

Simultaneous screening and classification of diabetic retinopathy and age-related macular degeneration based on fundus photos—a prospective analysis of the RetCAD system

Christos Skevas¹, Hanah Weindler¹, Max Levering¹, Jonne Engelberts², Mark van Grinsven², Toam Katz¹

¹University Clinic Eppendorf Hamburg, Hamburg 20246, Germany

²Thirona B.V., Toernooiveld 300 6525 EC, Nijmegen, The Netherlands

Correspondence to: Christos Skevas. University Clinic Eppendorf, Martinistr.52, Hamburg 20246, Germany. c.skevas@uke.de

Received: 2022-02-17 Accepted: 2022-09-19

Abstract

• **AIM:** To assess the accuracy of an artificial intelligence (AI) based software (RetCAD, Thirona, The Netherlands) to identify and grade age-related macular degeneration (AMD) and diabetic retinopathy (DR) simultaneously based on fundus photos.

• **METHODS:** This prospective study included 1245 eyes of 630 patients attending an ophthalmology day-care clinic. Fundus photos were acquired and parallel graded by the RetCAD AI software and by an expert reference examiner for image quality, and staging of AMD and DR. Adjudication was provided by a second expert examiner in case of disagreement between the AI software and the reference examiner. Statistical analysis was performed on eye-level and on patient-level, by summarizing the individual image level-gradings into an eye-level or patient-level score, respectively. The performance of the RetCAD system was measured using receiver operating characteristics (ROC) analysis and sensitivity and specificity for both AMD and DR were reported.

• **RESULTS:** The RetCAD achieved an area under the ROC (Az) of 0.926 with a sensitivity of 84.6% at a specificity of 84.0% for image quality. On image level, the RetCAD software achieved Az values of 0.964 and 0.961 with sensitivity/specificity pairs of 98.2%/79.1% and 83.9%/93.3% for AMD and DR, respectively. On patient level, the RetCAD software achieved Az values of 0.960 and 0.948 with sensitivity/specificity pairs of 97.3%/73.3% and 80.0%/90.1% for AMD and DR, respectively. After

adjudication by the second expert examiner sensitivity/specificity increases on patient-level to 98.6%/78.3% and 100.0%/92.3% for AMD and DR, respectively.

• **CONCLUSION:** The RetCAD offers very good sensitivity and specificity compared to manual grading by experts and is in line with that obtained by similar automated grading systems. The RetCAD AI software enables simultaneous grading of both AMD and DR based on the same fundus photos. Its sensitivity may be adjusted according to the desired acceptable sensitivity and specificity. Its simplicity cloud base integration allows cost-effective screening where routine expert evaluation may be limited.

• **KEYWORDS:** RetCAD; age-related macular degeneration; diabetic retinopathy

DOI:10.18240/ijo.2022.12.14

Citation: Skevas C, Weindler H, Levering M, Engelberts J, van Grinsven M, Katz T. Simultaneous screening and classification of diabetic retinopathy and age-related macular degeneration based on fundus photos—a prospective analysis of the RetCAD system. *Int J Ophthalmol* 2022;15(12):1985-1993

INTRODUCTION

Diabetic retinopathy (DR) and age-related macular degeneration (AMD) are some of the most common etiologies of visual impairment in adults in developed countries^[1-2]. Diabetes mellitus with its rising prevalence especially in the industrialized world is a disease with multiple risks. With an estimated prevalence of 592 million by 2035, it's spreading like a worldwide epidemic^[3]. The inconsiderate level of blood sugar in diabetic patients may lead to irreversible microvascular changes and complications. The 32.4 million blind and 191 million visually impaired people 0.8 million (2.6%) and 3.7 million (1.9%) respectively were because of DR. In 1990 only 2.1% of the blind and 1.3% if the visually impaired were DR related^[4]. With a good chance of preventing irreversible blindness, regular retinal checks are necessary for diabetic patients.

AMD is a central retinal disease common in patients over 50 years old. It is a leading cause of blindness worldwide and shows several environmental and genetic components with no effective prevention therapy. It is estimated that about 7.6% of United States population over 60y have intermediate or advanced AMD^[5]. Because of the demographic changes the prevalence will be rising within the next decades AMD can be categorized into early, intermediate and late AMD. Within late AMD there is dry AMD (geographic atrophy) and wet AMD [choroidal neovascularization (CNV)]. Currently there are several options for treating wet AMD with intravitreal injections, although a cure is not possible up till date. Dry AMD has no treatment options but must be regularly checked up.

The nature of these slow progressing diseases and the need for timely intervention depending on gradual changes in the clinical signs visible in the fundus justify scheduled clinical examinations of the fundus using a slit lamp or a fundus camera, with follow-ups over many years by an ophthalmologist. The purpose of these screening examinations is to classify the severity of these diseases triaging between further follow-up by an ophthalmologist and/or need for surgical and laser treatments or continuing regular screening. The manual screening demands constant availability of specialized personnel worldwide, repeated transport of patients to a nearby physician and availability of standardized tools for classification and triage. The manual screening causes high costs to the health system although only a small part is triaged for intervention. Moreover, among the elderly the physical availability of screening is often impaired.

Several automated systems for analyzing fundus photos and triaging AMD or DR were developed and tested. As it has been shown, this artificial intelligence (AI) based tools can have a high sensitivity and specificity concerning the detection of DR in unselected patients regularly visiting their eye-clinic as well as in a cohort of diabetic patients^[6-9]. There are also deep learning models able to grade AMD stages with a high sensitivity and specificity^[10-11] and even estimate the progression probability^[5]. These tools are based on fundus photos which are taken by a local technician and either locally interpreted or uploaded to a cloud where they are evaluated. Within a short amount of time the results are sent back to the acquiring technician with an exact numeric value to implicate for example the need to be reevaluated by an ophthalmologist. These image assessment software tools may be more cost-efficient concerning for example diabetes patients, because the increasing incidence of diabetes and standardized DR screening protocols, allowing for cost-effectively automation of this process^[12].

The RetCAD software allows either cloud based or local differentiating between a normal and diseased retina as well as

diagnosis of existence and grading of both DR and AMD by scoring the image for quality grading of either disease based on color fundus images alone.

SUBJECTS AND METHODS

Ethical Approval This prospective cross-sectional study was performed at the Department of Ophthalmology of the University Medical Centre Hamburg-Eppendorf. The study is registered and approved by the Ethics Review Board of the medical association Hamburg (registered study number: PV7377) and follows the recommendations of the Declaration of Helsinki. Written informed consent was obtained by each patient.

Fundus Photo Acquisition Patients were prospectively included from April 2020 to July 2020. Inclusion criteria were an age of at least 18y and eyes in which clear media allowed a sharp fundus photo. As the aim was to compare the detection capacity of the AI software, no pre-selection of known DR or AMD patients was done. This study included 3609 fundus photos of 630 patients consecutively attending a single day eye-clinic visit. Two examples of a sharp fundus photo can be seen in Figure 1. Fundus camera device and imaging protocol: All photos were taken by the same fundus camera (Topcon TRC-NW400) and the same technician (Levering M) acquiring 1 to 3 photos per eye with the focus on the optic nerve head (ONH), macula or the central fundus. All photos were taken without pupil dilation.

All fundus photographs were anonymized and uploaded to the cloud platform where the RetCAD software v1.3.1 (Thirona, The Netherlands) was applied to the images. All images were assigned a quality score, DR score and AMD score by the AI algorithm. The quality score ranged from 0 to 100 where a higher score indicates a better-quality fundus image. The DR and AMD scores range from 0 to 100, where a higher score means a more severe stage of DR or more severe AMD is detected. The software is calibrated in such a way that an AMD or DR score ≥ 50 indicates referable AMD or DR.

Manual Fundus Photo Grading All fundus photos were each individually graded by a single trained ophthalmologist (Weindler H), who's gradings were used as the reference standard in this study. Each image was first graded for image quality and if image quality was deemed sufficient, the image was graded for DR and AMD. The classification for DR followed the ETDRS classification whereas for AMD, the grading followed the AREDS classification.

DR grading: 0 to 5 (0=no DR, 1=mild DR (microaneurysms only), 2=moderate non-proliferative DR (NPDR; microaneurysms, bleeding, exudates, cotton wool spots), 3=severe NPDR (severe retinal bleedings, cotton wool spots, venous beading), 4=very severe NPDR (severe retinal bleedings in all 4 quadrants or significant venous beading in at least 2 quadrants or moderate

intraretinal microaneurysms in at least one quadrant), 5=proliferative DR. A DR stage of 0 or 1 is referred to a non-referable DR whereas DR stages 2 or higher are referred to as referable DR.

AMD grading: 0 to 3 (0=no AMD, 1=early AMD (medium-scaled drusen >63 μm to <125 μm), 2=intermediate DR (large drusen >125 μm), 3=advanced AMD (neovascular AMD and/or geographic atrophy). An AMD stage of 0 or 1 is referred to as non-referable AMD, whereas an AMD stage of 2 or 3 is referred to as referable AMD.

A second, senior ophthalmologist (Skevas C) provided adjudication in case there was a discrepancy between the AI and the first grader. This second grader followed the same classification scheme.

Outcome Measures RetCAD outcomes were compared to the reference standard on both eye-level and patient-level.

For eye-level, the eye-based AMD/DR score was determined as being the maximum AMD/DR score for the images belonging to that eye. For example, when an eye had 3 good quality images, the AMD/DR score was set as the maximum of these 3 AMD/DR scores. The reference standard was set in a similar way, by taking the highest AMD/DR grade of all the images belonging to that eye. Images which were graded as insufficient image quality by the reference grader were left out in the analysis (both for the AI and the manual grader). In other words, an eye level score could thus also be based on a single image if all other images were marked as insufficient image quality by the reference grader.

On patient-level, the same scoring methodology was used, but now all scores/grades from images from the patient were used. Also here, images graded as insufficient image quality by the reference grader were discarded.

Statistical Analysis The analysis was done in eye-level and in patient-level. Receiver operating characteristic (ROC) analyzes were carried out to evaluate the RetCAD software against the manual reference for the quality evaluation and the distinction between the transferable and non-transferable AMD and DR. The sensitivity and specificity of RetCAD are shown in the ROC curves for various threshold values. For the calibrated threshold of 50, the sensitivity and specificity of the RetCAD software were also determined. The general efficiency of the software is shown by the value of the area under the ROC curve (Az). The statistical analysis and creation of the diagrams were made with the Python module Scikit-learn 0.24.1. The distribution to the respective stages is visualized in box plots and histograms.

RESULTS

Image Quality Analysis of the Fundus Photographs All images were graded as either good image quality or insufficient image quality by the reference grader. Table 1 shows the

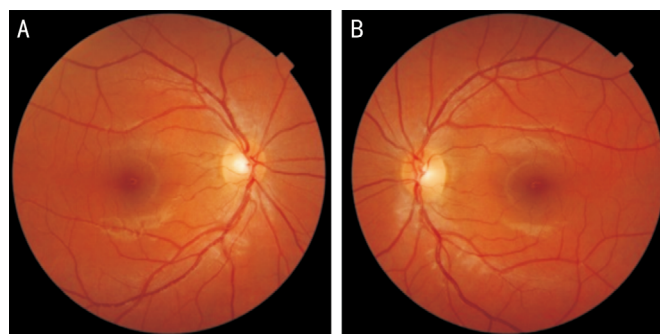


Figure 1 Example good quality fundus images of right eye (A) and left eye (B) of a patient Fundus photo analysis using RetCAD.

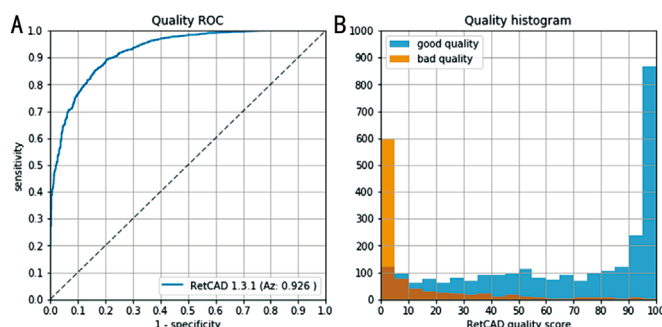


Figure 2 Receiver operating characteristic analysis for image quality (A) and image quality score distribution (B) Good quality images as graded by the reference grader (blue), and quality images as graded by the reference observer (orange).

Table 1 Image quality grading by the reference grader

Level	Frequency of good quality images						Total	
	0	1	2	3	4	5		6
Image-level [# images]	909	2700	-	-	-	-	-	3609
Eye-level [# eyes]	183	126	237	696	3	-	-	1245
Patient-level [# patients]	27	32	42	113	77	97	242	630

Each image was assigned an image quality score by RetCAD.

number of good quality images, good quality images per eye and good quality images per patient. An eye or patient was used for further analysis if at least 1 image was graded as good quality for the eye or patient, respectively.

Figure 2 shows the distribution of the image quality scores, where blue color means good quality according to the reference grader and orange indicate bad quality by the reference grader. The Y-axis shows the quantity of the images, on the X-axis is the quality score of RetCAD shown. On the extreme parts of the score the rating of RetCAD corresponds with the manually rated scores. The ROC Analysis shows the specificity and sensitivity of the differentiation between good and bad image quality by RetCAD. At a cut-off threshold of 25, the sensitivity is 84.6% whereas the specificity is 84.0%. The Az reaches a value of 0.926. Of 3609 images, 2700 have sufficient quality, while 909 have bad quality. If at least one image per eye, respectively one image per patient was of sufficient quality, it was integrated in the further analysis. Out

of 1245 photographed eyes, 1062 with good image quality could be used for the study. For the analysis on a per patient basis, the image quality was sufficient in 603 of 630 patients. This corresponds to 85.3% of all eyes and 95.7% of all patients who could be included.

The AMD gradings by the reference grader on eye level were as follows: stage 0: 911 eyes, stage 1: 94 eyes, stage 2: 31 eyes, stage 3: 26 eyes (Figure 3A). For DR, these were as follows: stage 0: 1007 eyes, stage 1: 24 eyes, stage 2: 31 eyes, stage 3, 4 and 5: 0 eyes (Figure 3B). This means that 57/1062 eyes were graded as referable AMD and 31/1062 eyes were graded as referable DR.

The box-plots in Figure 4 show a summary of the RetCAD AMD and DR scores for the different AMD and DR stages on eye-level. The orange bars represent the medians of the stages. For AMD, these are 20.78 for stage 0, 65.315 for stage 1, 82.85 for stage 2, and 81.795 for stage 3; whereas for DR, these are 15.71 for stage 0, 66.0 for stage 1 and 67.61 for stage 2.

Figure 5 shows the ROC graphs of the AI system for the detection of referable AMD and DR on eye-level, respectively. The RetCAD software obtained an Az value of 0.964 for detecting referable AMD and an Az value of 0.961 for the detection of referable DR. At the operating point with a cut-off of 50, the sensitivity of the RetCAD software for AMD is 98.2% and the specificity is 79.1%. For DR, the sensitivity is 83.9% and specificity is 93.3%.

The AMD gradings by the reference grading on patient level were as follows: stage 0: 497 patients, stage 1: 68 patients, stage 2: 16 patients, stage 3: 22 patients (Figure 6A). For DR, these were as follows: stage 0: 559 patients, stage 1: 19 patients, stage 2: 25 patients, stage 3, 4 and 5: 0 patients (Figure 6B). This means that 38/603 patients were graded as referable AMD and 25/603 patients were graded as referable DR.

The box-plots in Figure 7 shows a summary of the RetCAD AMD and DR scores for the different AMD and DR stages on patient-level. The orange bars indicate the median RetCAD values per grading stage.

Figure 8 shows the ROC graphs of the AI system for the detection of referable AMD and DR on patient-level, respectively. The RetCAD software obtained an Az value of 0.960 for detecting referable AMD and an Az value of 0.948 for the detection of referable DR. At the operating point with a cut-off of 50, the sensitivity of the RetCAD software for AMD is 97.3% and the specificity is 73.3%. For DR, the sensitivity is 80.0% and specificity is 90.1%.

Analysis of Disagreement An analysis of the disagreement between the reference grading and the RetCAD software was made on eye-level and patient-level when using the pre-defined fixed threshold of 50 for both AMD and DR for the RetCAD software.

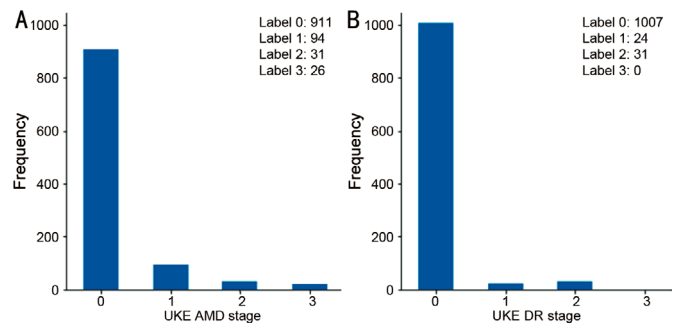


Figure 3 Frequency of age-related macular degeneration and diabetic retinopathy gradings by the reference grader on eye-level.

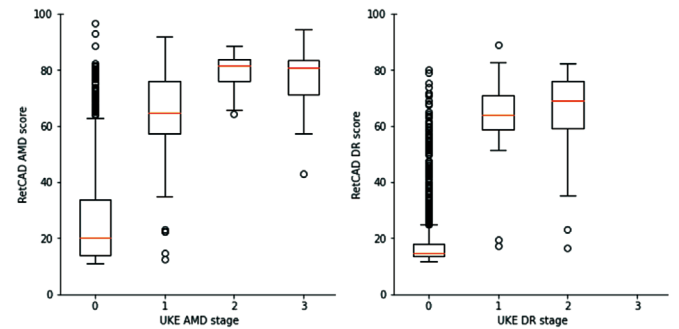


Figure 4 Box plot of RetCAD scores per age-related macular degeneration and diabetic retinopathy stage.

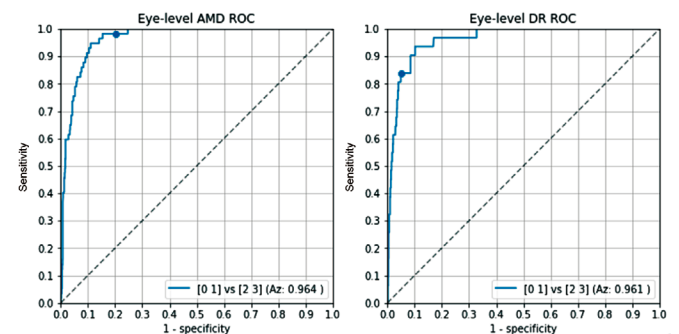


Figure 5 Receiver operating characteristic analysis on eye-level for referable age-related macular degeneration and referable diabetic retinopathy. The dots depict the operating point at a cut off of 50.

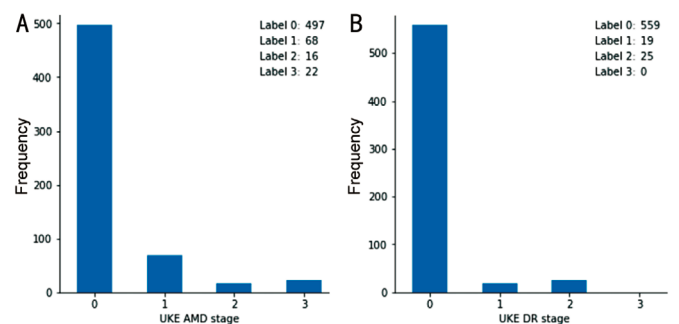


Figure 6 Frequency of age-related macular degeneration and diabetic retinopathy gradings by the reference grader on patient-level.

As can be seen in Figure 9, 57 eyes were graded as referable AMD of which 56 were identified by the RetCAD software at a cut-off of 50. The 1 eye that was missed by the AI was

confirmed to be referable AMD by the 2nd grader. Of the 1005 eyes graded as non-referable AMD, the AI software indicated that 210 eyes had referable-AMD of which grader 2 confirmed 38 eyes to have referable AMD, 155 eyes did not have referable AMD and 17 eyes were deemed with insufficient image quality.

As can be seen in Figure 10, 31 eyes were graded as referable DR of which 26 were identified by the RetCAD software at a cut-off of 50. The 5 eyes that were missed by the AI were confirmed to be non-referable DR by the 2nd grader. Of the 1031 eyes graded as non-referable DR, the AI software indicated that 69 eyes had referable DR. Of these 69 eyes, 12 eyes were confirmed to have referable DR, 54 did not have referable DR and 3 eyes were deemed with insufficient image quality by grader 2.

As can be seen in Figure 11, 38 patients were graded as referable AMD of which 37 were identified by the RetCAD software at a cut-off of 50. The 1 patient that was missed by the AI was confirmed to be referable AMD by the 2nd grader. Of the 565 patients graded as non-referable AMD, the AI software indicated that 151 patients had referable-AMD. Of these 151 patients, 31 patients were confirmed to have referable AMD, 115 did not have referable AMD and 5 patients were deemed with insufficient image quality by grader 2.

As can be seen in Figure 12, 25 patients were graded as referable DR of which 20 were identified by the RetCAD software at a cut-off of 50. The 5 patients that were missed by the AI were confirmed to be non-referable DR by the 2nd grader. Of the 578 patients graded as non-referable DR, the AI software indicated that 57 patients had referable-DR. Of these 57 patients, 12 patients were confirmed to have referable DR, 44 did not have referable DR and 1 patient was deemed with insufficient image quality according to grader 2.

DISCUSSION

The detection of DR and AMD are of great importance, because of their potential visual impairment properties and rising prevalence. As the technical possibilities are consistently improving worldwide, the implementing of AI in diagnostic approaches seems to be logical. The current study investigated the performance of the AI system RetCAD v1.3.1 in clinical use in terms of image quality and simultaneous screening and grading of the two most common retinal diseases DR and AMD using retinal photographs as compared with manual grading of the corresponding images. The previously trained software had been tested for effectiveness with public datasets before. In this study, RetCAD was tested for the first time in normal daily clinical practice on new non-pre-selected patients with actual retinal images and its performance in real clinical practice was evaluated.

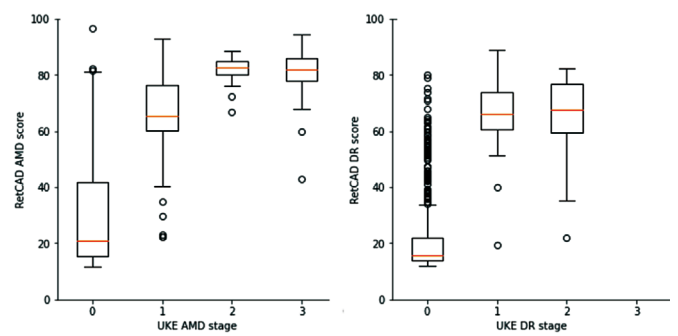


Figure 7 Box plot of RetCAD scores per age-related macular degeneration and diabetic retinopathy stage.

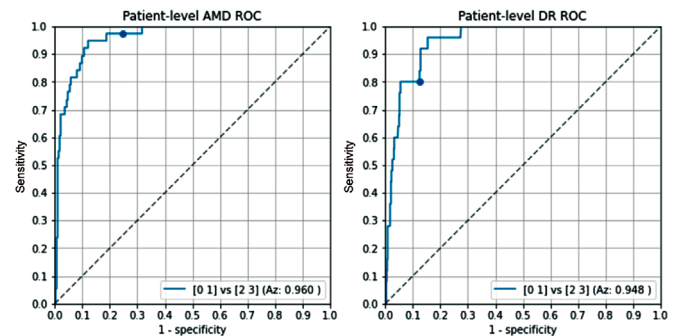


Figure 8 Receiver operating characteristic analysis on patient-level for referable age-related macular degeneration and referable diabetic retinopathy. The dots depict the operating point at a cut off of 50.

In order to achieve a high success rate in the screening or grading of clinical pictures, it is necessary to include only fundus images with sufficient image quality in the analysis. RetCAD evaluates the acquired fundus images with a corresponding quality score and can thus immediately exclude images of poor quality on site. Fundus images can be retaken immediately on site with the patient still present. A new appointment with the patient for a better image acquisition therefore is not necessary. Especially when the software is used in non-ophthalmic facilities, automatic differentiation between poor and sufficient image quality can be helpful for non-experts. In our results, the AI software had a sensitivity of 84.6% and specificity of 84.0%. Image quality scoring is a more subjective task as there are no clearly defined grading protocols for images quality. It was chosen to only remove images deemed bad quality by the reference observer as she was not able to grade them for AMD and/or DR. Images deemed bad quality only by the software were not removed to not decrease the number of subjects for testing. RetCAD does give an output for AMD and DR, regardless of image quality. Only the output might be less reliable when the image quality is deemed low. Of the 1245 photographed eyes, 1062 could be used resulting in 85.3% percent of the eyes.

The results for the automatic detection of AMD using the RetCAD AMD module were recorded at eye level and patient

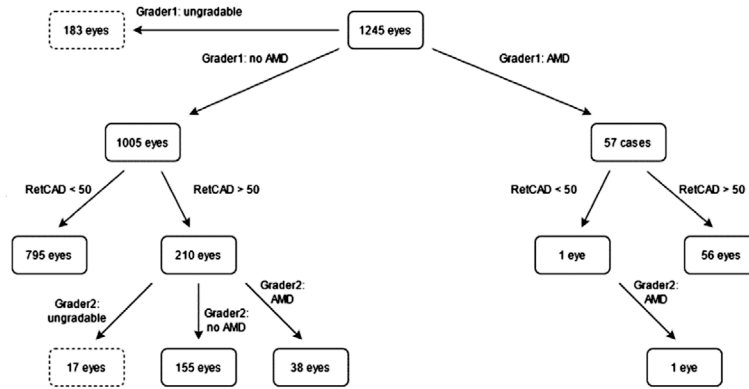


Figure 9 Age-related macular degeneration grading overview on eye-level.

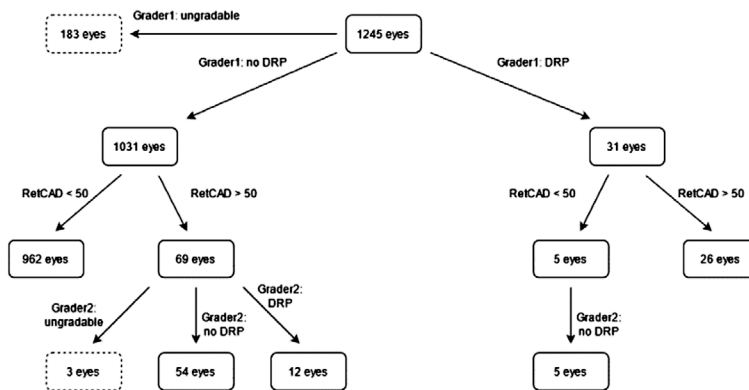


Figure 10 Diabetic retinopathy grading overview on eye-level.

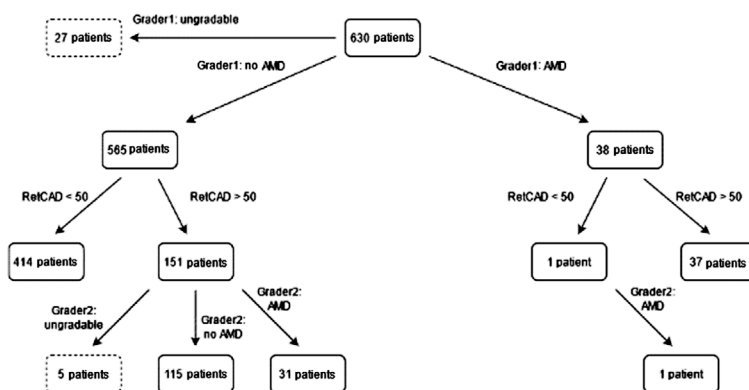


Figure 11 Age-related macular degeneration grading overview on patient-level.

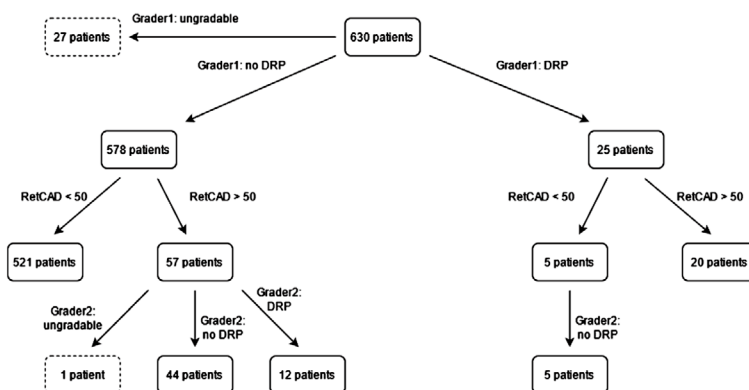


Figure 12 Diabetic retinopathy grading overview on patient-level.

level, with a general agreement with manual grading of 0.964 at eye-level slightly better than at patient-level with 0.960. The

sensitivity reaches a value of 98.2% and a specificity of 79.1% at eye-level. At patient-level the values are slightly lower with a

sensitivity of 97.3% and a specificity of 73.3%. The sensitivity and specificity most likely achieve a better value at the eye-level, since it can be determined more precisely due to the larger number of cases and the statistical significance is greater here. These high values indicate that over 98% of eyes with AMD were correctly classified as having the disease, which would result in a referral recommendation. The threshold of 50 at which a more detailed examination by a specialist is indicated was calibrated so that the specificity was slightly lower, with 79.1% of eyes correctly classified as healthy, but this minimized the risk of disease going undetected. The advantage of the AI output as a continuous value between 0 and 100 is that the cut-off threshold can be altered to put more emphasis on sensitivity or specificity. In a triage setting, one might prefer a high sensitivity to minimize the risk of the false negatives. In other settings, one might prefer a high specificity minimize the workload (*e.g.*, less false positive referrals) for retinal specialist who receive the referred patients.

Overall, only one patient with manually detected AMD requiring referral was not detected as such by the software, but 31 additional patients not noticed by the primary grader were automatically correctly diagnosed with AMD and approved by grader 2 as requiring referral. The results of RetCAD are therefore quite comparable to those of manual evaluation. When using a grader 2 adjudicated reference, a potential sensitivity of 98.6% (68/69 patients) and specificity of 78.3% (414/529 patients) could be achieved on patient-level.

With a performance of 91.6%, Burlina *et al*^[10] were able to achieve similar results for automated AMD grading using the DCNN AlexNet. However, these analyses were performed on public datasets and not on patients in a real clinic setting. In a retrospective work, Zapata *et al*^[13] achieved a sensitivity of 0.902 and a specificity of 0.825 with respect to automatic AMD classification. These values are based on a larger cohort but are still similar to RetCAD results. An average sensitivity of 88% and specificity of 90% could be shown in an analysis of automated AMD detection systems with a manual reference. Most of the involved studies determined the values on AREDS fundus photographs and not on patients in a realistic screening situation^[14]. RetCAD shows a higher sensitivity but a lower specificity using the selected threshold here.

The DR module of RetCAD v1.3.1. also achieves very good values at eye level in terms of Az (0.961) as well as sensitivity (83.9%) and specificity (93.3%). On patient-level, the Az was 0.948 with good sensitivity (80.0%) and specificity (90.1%). Grader 1 and grader 2 together identify 32 patients with DR. These are all identified by RetCAD resulting in a potential sensitivity of 100.0% (32/32 patients) and specificity of 92.3% (526/570 patients) when using a grader 2 adjudicated reference.

For the pioneer in the field of DR detection using AI and at the same time a comparable medical device, IDx-DR, a sensitivity of 97% and a specificity of 87% are reported for the public Messidor-2 datasets^[6]. In a clinical setting, values for specificity of 87.2% and sensitivity of 90.7% are obtained^[15]. In a recent project very similar to this study, IDx-DR achieved a sensitivity of 100% and a specificity of 82% when compared to three ophthalmologists^[16]. However, all patients were pre-selected with a diagnosis of diabetes prior to participation. With EyeArt v2.0, another program for automated DR diagnosis, a sensitivity of 91.3% and a specificity of 91.1% are achieved^[7]. In a recent study, sensitivity with the newer version EyeArt v2.1.0 is 95.7% for DR requiring referral and specificity is 54% for no DR and DR not requiring referral^[17]. None of these software diagnoses DR and AMD simultaneously.

One of the advantages of using RetCAD is its integration possibilities in clinical practice: either local or cloud-based, which makes it possible to evaluate the results regardless of location. These can be accessed online and offline as often as desired and used with any fundus camera. Cloud-based tele screening already provides promising prospects for future screening programs^[18]. González-Gonzalo *et al*^[19] also found correspondingly good functionality in RetCAD. Currently, RetCAD is one of the few available software packages that can simultaneously detect the two retinal diseases AMD and DR at the same time^[20]. Comparable AI systems are limited to the detection of just one disease^[7,13,21].

With this work, the AI software was evaluated in a real-world screening scenario without resorting to public databases, allowing patients to have an automated screening performed simultaneously in a short time during a routine visit. This real-world implementation is an important step in testing before such a medical device is used in everyday screening^[22-23]. Limitations of this work include the lack of late DR stages. Although these patients would likely not end up in routine eye screening, it could affect the results. Additionally, there was a lack of eyes with both DR and AMD in the same fundus. Moreover, the RetCAD quality score was not included in the further analysis. This can be justified by the fact that the number of images should not be limited further. Lastly, the secondary grader did not evaluate eyes and patients with matches between the reference grader and the AI, which could introduce a bias effect. However, the aim of this study was not to compare the two graders with each other, but to evaluate software performance against a predefined reference grade. Only in case of discrepancies was the secondary grader consulted for a decision. The above-mentioned limitations provide room for future research.

Additionally, reimbursement for such automated analysis systems plays an important role in the clinical uptake of

such systems in real clinical practice. With healthcare costs evermore increasing, next to the performance evaluation of these systems, also the cost-effectiveness of such systems should be evaluated.

This study confirmed that AI-based software systems can identify DR and AMD simultaneously within one screening. This software can be used in a real clinical setting to automatically detect abnormalities in fundus photographs with regard to AMD and DR disease during routine check-ups with a high degree of certainty, without the necessary need for a physician to be involved in the first step. The application is therefore a realistic resource-saving alternative to conventional AMD and DR screenings. A larger number of patients with more pronounced disease patterns are beneficial for future investigations to substantiate the obtained results. In conclusion, AI diagnostics for retinal diseases have the possibility to provide qualitative screenings to a broader population.

ACKNOWLEDGEMENTS

Foundation: Supported by the Horizon 2020 EUROSTARS programme [Eurostars programme: IMAGE-R (#12712)].

Conflicts of Interest: Skevas C, None; Weindler H, None; Levering M, None; Engelberts J, employee of Thirona; van Grinsven M, employee and shareholder of Thirona; Katz T, non-paid member of the Thirona advisory board.

REFERENCES

- Ambati J, Fowler BJ. Mechanisms of age-related macular degeneration. *Neuron* 2012;75(1):26-39.
- Wang W, Lo ACY. Diabetic retinopathy: pathophysiology and treatments. *Int J Mol Sci* 2018;19(6):1816.
- Guariguata L, Whiting DR, Hambleton I, Beagley J, Linnenkamp U, Shaw JE. Global estimates of diabetes prevalence for 2013 and projections for 2035. *Diabetes Res Clin Pract* 2014;103(2):137-149.
- Leasher JL, Bourne RR, Flaxman SR, Jonas JB, Keeffe J, Naidoo K, Pesudovs K, Price H, White RA, Wong TY, Resnikoff S, Taylor HR, Vision Loss Expert Group of the Global Burden of Disease Study. Global estimates on the number of people blind or visually impaired by diabetic retinopathy: a meta-analysis from 1990 to 2010. *Diabetes Care* 2016;39(9):1643-1649.
- Bhuiyan A, Wong TY, Ting DSW, Govindaiah A, Souied EH, Smith RT. Artificial intelligence to stratify severity of age-related macular degeneration (AMD) and predict risk of progression to late AMD. *Transl Vis Sci Technol* 2020;9(2):25.
- Abramoff MD, Lou YY, Erginay A, Clarida W, Amelon R, Folk JC, Niemeijer M. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Invest Ophthalmol Vis Sci* 2016;57(13):5200-5206.
- Bhaskaranand M, Ramachandra C, Bhat S, Cuadros J, Nittala MG, Sadda SR, Solanki K. The value of automated diabetic retinopathy screening with the EyeArt system: a study of more than 100, 000 consecutive encounters from people with diabetes. *Diabetes Technol Ther* 2019;21(11):635-643.
- Padhy SK, Takkar B, Chawla R, Kumar A. Artificial intelligence in diabetic retinopathy: a natural step to the future. *Indian J Ophthalmol* 2019;67(7):1004-1009.
- Tufail A, Rudisill C, Egan C, Kapetanakis VV, Salas-Vega S, Owen CG, Lee A, Louw V, Anderson J, Liew G, Bolter L, Srinivas S, Nittala M, Sadda S, Taylor P, Rudnicka AR. Automated diabetic retinopathy image assessment software: diagnostic accuracy and cost-effectiveness compared with human graders. *Ophthalmology* 2017;124(3):343-351.
- Burlina P, Pacheco KD, Joshi N, Freund DE, Bressler NM. Comparing humans and deep learning performance for grading AMD: a study in using universal deep features and transfer learning for automated AMD analysis. *Comput Biol Med* 2017;82:80-86.
- Grassmann F, Mengelkamp J, Brandl C, Harsch S, Zimmermann ME, Linkohr B, Peters A, Heid IM, Palm C, Weber BHF. A deep learning algorithm for prediction of age-related eye disease study severity scale for age-related macular degeneration from color fundus photography. *Ophthalmology* 2018;125(9):1410-1420.
- Avidor D, Loewenstein A, Waisbourd M, Nutman A. Cost-effectiveness of diabetic retinopathy screening programs using telemedicine: a systematic review. *Cost Eff Resour Alloc* 2020;18:16.
- Zapata MA, Royo-Fibla D, Font O, et al. Artificial intelligence to identify retinal fundus images, quality validation, laterality evaluation, macular degeneration, and suspected glaucoma. *Clin Ophthalmol* 2020;14:419-429.
- Dong L, Yang Q, Zhang RH, Wei WB. Artificial intelligence for the detection of age-related macular degeneration in color fundus photographs: a systematic review and meta-analysis. *EClinicalMedicine* 2021;35:100875.
- van der Heijden AA, Abramoff MD, Verbraak F, van Hecke MV, Liem A, Nijpels G. Validation of automated screening for referable diabetic retinopathy with the IDx-DR device in the Hoorn Diabetes Care System. *Acta Ophthalmol* 2018;96(1):63-68.
- Shah A, Clarida W, Amelon R, Hernaez-Ortega MC, Navea A, Morales-Olivas J, Dolz-Marco R, Verbraak F, Jorda PP, van der Heijden AA, Peris Martinez C. Validation of automated screening for referable diabetic retinopathy with an autonomous diagnostic artificial intelligence system in a Spanish population. *J Diabetes Sci Technol* 2021;15(3):655-663.
- 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(5):723-728.
- Obata R, Yoshinaga A, Yamamoto M, et al. Imaging of a retinal pigment epithelium aperture using polarization-sensitive optical coherence tomography. *Jpn J Ophthalmol* 2021;65(1):30-41.
- González-Gonzalo C, Sánchez-Gutiérrez V, Hernández-Martínez P, Contreras I, Lechanteur YT, Domanian A, van Ginneken B, Sánchez CI. Evaluation of a deep learning system for the joint automated detection of diabetic retinopathy and age-related macular degeneration. *Acta Ophthalmol* 2020;98(4):368-377.

- 20 Son J, Shin JY, Kim HD, Jung KH, Park KH, Park SJ. Development and validation of deep learning models for screening multiple abnormal findings in retinal fundus images. *Ophthalmology* 2020; 127(1):85-94.
- 21 Peng YF, Dharssi S, Chen QY, Keenan TD, Agrón E, Wong WT, Chew EY, Lu ZY. DeepSeeNet: a deep learning model for automated classification of patient-based age-related macular degeneration severity from color fundus photographs. *Ophthalmology* 2019;126(4):565-575.
- 22 Feng J, Emerson S, Simon N. Approval policies for modifications to machine learning-based software as a medical device: a study of bio-creep. *Biometrics* 2021;77(1):31-44.
- 23 Berensmann M, Gratzfeld M. Anforderungen für Die CE-kennzeichnung von apps und wearables. *Bundesgesundheitsbl* 2018; 61(3):314-320.