
Macular edema	Some apparent retinal thickening or hard exudates in the posterior pole		

EyeArt and endocrinologists (who are responsible for screening for mtmDR among patients with diabetes in Hong Kong).

Statistical analyses were performed using SPSS (Windows version 26.0; IBM Corp, Armonk [NY], USA) and R Statistical Software (version 4.2.2, R Foundation for Statistical Computing, Vienna, Austria). Diagnostic performances of EyeArt and endocrinologists were compared, with the ophthalmologists' gradings being the gold standard. The required sample size was calculated using the R package.[17] Assuming EyeArt has 95.5% sensitivity and 85% specificity, the estimated number of eyes with mtmDR required to show 95.5% sensitivity with an expected confidence interval (CI) of 0.2 was 30. The estimated number of eyes without mtmDR required to show 85% specificity with an expected CI of 0.15 was 90. Based on a Hong Kong study reporting the prevalence of mtmDR to be 9.8% and the percentage of ungradable eyes or incomplete data to be 15%,[18] we intended to recruit 181 patients. The 95% CI was calculated using the Agresti-Coull interval method due to the poor coverage of the normal approximation interval (Wald interval), especially when the proportion near 0 or 1, and the over-conservative estimates by the Clopper-Pearson interval.[19] Factors associated with diagnostic accuracy were determined using the Chi-squared test or Student's *t* test. Diagnostic performances of EyeArt and endocrinologists were compared using Cohen's kappa, Gwet's first-order agreement (AC1) coefficient, and McNemar's test. For McNemar's test, an exact p value was used when the discordance was small (<25); otherwise, a mid p value was used.[20]

Table 2. Characteristics of participants (n=164).

Characteristic	Value*
Age, y	60 (53-66)
Sex	
Female	57 (34.8)
Male	107 (65.2)
Ethnicity	
Chinese	162 (98.8)
Asian origin other than Chinese	1 (0.6)
Caucasian	1 (0.6)
Body mass index, kg/m^2	25.7 (22.8-29.2)
Systolic blood pressure, mmHg	138 (128-150)
Total cholesterol, mmol/L	4.09 (3.54-4.79)
Estimated glomerular filtration rate, mL/min/1.73m^2	84.5 (70.8-98.0)
Diabetes type	
Type 1	4 (2.4)
Type 2	160 (97.6)
Duration of diabetes, y	8 (2-11)
Glycated hemoglobin,%	7.4 (6.5-7.8)

* *Data are presented as median (interquartile range) or No. (%) of participants*

Results

Of 187 patients with diabetes recruited, 23 were excluded owing to incomplete data or lost fundus photographs, and the remaining 107 men and 57 women (328 eyes) were included in the analysis (**Table 2**). Most were Chinese (98.8%) and had type 2 diabetes (97.6%), with a median disease duration of 8 years and a median glycated hemoglobin of 7.4%.

According to the two ophthalmologists specializing in vitreoretinal diseases (the gold standard), 25 eyes were deemed ungradable and 303 gradable (**Table 3**). Of the latter, 27 were classified as having mtmDR; 10 of which were vtDR and six of these also had macular edema. In addition, 61 eyes had myopic maculopathy, and 56 eyes had other concomitant ophthalmic pathologies including hypertensive retinopathy (**Figure**), cataracts, increased cup-disc ratio, epiretinal membrane, age-related macular degeneration, and central retinal vein occlusion.

Table 3. Comparison between EyeArt and endocrinologists in detecting more-than-mild diabetic retinopathy (mtmDR) based on the gold standard by two ophthalmologists.

Evaluator	No. of eyes with mtmDR as determined by two ophthalmologists	
	Yes (n=27)	No (n=276)
EyeArt		
mtmDR +	22	25
mtmDR –	3	249
Ungradable	2	2
Endocrinologists		
mtmDR +	22	19
mtmDR –	5	257

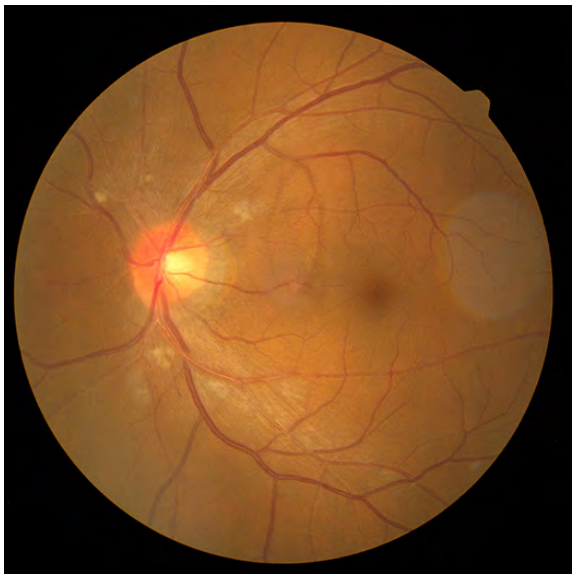

Figure. A fundus photograph of an eye with hypertensive retinopathy misclassified as having more-than-mild diabetic retinopathy by EyeArt.

Table 4. Diagnostic performances of EyeArt and endocrinologists.

Diagnostic performance	EyeArt*	Endocrinologists*
Sensitivity	88.0 (69.2-96.7)	81.5 (62.8-92.3)
Specificity	90.9 (86.8-93.8)	93.1 (89.4-95.6)
Positive predictive value	46.8 (33.3-60.8)	53.7 (38.7-67.9)
Negative predictive value	98.8 (96.4-99.8)	98.1 (95.5-99.3)

* *Data are presented as % (95% confidence interval)*

According to EyeArt, four eyes were deemed ungradable and 299 gradable. Of the latter, 47 were classified as having mtmDR; 20 of which were vtDR and 11 of these also had macular edema. Of the 25 eyes deemed ungradable by the ophthalmologists, 16 (64%) were correctly identified by EyeArt, three (12%) were misdiagnosed as having vtDR, and six (24%) were misdiagnosed as not having DR or having mild DR and were not referred to an ophthalmologist. Regarding the three false-negative cases, one did not have other ocular pathology, one had an epiretinal membrane, and one had hypertensive retinopathy. Among the 25 false-positive cases, 14 did not have other pathology, nine had hypertensive retinopathy, one had an increased cup-disc ratio, and one had central retinal vein occlusion. Inaccurate diagnosis by EyeArt was associated with the presence of hypertensive retinopathy (χ^2=17.44, p<0.001).

According to the endocrinologists, no eyes were deemed ungradable, and 41 eyes were classified as having mtmDR, three of which also had macular edema. Of the 25 eyes deemed ungradable by the ophthalmologists, 23 (92%) were classified by endocrinologists as not having DR and two (8%) as having mtmDR and were referred to an ophthalmologist.

Diagnostic performances of EyeArt and endocrinologists were comparable in terms of sensitivity (88.0% vs 81.5%), specificity (90.9% vs 93.1%), positive predictive value (46.8% vs 53.7%), and negative predictive value (98.8% vs 98.1%) [**Table 4**].

The variability between EyeArt and endocrinologists is shown in **Table 5**. For a conservative evaluation, which assumes all ungradable fundus photographs would be wrongly graded, the kappa coefficient between EyeArt and endocrinologists was 0.505 (p<0.001), which was considered moderate.[21] However, the kappa coefficient underestimate agreement when the prevalence is high.[22] This can be addressed using Gwet's AC1 coefficient,[23-26] which was 0.832 (p<0.001), indicating almost perfect agreement. McNemar's test was used to evaluate the equality of sensitivity and specificity in patients with and without mtmDR.[27] The exact p value for sensitivity was 1, and the mid p value for specificity was 0.216. Both p values exceeded the cutoff of 0.05, failing to reject the null hypothesis of equivalence.

Table 5. Variability between EyeArt and endocrinologists in detecting more-than-mild diabetic retinopathy (mtmDR).

EyeArt	Ophthalmologists*			
	mtmDR +		mtmDR –	
	Endocrinologists			
	mtmDR +	mtmDR –	mtmDR +	mtmDR –
mtmDR +	19	3	7	20
mtmDR –	3	2	12	237

* *Data are presented as No. of eyes*

Discussion

Generative AI has been successfully applied in various medical specialties including radiology, dermatology, and pathology. For DR screening, three AI systems—LumineticsCore, EyeArt, and AEYE Diagnostic Screening—have been approved by the FDA. Other AI systems for retinopathy screening have also been approved in Europe and China.[28]

EyeArt has 95.5% sensitivity and 85% specificity.[7] However, in a UK study involving >30 000 patients, EyeArt demonstrated 95.7% sensitivity but only 54% specificity.[29] This lower specificity may be due in part to grading DR on the NDESP scale, wherein ETDRS level 35 DR does not need referral. The UK's National Screening Committee still deemed EyeArt to be a safe and cost-effective screening tool.[30]

Our findings confirmed EyeArt's performance in diagnosing mtmDR, with 88.0% sensitivity and 90.9% specificity, which surpass the 85% sensitivity and 82.5% specificity that the FDA sets for evaluating AI.[6] The relatively lower positive predictive value (46.8%) is likely due to the low prevalence of mtmDR in our sample. Our findings support the use of EyeArt for routine DR screening of Chinese patients with diabetes attending specialist outpatient clinics. Based on the point estimates, EyeArt had slightly higher sensitivity but lower specificity than the endocrinologists. Diagnostic performances of EyeArt and endocrinologists were comparable, with a Gwet's AC1 coefficient of 0.832. Further studies are warranted to evaluate a two-step screening approach, in which fundus photographs labeled mtmDR-positive by EyeArt are subsequently reviewed by endocrinologists. It must be ensured that the use of EyeArt for DR screening will not compromise patient care. EyeArt is simple to use, requires limited training, and is easily scalable, allowing timely screening amid limited availability of endocrinologists or ophthalmologists. Moreover, EyeArt can be used for point-of-care screening, providing instant feedback on image quality and allowing instant retaking of the fundus photographs if the quality were suboptimal. Furthermore, immediate availability of grading alleviates patient anxiety, enabling better patient experience and improving adherence to eye care recommendations.[31]

The present study had several limitations. First, this was a single-center study that excluded patients who were regularly followed by an ophthalmologist or optometrist; thus, patients with mtmDR may have been underrepresented. Although the actual population attending the DR screening program was included, the results may not be generalizable to primary care settings due to the spectrum effect. Second, this cross-sectional study used retrospective data. As all consecutive visits were included, the diagnostic performance was unlikely to be affected. However, a prospective design would have allowed retaking of fundus photographs in cases of ungradable images and evaluation of patient acceptance, user experiences of AI, and the potential benefit of immediate grading on compliance with subsequent follow-ups. Third, the 95% CI of sensitivity in diagnosing mtmDR overlapped the null hypothesis required for FDA approval. To replicate the results that fulfil the FDA regulatory requirements of 85% sensitivity and 82.5% specificity with a null hypothesis of 75% sensitivity and 77.5% specificity, a larger sample size is needed. Although EyeArt is approved for vtDR detection, we could not evaluate its performance in diagnosing vtDR owing to the limited number of patients with vtDR. Lastly, although hypertensive retinopathy was associated with inaccurate diagnosis by EyeArt, only 36 of the 328 eyes had hypertensive retinopathy, and the prevalence of other ophthalmic pathologies was even lower. Further studies targeting concomitant ophthalmic pathologies are warranted to determine factors affecting diagnostic accuracy.

Conclusion

EyeArt demonstrated high sensitivity and specificity in diagnosing mtmDR among patients with diabetes, with almost perfect agreement with the endocrinologists. EyeArt can be used to screen for DR; nonetheless, a multicenter prospective study with a larger sample is warranted to validate its diagnostic performance.

Contributors

All authors designed the study, acquired the data, analyzed the data, drafted the manuscript, and critically revised the manuscript for important intellectual content. All authors had full access to the data, contributed to the study, approved the final version for publication, and take responsibility for its accuracy and integrity.

Conflicts of interest

As an editor of the journal, SCLA was not involved in the peer review process. Other authors have disclosed no conflicts of interest.

Funding/support

Licenses of using EyeArt for 250 patients and the technical support of EyeArt were provided by Eyenuk. The funder had no role in study design, data collection/analysis/interpretation or manuscript preparation.

Data availability

All data analyzed in this study are available from the corresponding author upon reasonable request.

Ethics approval

This study was approved by the Hong Kong East Cluster Research Ethics Committee (reference: HKECREC-2022-019). The requirement for patient consent was waived by the committee due to the retrospective nature of the study. The patients were treated in accordance with the tenets of the Declaration of Helsinki.

Acknowledgements

The authors thank Dr Callie Ka Li Ko for assistance in grading fundus photographs.

References

1. *International Diabetes Federation. IDF Diabetes Atlas 2025. Accessed 25 June 2024. Available from: https://diabetesatlas.org/resources/idf-diabetes-atlas-2025/*
2. *American Diabetes Association Professional Practice Committee. 16. Diabetes care in the hospital: standards of care in diabetes-2024. Diabetes Care 2024; 47(Suppl 1):S295-306.* Crossref
3. *Teo ZL, Tham YC, Yu M, et al. Global prevalence of diabetic retinopathy and projection of burden through 2045: systematic review and meta-analysis. Ophthalmology 2021; 128:1580-91.* Crossref
4. *The Diabetic Retinopathy Barometer Report Global Findings. Accessed 25 June 2024. Available from: https://www.iapb.org/wp-content/uploads/DR-Global-Report-1.pdf*
5. *Li KKW, Wong L, McGhee SM, et al. Diabetic retinopathy screening for specialist care. Hong Kong Med J 2019; 25(Suppl 9):S4-7.*
6. *Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. NPJ Digit Med 2018;1:39.* Crossref
7. *Ipp E, Liljenquist D, Bode B, et al. Pivotal evaluation of an artificial intelligence system for autonomous detection of referrable and vision-threatening diabetic retinopathy. JAMA Netw Open 2021; 4:e2134254.* Crossref
8. *Bernhardt M, Jones C, Glocker B. Potential sources of dataset bias complicate investigation of underdiagnosis by machine learning algorithms. Nat Med 2022; 28:1157-8.* Crossref

9. Burlina P, Joshi N, Paul W, Pacheco KD, Bressler NM. Addressing artificial intelligence bias in retinal diagnostics. Transl Vis Sci Technol 2021; 10:13. Crossref
10. Au SCL, Shum GTH, Chong SSY, Ko CKL. Accuracy of the EyeArt artificial intelligence system in detecting referable diabetic retinopathy in patients with diabetes in Hong Kong. Hong Kong J Ophthalmol 2023; 27:54-9. Crossref
11. Wilkinson CP, Ferris FL 3rd, Klein RE, et al. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. Ophthalmology 2003; 110:1677-82. Crossref
12. National Eye Institute Strategic Plan. Vision for the Future 2021-2025. Accessed 25 June 2024. Available from: https://www.nei.nih.gov/sites/default/files/2021-12/NEI-StrategicPlan-VisionForTheFuture_508_edit.pdf
13. NHS Diabetic Eye Screening Programme: grading definitions for referable disease. Accessed 25 June 2024. Available from: https://www.gov.uk/government/publications/diabetic-eye-screening-retinal-image-grading-criteria/nhs-diabetic-eye-screening-programme-grading-definitions-for-referable-disease-start-date-october-01
14. Wong TY, Sun J, Kawasaki R, et al. Guidelines on diabetic eye care: the International Council of Ophthalmology recommendations for screening, follow-up, referral, and treatment based on resource settings. Ophthalmology 2018; 125:1608-22. Crossref
15. EyeArt. Diabetic Retinopathy Screening Report. Accessed 25 June 2024. Available from: https://www.eyenuk.com/wp-content/uploads/2019/03/EyeArt_OUS_combined_20190327.pdf
16. Ohno-Matsui K, Kawasaki R, Jonas JB, et al. International photographic classification and grading system for myopic maculopathy. Am J Ophthalmol 2015; 159:877-83.e7. Crossref
17. Haynes A, Lenz A, Stalder O, Limacher A. presize: an R-package for precision-based sample size calculation in clinical research. J Open Source Softw 2021; 6:3118. Crossref
18. Lian JX, Gangwani RA, McGhee SM, et al. Systematic screening for diabetic retinopathy (DR) in Hong Kong: prevalence of DR and visual impairment among diabetic population. Br J Ophthalmol 2016; 100:151-5. Crossref
19. Agresti A, Coull BA. Approximate is better than 'exact' for interval estimation of binomial proportions. Am Stat 1998; 52:119-26. Crossref
20. Pembury Smith MQR, Ruxton GD. Effective use of the McNemar test. Behav Ecol Sociobiol 2020; 74:133. Crossref
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33:159-74. Crossref
22. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther 2005; 85:257-68. Crossref
23. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. Br J Math Stat Psychol 2008; 61:29-48. Crossref
24. Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. BMC Med Res Methodol 2013; 13:61. Crossref
25. Konstantinidis M, Le LW, Gao X. An empirical comparative assessment of inter-rater agreement of binary outcomes and multiple raters. Symmetry 2022; 14:262. Crossref
26. Xie Q. Agree or disagree? A demonstration of an alternative statistic to Cohen's Kappa for measuring the extent and reliability of agreement between observers. Accessed 25 June 2024. Available from: https://nces.ed.gov/FCSM/pdf/J4_Xie_2013FCSM.pdf
27. Trajman A, Luiz RR. McNemar chi2 test revisited: comparing sensitivity and specificity of diagnostic examinations. Scand J Clin Lab Invest 2008; 68:77-80. Crossref
28. Li S, Zhao R, Zou H. Artificial intelligence for diabetic retinopathy. Chin Med J (Engl) 2021; 135:253-60. Crossref
29. Heydon P, Egan C, Bolter L, et al. Prospective evaluation of an artificial intelligence-enabled algorithm for automated diabetic retinopathy screening of 30 000 patients. Br J Ophthalmol 2021; 105:723-8. Crossref
30. Automated grading in the Diabetic Eye Screening Programme. External review against programme appraisal criteria for the UK National Screening Committee. Accessed 25 June 2024. Available from: https://assets.publishing.service.gov.uk/media/619e80088fa8f5037e8ccb2f/Evidence_summary_AI_in_DESP_2021.pdf
31. Liu J, Gibson E, Ramchal S, et al. Diabetic retinopathy screening with automated retinal image analysis in a primary care setting improves adherence to ophthalmic care. Ophthalmol Retina 2021; 5:71-7. Crossref