Artificial Intelligence Detection of Diabetic Retinopathy: Subgroup Comparison of the EyeArt System with Ophthalmologists' Dilated Exams

Jennifer I. Lim, MD, Carl D. Regillo, MD, SriniVas R. Sadda, MD, Eli Ipp, MD, Malavika Bhaskaranand, PhD, Chaithanya Ramachandra, PhD, Kaushal Solanki, PhD, for the EyeArt Study Subgroup

PII: S2666-9145(22)00117-8
DOI: https://doi.org/10.1016/j.xops.2022.100228
Reference: XOPS 100228

To appear in: Ophthalmology Science

Received Date: 3 January 2022
Revised Date: 30 August 2022
Accepted Date: 22 September 2022

Please cite this article as: RRH: AI detection of DR: EyeArt system vs. ophthalmologists' dilated exams

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2022 Published by Elsevier Inc. on behalf of American Academy of Ophthalmology.
Artificial Intelligence Detection of Diabetic Retinopathy: Subgroup Comparison of the EyeArt System with Ophthalmologists’ Dilated Exams

**Authors:** Jennifer I. Lim MD\(^1\), Carl D. Regillo MD\(^2\), SriniVas R. Sadda MD\(^3\), Eli Ipp MD\(^4\), Malavika Bhaskaranand PhD\(^5\), Chaithanya Ramachandra PhD\(^5\), Kaushal Solanki PhD\(^5\) for the EyeArt Study

Subgroup*  

1 University of Illinois University of Illinois at Chicago, Dept. of Ophthalmology, Illinois Eye and Ear Infirmary, Chicago, Illinois (jennylim@uic.edu)  
2 Wills Eye Hospital, Thomas Jefferson University, Philadelphia, PA (cregillo@gmail.com)  
3 Doheny Eye Center, Arcadia, CA (SSadda@doheny.org)  
4 The Lundquist Institute, Harbor-UCLA Medical Center, Torrance, California (ipp@lundquist.org)  
5 Eyenuk, Inc., Los Angeles, CA (malaviak, chait, solanki@eyenuk.com)

**Contact information of corresponding author**

Jennifer I Lim MD  
Marion H Schenk Chair in Ophthalmology  
Vice Chair for Diversity & Inclusion  
Professor of Ophthalmology  
Director of Retina Service  
1855 W. Taylor St  
Mail Code 648  
Chicago IL 60612  
312-413-0704 office phone  
312-413-7929 FAX  
jennylim@uic.edu

Portions of these data were presented at the 2019 Annual Meeting of the American Society of Retina Specialists.

**Financial support:** The study was funded in part by grants from the National Institutes of Health (NIH): 9SB1EY027241 (Eyenuk), 2R44EY026864 (Eyenuk) and EY01792 (Core Grant University of Illinois at Chicago), unrestricted grant from Research to Prevent Blindness (University of Illinois at Chicago) and in part by Eyenuk, Inc., Los Angeles, CA, USA.

**Disclosures of the authors:**

JIL: Consultant to Alcon, Eyenuk consultant (starting October 2020 but not during the trial)  
El: Research support only through The Lundquist Institute from California Community Foundation from Genentech Foundation, Eyenuk, Inc., and Norris Foundation.
SRS: Consultant: Optos, Centervue, Heidelberg Engineering, Allergan, Amgen, Roche/Genentech, Novartis, Merck, 4DMT, Oxurion; Research Instruments: Optos, Centervue, Heidelberg Engineering, Carl Zeiss Meditec, Topcon, Nidek; Grant Support: NIH STTR/SBIR grant (Eyenuk) - investigator for Doheny subcontract

MB, CR, KS: Employees of Eyenuk, Inc.

**Short title:** AI detection of DR: EyeArt system vs. ophthalmologists’ dilated exams

**Key Words:** Artificial intelligence, screening, diabetic retinopathy
Compared to the clinical reference standard, sensitivity of the EyeArt system (97%) for detecting more than mild DR was higher than the sensitivity of dilated ophthalmoscopy by either general ophthalmologists (21%) or retina specialists (60%).
Abstract: (350/350 words)

Objective: To compare general ophthalmologists, retina specialists, and the EyeArt Artificial Intelligence (AI) system to the clinical reference standard for detecting more than mild DR (mtmDR).

Design: Prospective, pivotal, multi-center trial conducted from April 2017 to May 2018.

Setting: Subgroup included participants enrolled at 10 US retina and non-retina centers.

Participants: Participants were ≥18 years, had diabetes mellitus, and underwent dilated ophthalmoscopy. 521 of 893 participants met these criteria and completed the study protocol.

Testing: Participants underwent 2-field fundus photography (macula centered, disc centered) for the EyeArt system, dilated ophthalmoscopy, and 4-wide field stereoscopic dilated fundus photography for reference standard grading.

Main Outcome Measures: For mtmDR detection, sensitivity and specificity of EyeArt gradings of 2-field, fundus photographs and ophthalmoscopy grading versus a rigorous clinical reference standard comprising reading center grading of 4-wide field stereoscopic dilated fundus photographs using the Early Treatment Diabetic Retinopathy Study (ETDRS) severity scale. The AI system provided automatic eye-level results regarding mtmDR.

Results: 521 participants (999 eyes) at 10 centers underwent dilated ophthalmoscopy: 406 by non-retina and 115 by retina specialists. Reading Center graded 207 positive and 792 eyes negative for mtmDR. Of these 999 eyes, 26 eyes were ungradable by the EyeArt system, leaving 973 eyes with both EyeArt and Reading Center gradings. Retina specialists correctly identified 22 of 37 eyes as positive (sensitivity 59.5%) and 182 of 184 eyes as negative (specificity 98.9%) for mtmDR versus the EyeArt AI system which identified 36 of 37 as positive (sensitivity 97%) and 162 of 184 eyes as negative (specificity of 88%) for mtmDR. General ophthalmologists correctly identified 35 of 170 eyes as positive (sensitivity 20.6%) and
607 of 608 eyes as negative (specificity 99.8%) for mtmDR compared with the EyeArt AI system which identified 164 of 170 as positive (sensitivity 96.5%) and 525 of 608 eyes as negative (specificity 86%) for mtmDR.

Conclusions: The AI system had higher sensitivity for detecting mtmDR than either general ophthalmologists or retina specialists as compared to the clinical reference standard. It can potentially serve as a low-cost point-of-care DR detection tool and help address the diabetic eye screening burden.

Trial Registration: ClinicalTrials.gov Identifier NCT03112005

Tweet: EyeArt AI system for detection of more than mild DR is more sensitive than ophthalmoscopy by even retina specialists.
Artificial intelligence (AI) systems using digital fundus photography instruments have been developed for diabetic retinopathy (DR) screening in order to partially address the increased demand for screening related to a burgeoning diabetic population in the world. According to the International Diabetes Federation report of 2019, there are currently about 463 million persons with diabetes mellitus and this number is expected to increase to 700 million by 2045, an increase of almost 51%. In the US, at the present time, best case scenario estimates are that half of diabetic patients undergo screening. Screening for detection of DR is crucial as earlier treatment of DR is linked to better outcomes. Other options to address the screening burden include telemedicine screening programs, point of care screening or traveling/mobile systems to provide access to patients. There are fixed costs associated with telescreening which include the cost of individuals to obtain photographs, costs to transmit photos to expert readers, and the time of the specialist reading the images. Use of traveling screening mobile units incurs the additional cost of transportation.

Advantages of AI screening systems include convenience of point-of-care access and the potentially lower operating cost as a result of automatic interpretation of the images and referral to an eye care specialist only as needed. Drawbacks of current telescreening and AI systems include the lack of an ability to detect and correct refractive errors, and lack of evaluation of the anterior segment and peripheral retina as would be performed during an in-person examination.

Two fully autonomous AI systems for detecting DR without human oversight have been cleared by the FDA. The IDx-DR system is an FDA-cleared AI point-of-care screening system with 87% sensitivity, 90% specificity and 96% imageability for more than mild DR (mtmDR). The EyeArt system (Eyenuk, Inc., Los Angeles, California) is also an FDA-cleared AI-based system that can enable point of care screening with 96% sensitivity, 88% specificity, and 97% imageability for detecting eyes with mtmDR. Unlike IDx-
DR, the EyeArt system can also detect eyes with vtDR with 97% sensitivity and 90% specificity and provide DR detection results at eye-level. Moreover, to generate disease detection results, the EyeArt system requires fewer (12.6%) patients to be dilated compared to the IDx-DR (23.6%).

The EyeArt system was developed using deep learning and uses multiple deep neural networks for specific classification tasks on images and the outcomes of these various networks are combined in a clinically aligned framework. It was trained on 375,000 images and then validated on 250,000 images. The AI analysis of the EyeArt system is cloud-based with a user interface that is installed on the user’s computer. The system requires two 45-degree field of view images, one centered on the optic disc and the second one centered on the macula, captured using a digital fundus camera. The images may be taken without dilation. The patient’s fundus photos are uploaded to the cloud and are interpreted by the AI system within 60 seconds for determination of mtmDR and vtDR. mtmDR is defined as the presence of moderate NPDR or higher stage or the presence of diabetic macular edema (DME), and vtDR is defined as the presence of severe NPDR or proliferative DR or the presence of DME. The patient disposition report generated for the patient visit contains mtmDR and vtDR determination for each eye. This report gives the follow-up recommendation as either referral to an ophthalmologist or return for screening in another year. The system is intended as a screening method to detect mtmDR and vtDR in eyes which have no prior treatment for DR.

In a prospective, pivotal study at 15 sites (primary care clinics, ophthalmology, retina practices), the EyeArt System was shown to have high sensitivity and specificity to enable point of care screening. In this pivotal study, the clinical reference standard was standardized, adjudicated grading of stereoscopic four field, 45 degrees digital fundus photographs on the Early Treatment Diabetic Retinopathy Study (ETDRS) scale, as assessed by the Wisconsin Reading Center. The ability of the EyeArt system to detect the presence of mtmDR was compared against the reading center determination. The study showed the AI system to have a sensitivity of 96%, specificity of 88% and
gradeability of 87.4% without dilation, increasing to 97.4% with dilation as needed. The EyeArt AI system compared favorably with the clinical reference standard of 4-field stereoscopic images and met the predetermined sensitivity and specificity endpoints for the detection of referable DR in diabetic individuals (p<0.0001), thus making the system suitable for point-of-care DR screening for triage and identification of patients requiring referral.

The purpose of this study was to compare general ophthalmologists, retina specialists and the EyeArt AI screening system to the clinical reference standard of reading center evaluation of fundus photographs for detection of mtmDR. Previous studies have compared dilated ophthalmoscopy against the ETDRS reference standard, AI systems against the ETDRS reference standard, AI systems against expert grading of non-stereo images, and AI systems against dilated ophthalmoscopy. To the best of our knowledge, this is the first study that evaluates and compares the performances of an AI system and dilated ophthalmoscopy against the ETDRS reference standard on the same large and diverse cohort of subjects enrolled at multiple centers with geographic diversity. Moreover, it also separately reports and compares the performance of dilated ophthalmoscopy by general ophthalmologists and retina specialists. Given that dilated ophthalmoscopy is the current standard of care for DR screening and that AI systems are alternatives that expand the reach of DR screening to primary care centers, this study comparing the two has significant clinical importance and impact.

Methods

This analysis focuses on the subset of data from 10 of the total 15 sites in the pivotal study (registered and publicly available at ClinicalTrials.gov, Identifier NCT03112005) where, a general ophthalmologist (i.e., non-retina practices) or a retina specialist (i.e., retina practices) participated in the trial to perform dilated eye exams. The other 5 sites did not have a participating general ophthalmologist or a retina specialist who could perform dilated eye exams. The exclusion of subjects
enrolled at these 5 sites due to lack of dilated ophthalmoscopy results does not introduce selection bias and does not alter the conclusions of this study. Institutional Review Board (IRB)/Ethics Committee approval was obtained and the study was conducted in adherence to the tenets of the Declaration of Helsinki. Following IRB approval and informed consent, patients were deemed eligible if they were diagnosed with diabetes mellitus, were 18 years of age or older and could tolerate fundus photography. Patients were ineligible if they had persistent visual impairment (defined by the American Foundation for the Blind as legal blindness or low vision where a person has difficulty accomplishing or cannot accomplish visual tasks even with prescribed corrective lenses), history of macular edema, treatment for any form of DR (intravitreal injections, laser photocoagulation, intraocular surgery other than for cataracts), or known retinal disease (vascular occlusion, retinal detachment). Enrollment criteria did not require knowledge of the patient’s ophthalmic history including DR diagnosis.

Study participants underwent non-mydriatic imaging using the EyeArt system, followed by dilation and both dilated ophthalmoscopy and fundus photography. Trained ophthalmic photographers obtained two 45-degrees, non-mydriatic color fundus digital images of the right eye followed by the left eye using a digital fundus camera with resolution of 1.69 megapixels or higher (Canon CR-2 AF or CR-2 Plus AF). For each eye, one image was centered on the optic disc and the other was centered on the macula. These non-mydriatic fundus photos were uploaded to the cloud, where the EyeArt AI system determined whether mtmDR and vtDR (as defined earlier) were present. If the image quality precluded an interpretation of level of DR by the AI system, operators received immediate feedback from the EyeArt system, and the eye underwent dilation followed by repeat imaging and upload for automated analysis. Otherwise, after the completion of non-mydriatic photos and dilation, the study participant proceeded to the remainder of the study procedures.

Once pupillary dilation was achieved, the participant underwent dilated ophthalmoscopy (slit lamp and indirect ophthalmoscopy) by the ophthalmologist (general or retina specialist), who graded
each eye for both presence or absence of clinically significant DME (CSDME) as per the ETDRS and
presence of any DR, non-proliferative (graded as mild, moderate or severe) DR (NPDR) or proliferative
DR (PDR). The ophthalmologist had access to the participant’s ophthalmic history when available. Four
mydriatic 45-degrees field of view stereoscopic images (nasal retina including optic disc, centered on
macula, supratemporal to disc and inferotemporal to disc) were then taken of each eye. These four 45-
degrees field of view images are equivalent to the seven-field ETDRS 30-degrees images. Images were
sent to the Wisconsin Fundus Photograph Reading Center (WFPRC) for ETDRS grading of the presence or
absence of DME as well as for ETDRS level grading of the DR. At the WFPRC, two independent certified
graders masked to the EyeArt results examined the images using standardized procedures to establish
the clinical reference standard. Between-grader differences exceeding prespecified criteria were
adjudicated by a third, more senior grader.

For ophthalmoscopy performance analyses, eyes were included if: (1) the color fundus
photographs were evaluable by the reading center, (2) a dilated examination with a DR grading was
performed by an ophthalmologist and (3) an EyeArt analysis was obtained. For EyeArt performance
analyses, eyes that were ungradable per the EyeArt system were considered equivalent to mtmDR
positive, since in regular clinical use of the EyeArt system, patients with mtmDR positive results or
ungradable results would be referred to an eye care provider for further evaluation. The WFPRC
gradings of the 4-wide field mydriatic photographs were used as the clinical reference standard for
determination of sensitivity and specificity when comparing ophthalmoscopy and the AI EyeArt DR
gradings. False negative rates were determined for each method.

Sensitivity was defined as the accuracy among positive detections, calculated as the ratio of positive
tests to reference standard positives. Specificity was defined as the accuracy among negative detections,
calculated as the ratio of negative tests to reference standard negatives. Imageability/gradability was
defined as the percentage of eyes that received a disease detection result (positive or negative) among
all eyes determined gradable by the WFPRC. For each of these performance measures, 95% confidence interval (CI) calculations were performed to account for the correlation between eyes of the same patient using methods presented by Kang and Lee.\

**Results**

Of the 893 participants who enrolled in the clinical trial and completed study procedures, 521 participants also underwent dilated ophthalmoscopy. Of the 521 participants, 221 eyes (from 115 participants) evaluated at retina practices and 778 eyes (from 405 participants) evaluated at general ophthalmology (non-retina) practices had evaluable clinical reference standard photos and 43 eyes (including 2 eyes from 1 participant) did not have evaluable clinical reference standard photos. One participant’s clinical reference standard images for each eye were deemed not evaluable by the Reading Center and therefore this participant could not be included in the analyses, which require comparisons to the reference standard images.

The demographic characteristics of these subjects are presented in Table 1. Of these 999 total eyes, 792 were read by the reading center as mtmDR negative and 207 were mtmDR positive. Of these total 999 eyes, 26 eyes were ungradable by the EyeArt System, leaving 973 eyes that had both EyeArt and Reading Center gradings. The reading center gradings are shown on Table 2. Of the 973 eyes, 669 had no apparent DR, 108 had mild NPDR, 183 had moderate NPDR, 1 had severe NPDR, 11 had PDR and 1 had questionable DR (ETDRS level 14 or 15) per the Reading Center gradings. There were therefore a total of 196 eyes with mtmDR.

EyeArt and dilated ophthalmologist examination gradings were compared to reading center gradings. The breakdown of the EyeArt and dilated ophthalmoscopy results is shown in Figure 1 and the sensitivity, specificity, and imageability/gradability are reported in Table 3.
Overall, sensitivity for detection of mtmDR was higher for the EyeArt System than for dilated ophthalmoscopy: 96.4% (95% CI, 93.1% - 99.8%) EyeArt versus 27.7% (95% CI, 20.1%-35.2%) for ophthalmoscopy. Specificity was lower for the EyeArt system than for ophthalmoscopy: 88.4% (95% CI, 85.8% - 91.1%) EyeArt versus 99.6% (95% CI, 99.1% - 100.0%) dilated ophthalmoscopy.

Evaluation of the dilated ophthalmoscopy results showed the retina specialists to have higher sensitivity than the general ophthalmologists for detection of mtmDR: 59.5% sensitivity retina specialists versus 20.7% for general ophthalmologists. The incidental demographic differences among the participants seen by the retina specialists and general ophthalmologists may have an impact on the performance differences in these sub-cohorts. Retina specialists correctly identified 22 of 37 eyes positive for mtmDR resulting in a sensitivity of 59.5% and 182 of 184 eyes negative for mtmDR resulting in a specificity of 98.9%. For this cohort, the EyeArt AI system correctly referred 36 of 37 eyes resulting in a sensitivity of 97% and correctly identified 162 of 184 eyes as negative for mtmDR resulting in a specificity of 88%. General ophthalmologists correctly identified 35 of 170 eyes as positive for mtmDR resulting in sensitivity of 20.7%, and correctly identified 607 of 608 eyes as negative for mtmDR, resulting in a specificity of 99.8%. For this cohort, the EyeArt AI system correctly referred 164 of 170 eyes resulting in sensitivity of 96.5% and correctly identified 525 of 608 eyes as negative for mtmDR, resulting in a specificity of 86%.

Among the 207 reference standard mtmDR positives, there were 18 false negatives with the EyeArt AI system. Of these 18 eyes that were not identified as mtmDR positive the EyeArt AI system, read 7 as negative and 11 as ungradable. All 7 reference standard positives read as negative by EyeArt were graded by the Reading Center as ETDRS level 35 without CSDME (Figure 2A, B). The corresponding widefield photos show retinal hemorrhages that were not detected by the EyeArt images (Figure 3). In each case, only one of the four widefield images was positive for DR. The larger hemorrhage (Figure 3A) was in a retinal location beyond the 2-field area imaged by the EyeArt system.
Of the 11 reference standard positives read as ungradable by EyeArt, 8 eyes were read by the Reading Center as ETDRS level 35 (12 eyes), 2 as level 43 and one as level 47; none had CSDME. Thus, none were positive for vtDR as per the Reading Center. There were 149 false negatives as read by the ophthalmologists. Of these 149 false negatives, 15/149 (10%) were evaluated by retina specialists and the remainder (134/149, 90%) by general ophthalmologists. None of these retina specialists' false negatives were read by the Reading Center as vtDR. In contrast, for general ophthalmologists, 26/134 (19%) eyes had vtDR. These included 5 eyes with PDR, 7 with level 47 and CSDME, 3 with level 43 and CSDME, 10 with level 35 and CSDME and 1 with level 14 and CSDME. Further analysis of all of the false negatives showed that 2 centers accounted for 90 of 134 false negative (67%) with 21 of the 90 (23%) having vtDR. Thus, for the other 5 non-retina centers, the rate of vtDR among the false negatives was 5 of 44 (11%).

The false positive rate was higher for the EyeArt AI System (90/792, 11%) than for the ophthalmologists (3/792, 0.3%). The EyeArt AI false positives were comprised mostly of eyes with mild NPDR (44 eyes) and 17 other diagnoses including drusen, nevi, epiretinal membrane, retinal vein occlusion, geographic atrophy, asteroid hyalosis, optic nerve swelling, peripapillary atrophy, uveitis, RPE disturbance, and fibrous tissue (Figure 2 C-F). For the ophthalmoscopy, the false positive diagnoses per reading center were no apparent DR (2 eyes) for retina specialists and chorioretinal scars for general ophthalmologists (1 eye).

Discussion

In a point of care screening prospective study, the EyeArt AI system's sensitivity for detection of mtmDR was much higher than either a general ophthalmologist or retina specialist as compared to the clinical reference standard of grading of fundus photos by a central reading center. Furthermore, the AI system, similar to the retina specialist group, did not miss any cases of vtDR, in contrast to the general
ophthalmologist group. Although the AI system had a higher rate of false positives than either a general ophthalmologist or retina specialist, these were, in general, eyes with mild NPDR or had other ocular pathology that in the real world, would benefit from being referred and evaluated by an ophthalmologist. The demographic diversity, geographic diversity and the wide spectrum of DR severity included in this study population support the generalizability of this study’s results to the intended DR screening population.

The significant discrepancy between the clinical reference standard grading and the retina specialists, is due mainly to undergrading of moderate NPDR as mild NPDR. None of the eyes had severe NPDR or higher. It is well known that color fundus photos result in higher DR grades as compared to ophthalmoscopic examination. Scanlon et al\textsuperscript{11} explain that the difference in referable retinopathy detection between ophthalmoscopy and seven field stereophotography was primarily due to difficulty in distinguishing hemorrhages and microaneurysms. ETDRS grading based on four-wide field (or seven field) stereo photographs is considered the clinical reference standard for DR diagnosis including by the FDA for this study and other comparison studies. An AI screening-based system is ideally one that has a high sensitivity in order not to miss significant pathology that could cause visual loss. Conversely, an ideal AI screening system would have a lower specificity in the trade-off between sensitivity and specificity when screening for potentially visual threatening pathology. The EyeArt system fulfills these requirements and is a potentially useful screening tool with its high sensitivity for detecting mtmDR.

Compared to dilated ophthalmoscopy for DR screening, advantages include the low cost and the convenient location as a point of care tool. Unlike human based telescreening systems, the EyeArt system provides an immediate determination of the presence of mtmDR that is available to the patient prior to leaving the primary care office, which has been shown to improve adherence to follow up care\textsuperscript{16}. A referral can then be made, and the importance of follow-up can be stressed to the patient.
Disadvantages include the somewhat lower specificity which may lead to over-referral, although a significant portion of false positives (i.e., over-referrals) had mild NPDR or other ocular pathology that would benefit from evaluation by an ophthalmologist. In addition, patients who do not have mtmDR may not be referred for an ophthalmic examination that could screen for other ocular diseases. Other ocular conditions such as refractive errors, glaucoma and cataract are not addressed in these individuals. Thus, use of the system should be tempered with the advice to seek a general ophthalmic examination if the patient has blurred vision or if the patient is in a high-risk group for conditions such as glaucoma.

Given the current low rate of compliance with the recommendation for an annual diabetic retina examination, this system can be a useful adjunct in the detection of mtmDR and appears to be more accurate than clinical ophthalmoscopy for routine retinal screening.
Acknowledgements: The authors would like to acknowledge the contributions of the EyeArt Study Subgroup (Harvey Dubiner MD, Grace Levy-Clarke MD, Richard Pesavento MD, Mark D. Sherman MD, Steven Silverstein MD, Brian Kim MD, Gerald B. Walman MD), the Wisconsin Fundus Photography Reading Center (Barbara A. Blodi, Amitha Domalpally, Susan Reed, James Reimers, Kris Lang, Holy Cohn, Ruth Shaw, Sheila Watson, Andrew Ewen, Nancy Barrett, and Maria Swift), and Jeffrey Gornbein.

Author Contributions:

Concept and Design: Malavika Bhaskaranand, Chaithanya Ramachandra, Kaushal Solanki
Acquisition, analysis, or interpretation of data: All authors
Drafting of the manuscript: All authors
Critical revision of the manuscript for important intellectual content: All authors
Statistical analysis: Malavika Bhaskaranand, Chaithanya Ramachandra
Obtained funding: Malavika Bhaskaranand, Chaithanya Ramachandra, Kaushal Solanki
Supervision: Malavika Bhaskaranand, Chaithanya Ramachandra, Kaushal Solanki

Jennifer I. Lim and Malavika Bhaskaranand had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
References


   https://www.afb.org/research-and-initiatives/statistics/key-definitions-statistical-terms


**Figure Legends**

1. STARD chart of participants with dilated ophthalmoscopy (n=999 eyes) showing the number of eyes that were positive, negative, or ungradable for mtmDR per grading by the Reading Center, EyeArt system and ophthalmoscopy by general ophthalmologists and retina specialists.

Intent to screen (ITS) cohort with clinical reference standard and EyeArt analysis

719: dilated ophthalmoscopy not done

Reading center mtmDR negative
792

Reading center mtmDR positive
207

mtmDR negative
687

mtmDR positive
90

Ungradable
15

mtmDR negative
7

mtmDR positive
189

Ungradable
11

Ophthalmoscopy results

mtmDR negative
805

mtmDR positive
2

mtmDR negative
90

mtmDR positive
0

mtmDR negative
14

mtmDR positive
1

mtmDR negative
7

mtmDR positive
137

mtmDR negative
55

Ungradable
1

mtmDR negative
5

mtmDR positive
6
2. 2-field (EyeArt) images for eyes that were read as false positive or false negative per the EyeArt system compared to the grading by the Reading Center. A. EyeArt 2-field images read as negative for more than diabetic retinopathy (mtmDR) whereas the Reading Center 4-widefield images were graded as ETDRS 35C. No hemorrhages are visible on the images shown. B. EyeArt 2-field images read as negative for mtmDR whereas the Reading Center 4-widefield images were graded ETDRS 35C. Note there is a small hemorrhage just inferior to the superior arcade arteriole. C. EyeArt 2-field images read as vision threatening diabetic retinopathy (vtDR) whereas the Reading Center diagnosed central vein occlusion. D. EyeArt 2-field images read as vtDR whereas the Reading Center diagnosed swollen optic discs of unknown cause. E. EyeArt 2-field images read as mtmDR) whereas the Reading Center diagnosed an epiretinal membrane. Note the striae present in the macular area.

3. Color widefield 45-degree photographs corresponding to the representative 2-field EyeArt false negatives for more than mild DR (mtmDR) shown in Figure 2. A. This inferior field is the only one of the four widefield images that revealed any DR. This image shows the presence of 2 small retinal hemorrhages and one larger hemorrhage. The larger hemorrhage was not in an area imaged by the EyeArt system. B. This widefield photograph which reveals a small hemorrhage located just inferior to the superior arcade artery. This hemorrhage was visible on the 2-field EyeArt image but was read falsely negative for mtmDR. The other 3 widefield images of this eye did not show any retinopathy.
Table 1. Demographic characteristics for subjects with dilated ophthalmoscopy exams at non-retina sites (N=406), retina sites (N=115), and total (N=521).

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Total n/N (%)</th>
<th>Non-retina sites n/N (%)</th>
<th>Retina sites n/N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n/N</td>
<td>n/N (%)</td>
<td>n/N (%)</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;65 years</td>
<td>370/521</td>
<td>(71.0%)</td>
<td>291/406 (71.7%)</td>
</tr>
<tr>
<td>≥65 years</td>
<td>151/521</td>
<td>(29.0%)</td>
<td>115/406 (28.3%)</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>244/521</td>
<td>(46.8%)</td>
<td>176/406 (43.3%)</td>
</tr>
<tr>
<td>Female</td>
<td>277/521</td>
<td>(53.2%)</td>
<td>230/406 (56.7%)</td>
</tr>
<tr>
<td><strong>Ethnicity\textsuperscript{a}</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>169/521</td>
<td>(32.4%)</td>
<td>149/406 (36.7%)</td>
</tr>
<tr>
<td>Non-Hispanic/Latino</td>
<td>352/521</td>
<td>(67.6%)</td>
<td>257/406 (63.3%)</td>
</tr>
<tr>
<td><strong>Race\textsuperscript{a}</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>2/521 (0.4%)</td>
<td>2/406 (0.5%)</td>
<td>0/115 (0.0%)</td>
</tr>
<tr>
<td>Asian</td>
<td>18/521</td>
<td>(3.5%)</td>
<td>14/406 (3.4%)</td>
</tr>
<tr>
<td>Black or African American</td>
<td>105/521</td>
<td>(20.2%)</td>
<td>58/406 (14.3%)</td>
</tr>
<tr>
<td>Native Hawaiian or other Pacific Islander</td>
<td>3/521 (0.6%)</td>
<td>3/406 (0.7%)</td>
<td>0/115 (0.0%)</td>
</tr>
<tr>
<td>Other</td>
<td>44/521</td>
<td>(8.4%)</td>
<td>24/406 (5.9%)</td>
</tr>
<tr>
<td>White</td>
<td>349/521</td>
<td>(67.0%)</td>
<td>305/406 (75.1%)</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Race and ethnicity were self-reported.
Table 2: ICDR and CSDME levels of 973 analyzable eyes (from STARD chart in Figure 1) of participants with dilated ophthalmoscopy. The green shading indicates cases that are considered mtmDR positive per the reading center.

<table>
<thead>
<tr>
<th>DR level</th>
<th>No apparent DR</th>
<th>Mild NPDR</th>
<th>Moderate NPDR</th>
<th>Severe NPDR</th>
<th>PDR</th>
<th>Questionable</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>669</td>
<td>108</td>
<td>122</td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>908</td>
</tr>
<tr>
<td>Present</td>
<td>0</td>
<td>0</td>
<td>51</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>54</td>
</tr>
<tr>
<td>Questionable</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>669</strong></td>
<td><strong>108</strong></td>
<td><strong>183</strong></td>
<td><strong>1</strong></td>
<td><strong>11</strong></td>
<td><strong>1</strong></td>
<td><strong>973</strong></td>
</tr>
</tbody>
</table>
Table 3. Performance characteristics (sensitivity, specificity, imageability/gradability) for subject eyes with dilated ophthalmoscopy exams at non-retina sites (N=778), retina sites (N=221), and total (N=999).

<table>
<thead>
<tr>
<th>Cohort (Number of eyes with CRS)</th>
<th>Screening methodology</th>
<th>Sensitivity [95% CI]</th>
<th>Specificity [95% CI]</th>
<th>Imageability/gradability [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects with dilated ophthalmoscopy (n=999 eyes)</td>
<td>Dilated ophthalmoscopy</td>
<td>27.7% [20.1%-35.2%]</td>
<td>99.6% [99.1% - 100.0%]</td>
<td>99.9% [99.7% - 100.0%]</td>
</tr>
<tr>
<td></td>
<td>EyeArt system (with dilation-if-needed) (147 eyes dilated)</td>
<td>96.4% [93.1% - 99.8%]</td>
<td>88.4% [85.8% - 91.1%]</td>
<td>97.4% [96.0% - 98.8%]</td>
</tr>
<tr>
<td>Subjects enrolled at non-retina specialty centers (n=778 eyes)</td>
<td>Dilated ophthalmoscopy</td>
<td>20.6% [13.1% - 28.0%]</td>
<td>99.8% [99.5% - 100.0%]</td>
<td>99.9% [99.6% - 100.0%]</td>
</tr>
<tr>
<td></td>
<td>EyeArt system (with dilation-if-needed) (109 eyes dilated)</td>
<td>96.5% [92.9% - 100.0%]</td>
<td>86.3% [83.0% - 89.7%]</td>
<td>97.6% [96.0% - 99.1%]</td>
</tr>
<tr>
<td>Subjects enrolled at retina specialty centers (n=221 eyes)</td>
<td>Dilated ophthalmoscopy</td>
<td>59.5% [40.2% - 78.7%]</td>
<td>98.9% [97.1% - 100.0%]</td>
<td>100.0% [100.0% - 100.0%]</td>
</tr>
<tr>
<td></td>
<td>EyeArt system (with dilation-if-needed) (38 eyes dilated)</td>
<td>97.3% [90.3% - 100.0%]</td>
<td>88.0% [82.5% - 93.5%]</td>
<td>96.8% [93.6% - 100.0%]</td>
</tr>
</tbody>
</table>
Table 1. Demographic characteristics for subjects with dilated ophthalmoscopy exams at non-retina sites (N=406), retina sites (N=115), and total (N=521).

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Total n/N (%)</th>
<th>Non-retina sites n/N (%)</th>
<th>Retina sites n/N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n/N (%)</td>
<td>n/N (%)</td>
<td>n/N (%)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;65 years</td>
<td>370/521 (71.0%)</td>
<td>291/406 (71.7%)</td>
<td>79/115 (68.7%)</td>
</tr>
<tr>
<td>≥65 years</td>
<td>151/521 (29.0%)</td>
<td>115/406 (28.3%)</td>
<td>36/115 (31.3%)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>244/521 (46.8%)</td>
<td>176/406 (43.3%)</td>
<td>68/115 (59.1%)</td>
</tr>
<tr>
<td>Female</td>
<td>277/521 (53.2%)</td>
<td>230/406 (56.7%)</td>
<td>47/115 (40.9%)</td>
</tr>
<tr>
<td>Ethnicity\a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>169/521 (32.4%)</td>
<td>149/406 (36.7%)</td>
<td>20/115 (17.4%)</td>
</tr>
<tr>
<td>Non-Hispanic/Latino</td>
<td>352/521 (67.6%)</td>
<td>257/406 (63.3%)</td>
<td>95/115 (82.6%)</td>
</tr>
<tr>
<td>Race\a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>2/521 (0.4%)</td>
<td>2/406 (0.5%)</td>
<td>0/115 (0.0%)</td>
</tr>
<tr>
<td>Asian</td>
<td>18/521 (3.5%)</td>
<td>14/406 (3.4%)</td>
<td>4/115 (3.5%)</td>
</tr>
<tr>
<td>Black or African American</td>
<td>105/521 (20.2%)</td>
<td>58/406 (14.3%)</td>
<td>47/115 (40.9%)</td>
</tr>
<tr>
<td>Native Hawaiian or other Pacific Islander</td>
<td>3/521 (0.6%)</td>
<td>3/406 (0.7%)</td>
<td>0/115 (0.0%)</td>
</tr>
<tr>
<td>Other</td>
<td>44/521 (8.4%)</td>
<td>24/406 (5.9%)</td>
<td>20/115 (17.4%)</td>
</tr>
<tr>
<td>White</td>
<td>349/521 (67.0%)</td>
<td>305/406 (75.1%)</td>
<td>44/115 (38.3%)</td>
</tr>
</tbody>
</table>

\a Race and ethnicity were self-reported.
Table 2: ICDR and CSDME levels of 973 analyzable eyes (from STARD chart in Figure 1) of participants with dilated ophthalmoscopy. The green shading indicates cases that are considered mtmDR positive per the reading center.

<table>
<thead>
<tr>
<th>CSDME</th>
<th>No apparent DR</th>
<th>Mild NPDR</th>
<th>Moderate NPDR</th>
<th>Severe NPDR</th>
<th>PDR</th>
<th>Questionable</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>669</td>
<td>108</td>
<td>122</td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>908</td>
</tr>
<tr>
<td>Present</td>
<td>0</td>
<td>0</td>
<td>51</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>54</td>
</tr>
<tr>
<td>Questionable</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>669</td>
<td>108</td>
<td>183</td>
<td>1</td>
<td>11</td>
<td>1</td>
<td>973</td>
</tr>
</tbody>
</table>
Table 3. Performance characteristics (sensitivity, specificity, imageability/gradability) for subject eyes with dilated ophthalmoscopy exams at non-retina sites (N=778), retina sites (N=221), and total (N=999).

<table>
<thead>
<tr>
<th>Cohort (Number of eyes with CRS)</th>
<th>Screening methodology</th>
<th>Sensitivity [95% CI]</th>
<th>Specificity [95% CI]</th>
<th>Imageability/gradability [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects with dilated ophthalmoscopy (n=999 eyes)</td>
<td>Dilated ophthalmoscopy</td>
<td>27.7% [20.1%-35.2%]</td>
<td>99.6% [99.1%-100.0%]</td>
<td>99.9% [99.7%-100.0%]</td>
</tr>
<tr>
<td></td>
<td>EyeArt system (with dilation-if-needed) (147 eyes dilated)</td>
<td>96.4% [93.1%-99.8%]</td>
<td>88.4% [85.8%-91.1%]</td>
<td>97.4% [96.0%-98.8%]</td>
</tr>
<tr>
<td>Subjects enrolled at non-retina specialty centers (n=778 eyes)</td>
<td>Dilated ophthalmoscopy</td>
<td>20.6% [13.1%-28.0%]</td>
<td>99.8% [99.5%-100.0%]</td>
<td>99.9% [99.6%-100.0%]</td>
</tr>
<tr>
<td></td>
<td>EyeArt system (with dilation-if-needed) (109 eyes dilated)</td>
<td>96.5% [92.9%-100.0%]</td>
<td>86.3% [83.0%-89.7%]</td>
<td>97.6% [96.0%-99.1%]</td>
</tr>
<tr>
<td>Subjects enrolled at retina specialty centers (n=221 eyes)</td>
<td>Dilated ophthalmoscopy</td>
<td>59.5% [40.2%-78.7%]</td>
<td>98.9% [97.1%-100.0%]</td>
<td>100.0% [100.0%-100.0%]</td>
</tr>
<tr>
<td></td>
<td>EyeArt system (with dilation-if-needed) (38 eyes dilated)</td>
<td>97.3% [90.3%-100.0%]</td>
<td>88.0% [82.5%-93.5%]</td>
<td>96.8% [93.6%-100.0%]</td>
</tr>
</tbody>
</table>
Compared to the clinical reference standard, sensitivity of the EyeArt system (97%) for detecting more than mild DR was higher than the sensitivity of dilated ophthalmoscopy by either general ophthalmologists (21%) or retina specialists (60%).