



# The SEE Study: Safety, Efficacy, and Equity of Implementing Autonomous Artificial Intelligence for Diagnosing Diabetic Retinopathy in Youth

Risa M. Wolf,<sup>1</sup> T.Y. Alvin Liu,<sup>2</sup>  
 Chrystal Thomas,<sup>1</sup> Laura Prichett,<sup>3</sup>  
 Ingrid Zimmer-Galler,<sup>2</sup> Kerry Smith,<sup>2</sup>  
 Michael D. Abramoff,<sup>4,5,6,7,8</sup> and  
 Roomasa Channa<sup>2,9</sup>

*Diabetes Care* 2021;44:781–787 | <https://doi.org/10.2337/dc20-1671>

## OBJECTIVE

Diabetic retinopathy (DR) is a leading cause of vision loss worldwide. Screening for DR is recommended in children and adolescents, but adherence is poor. Recently, autonomous artificial intelligence (AI) systems have been developed for early detection of DR and have been included in the American Diabetes Association's guidelines for screening in adults. We sought to determine the diagnostic efficacy of autonomous AI for the diabetic eye exam in youth with diabetes.

## RESEARCH DESIGN AND METHODS

In this prospective study, point-of-care diabetic eye exam was implemented using a nonmydriatic fundus camera with an autonomous AI system for detection of DR in a multidisciplinary pediatric diabetes center. Sensitivity, specificity, and diagnosability of AI was compared with consensus grading by retinal specialists, who were masked to AI output. Adherence to screening guidelines was measured before and after AI implementation.

## RESULTS

Three hundred ten youth with diabetes aged 5–21 years were included, of whom 4.2% had DR. Diagnosability of AI was 97.5% (302 of 310). The sensitivity and specificity of AI to detect more-than-mild DR was 85.7% (95% CI 42.1–99.6%) and 79.3% (74.3–83.8%), respectively, compared with the reference standard as defined by retina specialists. Adherence improved from 49% to 95% after AI implementation.

## CONCLUSIONS

Use of a nonmydriatic fundus camera with autonomous AI was safe and effective for the diabetic eye exam in youth in our study. Adherence to screening guidelines improved with AI implementation. As the prevalence of diabetes increases in youth and adherence to screening guidelines remains suboptimal, effective strategies for diabetic eye exams in this population are needed.

Diabetes is a significant public health problem worldwide, and in the U.S., it affects almost 30 million people (1). In youth, diabetes is one of the most common chronic childhood diseases, with an incidence that has been increasing over the past decade for both type 1 diabetes (T1D) and type 2 diabetes (T2D) (2). Children with diabetes are at risk for diabetes-related complications, including diabetic retinopathy (DR), which

<sup>1</sup>Division of Endocrinology, Department of Pediatrics, Johns Hopkins School of Medicine, Baltimore, MD

<sup>2</sup>Wilmer Eye Institute, Johns Hopkins School of Medicine, Baltimore, MD

<sup>3</sup>Biostatistics, Epidemiology, and Data Management Core, Johns Hopkins School of Medicine, Baltimore, MD

<sup>4</sup>Department of Ophthalmology and Visual Sciences, The University of Iowa, Iowa City, IA

<sup>5</sup>IDx, Coralville, IA

<sup>6</sup>Iowa City VA Medical Center, Iowa City, IA

<sup>7</sup>Department of Biomedical Engineering, The University of Iowa, Iowa City, IA

<sup>8</sup>Department of Electrical and Computer Engineering, The University of Iowa, Iowa City, IA

<sup>9</sup>Department of Ophthalmology and Visual Sciences, University of Wisconsin, Madison, WI

Corresponding author: Roomasa Channa, [rchanna@wisc.edu](mailto:rchanna@wisc.edu)

Received 5 July 2020 and accepted 23 December 2020

This article contains supplementary material online at <https://doi.org/10.2337/figshare.13484976>.

This article is featured in a podcast available at <https://www.diabetesjournals.org/content/diabetes-core-update-podcasts>.

© 2021 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <https://www.diabetesjournals.org/content/license>.

is a leading cause of blindness and vision loss in young adults (3). The risk for development and progression of retinopathy can be partially mitigated by intensive glycemic control (4). However, only 17% of pediatric patients achieve the American Diabetes Association (ADA)-recommended HbA<sub>1c</sub> target of <7.5% for optimal glycemic control (5). Thus, evaluating for complications, and particularly DR, is of utmost importance to promote early detection and intervention to prevent vision loss.

The ADA recommends screening for DR in patients with T1D within 3–5 years of diagnosis and at the time of diagnosis for patients with T2D, with yearly follow-up exams thereafter for both types of diabetes (6). Studies show that only 35–72% of youth with diabetes undergo recommended ophthalmic exams in accordance with clinical practice guidelines. Furthermore, minority youth and children from lower socioeconomic backgrounds are less likely to undergo recommended screening compared with their White counterparts, even with insurance coverage (7). This may be due to the length of time associated with ophthalmic examinations and the additional time off from work and school required to attend ophthalmology visits in addition to quarterly diabetes care visits. While the current prevalence of DR in youth is reportedly low, this is likely to increase with the rising incidence of pediatric diabetes, especially T2D, which is associated with an earlier onset of DR (2,8,9).

To improve accessibility and screening rates, digital fundus photography using nonmydriatic cameras has been implemented in adult and pediatric clinics (10–12). Digital fundus photography can be performed without the need for pupil dilation and takes a few minutes compared with traditional dilated eye exams (13,14). These cameras produce high-quality images that can be assessed for the presence of DR by eye care providers (optometrists or ophthalmologists) or trained readers in a deferred manner on site or remotely. Use of digital retinal images is sensitive and specific for detection of DR changes, demonstrating substantial agreement with stereoscopic photos of the seven standard fields (15,16), and in some cases, has been found to be more sensitive than dilated eye exams by eye care providers (16,17).

Recently, fully autonomous artificial intelligence (AI)-based systems have been developed for detection of DR and diabetic macular edema (18–21). With autonomous AI systems, retinal images are taken with a nonmydriatic fundus camera and assessed for presence or absence of DR in real time, without the supervision of an eye care provider, making it useful in settings where specialists are not readily available. Similar to other point-of-care (POC) initiatives that have demonstrated improved adherence in patients with diabetes (22), autonomous AI is also expected to improve adherence for the diabetic eye exam, and has already been used and Food and Drug Administration (FDA) approved for adults (18,19). Trials of this technology in the adult primary care setting have demonstrated sensitivity of 87.2%, specificity of 90.7%, with 96% diagnosability (defined as percentage of patients evaluated for DR who receive an interpretable result from the AI algorithm) in detecting more-than-mild DR (mtmDR) and diabetic macular edema. While this technology has been shown to be effective and safe in adults, and is included in the ADA's 2020 Standard of Diabetes Care (23), it has not been studied in the pediatric diabetes population and is not currently FDA approved for use in patients <21 years of age.

The purpose of this study was to prospectively assess the diagnostic efficacy of autonomous AI for diabetic eye exams in pediatric patients in a real-world setting. We also measured adherence to ophthalmic screening guidelines before and after implementation.

## RESEARCH DESIGN AND METHODS

This prospective study was conducted at a multidisciplinary pediatric diabetes clinic affiliated with Johns Hopkins University School of Medicine. The Johns Hopkins University School of Medicine institutional review board approved this study in accordance with the Declaration of Helsinki. Patients aged 5–21 years with T1D, T2D, or cystic fibrosis-related diabetes who were being seen for regular diabetes care between December 2018 and November 2019 were included. At the time this study was initiated, ADA 2018 guidelines for DR screening included youth with T1D aged  $\geq 10$  years with diabetes duration of at least 3 years and T2D at diagnosis. Exclusion criteria

included ophthalmic abnormalities, such as media opacities or strabismus.

Parents/caregivers were notified about the study through secure electronic messaging in advance of their diabetes appointment. A research assistant approached them at their visit and obtained informed consent if they were interested in participating. Participants' clinical data were extracted from their electronic medical record. Parental education and household income were self-reported.

Color photos of the retina were obtained using a TRC-NW400 camera (Topcon, Tokyo, Japan) installed with IDx-DR 2.0 US autonomous AI. No pupil dilation was used. Images were acquired by a trained research assistant with no prior experience in ophthalmic imaging. Two color images were acquired per eye, one macula centered and the other disc centered, for a total of four photos. The IDx-DR system for interpretation of fundus images using autonomous AI is built into the camera and designed to check images for quality and to alert the operator if additional attempts are needed to obtain better-quality images while the patient is still in the exam chair. Image acquisition time was measured for the first 40 patients to assess the feasibility of implementation in a busy diabetes clinic, and a learner operator curve was created. All images were interpreted by the IDx-DR autonomous AI system as either none or mild DR or mtmDR. AI interpretation was received by the clinic within 1 min of image acquisition but was not shared with the patient during the study.

The same color photos of the retina obtained from the TRC-NW400 camera were then uploaded into the electronic medical record in Digital Imaging and Communications in Medicine format and reviewed independently by two retina specialists who were masked to AI interpretation. Retina specialists performed DR grading on the basis of International Clinical Diabetic Retinopathy classification (24). Grades for DR were as follows: none, mild nonproliferative DR (NPDR), moderate NPDR, severe NPDR, and proliferative DR, but results were consolidated into two categories—none and mild DR, and mtmDR—for comparison with AI output. If two graders reported discrepant interpretations, then images were sent to a third retina specialist for adjudication. All three retina specialists also screened for suspicion of macular

edema on the basis of the presence of exudates, hemorrhages, or microaneurysms at the level of the macula. All patients determined to have any DR by at least two readers were referred for a dilated eye exam. Patients with images that could not be read by either retina specialist were also recommended to have a dilated eye exam. Final adjudicated grading or dilated eye exam by retina specialists was considered the reference standard for this study. DR grading for the patient was based on the eye with more severe retinopathy. These results were shared with parents/caregivers and diabetes providers.

### Statistical Analysis

Sensitivity and specificity were calculated by creating  $2 \times 2$  tables, with the final adjudicated image interpretation being used as the reference standard. If images could not be interpreted and a dilated eye exam was obtained, the dilated eye exam results were used as the reference standard. Exact 95% CIs were calculated for sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Diagnosability, defined as the percentage of patients evaluated for DR who receive an interpretable result from the AI algorithm, was determined. While the diagnostic accuracy and feasibility of AI were determined on the basis of the results of the entire cohort, we also performed a subset analysis on participants who met ADA criteria for the diabetic eye exam. Prevalence of DR is reported as presence of any DR among patients. Sensitivity and specificity of AI is based on the output of no-more-than-mild DR (none or mild DR) versus mtmDR. Logistic regression was used to evaluate equity, or the effect of race, ethnicity, age, and sex, on sensitivity, specificity, and diagnosability (25,26). Demographic and clinical characteristics among the subgroups were analyzed using  $\chi^2$  tests for categorical data and Student *t* tests for continuous variables. Poisson regression was used to assess the relationship between characteristics of the population and the number of attempts needed to capture a useable image in AI screening. Predictors of false-positive outcomes and DR prevalence were each assessed using a series of univariate and multivariate logistic regression models. Retina specialist interrater agreement was calculated using weighted Cohen  $\kappa$ .

To account for diagnostic access bias (27), we calculated the adherence-corrected sensitivity (ACS) for the study sample. The calculation for ACS assumes that the prevalence in the unscreened (nonadherent) population is the same as the screened (adherent) population and allows estimation of the positives in the unscreened group. The ACS provides a population-level estimate of the true proportion of cases that are detected, even when a subset of the population does not have access to the test because of diagnostic access bias, irrespective of the reason for not undergoing the test. If two tests have equal sensitivity, but diagnostic access bias is lower for one than the other, ACS will be higher for the former. The ACS is calculated using the following equation:  $ACS = \text{true positives} / (\text{all positives} + \text{estimate of positives in nonadherent sample})$ . Estimate of positives in the nonadherent sample was calculated as  $[(1 - \text{adherence}) \times \text{prevalence} \times \text{study sample size}]$ .

### RESULTS

Over a 12-month period, parents/caregivers of 327 pediatric patients with diabetes were approached in the diabetes clinic and invited to participate. Of these, 310 (95%) consented to participation and screening (Fig. 1). Mean age was  $12.2 \pm 3.6$  years, 47% were male, and 57% were White, 32% Black, 4% Hispanic, and 7% Asian/other ethnicity. Patients predominantly (82%) had T1D. Mean age at diagnosis was  $8.9 \pm 4.4$  years, and median duration of diabetes was 2.52 years (interquartile range 0.75, 5.33). HbA<sub>1c</sub> was  $\geq 7.5\%$  in the majority (70%) of patients. Of participants with T1D, 47% used an insulin pump, and 63% used a continuous glucose monitor. Of those with T2D, 55% used insulin therapy (Table 1).

Of the 310 participants, 308 were graded as having sufficient image quality by the reference standard (Table 2). Of the participants with exams of sufficient quality, 295 (95.8%) had no DR, 6 (2.0%) had mild NPDR, and 7 (2.3%) had moderate NPDR according to the reference standard. No patients had severe NPDR or proliferative DR. Overall prevalence of any DR in this population was 4.2% (4.0% and 4.2% in patients with T1D and T2D, respectively). The interrater agreement for the initial two graders was 92.2%, with a weighted Cohen  $\kappa$  of 0.247. None

of the patients had any suspicion of macular edema.

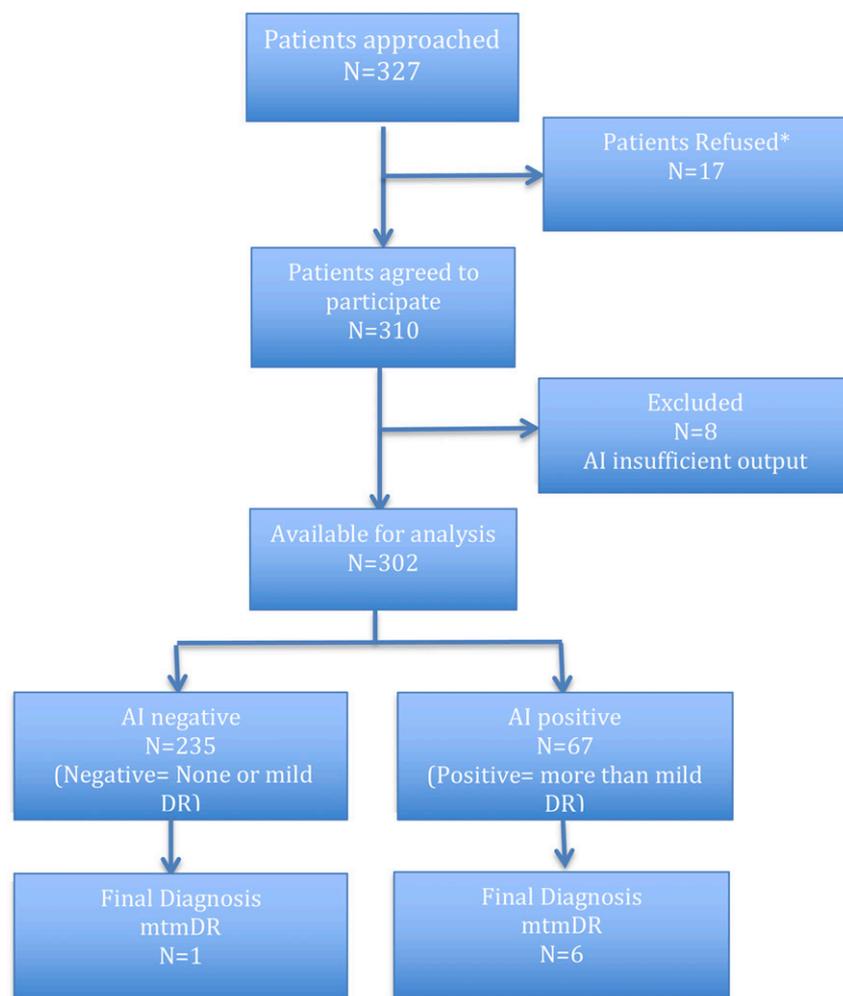
Of the 310 participants, AI gave an interpretation in 302 (97.5%). The images that were not interpretable by AI were due to the participant's inability to keep his or her eyes open during the photographic flash or to focus for the disc-centered images. There was no difference in clinical characteristics (age, sex, race, behavioral diagnosis) between these patients and the overall cohort. The sensitivity and specificity of autonomous AI to detect mtmDR was 85.7% (95% CI 42.1–99.6%) and 79.3% (74.3–83.8%) compared with the reference standard. The PPV and NPV for mtmDR were 9% (3.7–20.2%) and 99.6% (97.6–99.9%), respectively. There was no significant effect of race, ethnicity, age, and sex on specificity and diagnosability.

We performed a subset analysis in patients who met ADA criteria for the diabetic eye exam (149 patients; 48.1% of the entire cohort), of whom 6 had mild DR and 6 had mtmDR. In this subset, the sensitivity and specificity of AI to detect mtmDR was similar to that in the rest of the group at 83.3% (95% CI 35.9–99.6%) and 74.8% (66.8–81.8%), respectively. The PPV and NPV in this subset for mtmDR were 12.5% (4.5–28.8%) and 99.0% (94.8–99.9%), respectively. Participants in this subset were older, included more patients with T2D, and more Black participants compared with the overall cohort ( $P < 0.01$ ).

One hundred fifty-two (49%) participants reported having a diabetic eye exam before participation in this study. Of this group, only 11.3% had a record of the dilated eye exam results in their chart. After implementation, adherence improved from 49% to 95%. On the basis of our preimplementation adherence rate for diabetic eye exams of 49%, the corresponding ACS for detection of mtmDR was 30%. With AI implementation, ACS was 85%.

One participant had a false-negative result for mtmDR compared with the reference standard. Review of the retinal images revealed an isolated retinal hemorrhage nasal to the disc in one eye, with no microaneurysms in either eye (see Fig. 2A).

There were 61 patients who were false positive for mtmDR compared with the reference standard (Fig. 2B). All false-positive images were reviewed, and the



**Figure 1**—Standards for the Reporting of Diagnostic Accuracy Studies diagram describing recruitment numbers and device mtmDR output. \*Reasons for refusal: lack of time, already had eye exam, and legal guardian not present.

reasons were classified as follows: 1) images having a very shiny internal limiting membrane (ILM) with diffuse sheen or individual shiny dots likely being mistaken for exudates ( $n = 44$ ) or 2) images noted to have quality issues (slightly blurred, edge artifact, too dark, or a diffuse reddish hue) ( $n = 17$ ). The ILM sheen is commonly seen in the retina of children; thus, we evaluated whether age was associated with false-positive results. On univariate and multivariate analysis, age, sex, and ethnicity were not associated with false-positive results.

In 184 (60%) patients, it took one attempt to obtain adequate images; in 115 (37.5%), it took two or more attempts; and in 8 (2.5%), a result could not be obtained. On multivariate analysis, age, sex, ethnicity, and behavioral disorders (attention deficit hyperactivity disorder, autism) were not significantly

associated with the number of attempts required to get an adequate image. The average imaging time for the first 40 patients was 7 min and 13 s (SD 2.5 min), with patients 31–40 taking an average of 6 min and 35 s.

## CONCLUSIONS

This is the first study to apply an autonomous AI system for the diabetic eye exam in the pediatric population. The results show that AI in pediatric diabetes has a sensitivity of 85.7%, specificity of 79.3%, and diagnosability rate of 97% for detection of mtmDR compared with the reference standard as defined by retina specialists. The high sensitivity, specificity, and NPV demonstrate the safety and efficacy of its use in pediatrics, comparable to the diagnostic accuracy of the system in an adult primary care setting (19). Furthermore, the high diagnosability,

real-time AI feedback for image acquisition; brief imaging time; and ability to use a nonophthalmology-trained operator allow for successful integration into the existing workflow of a busy multidisciplinary diabetes clinic. This is especially timely because the most recent ADA guidelines include AI systems as a new approach to the diabetic eye exam.

In this study, we used adjudicated grades by retina specialists as the reference standard to allow us to compare our results with a more real-world teleophthalmology-like scenario where a reading center is typically not used to detect referable DR. The retina specialists graded the same images as the AI, as has been done in some studies (20). However, other prospective studies have compared the AI output with a surrogate outcome, a reference standard validated to predict patient outcome on the basis of the Early Treatment Diabetic Retinopathy Study (ETDRS) severity scale (19,28). The ETDRS severity scale was established after decades of epidemiological research in adults, where grading of fundus photos by trained reading center graders is linked with a specific likelihood of vision loss (28). This serves as a surrogate outcome to clinical outcomes, such as vision loss and blindness, that may take years to manifest.

The prevalence of DR in our prospective cohort is similar to reported rates, with recent studies showing prevalence near 4–5% in T1D (13,29) and 4–13.7% in adolescents with T2D (30–32). The low prevalence and low risk for DR in this patient population create an ideal screening scenario. Implementation of POC detection of DR using AI improved screening rates in this cohort, where only one-half (49%) reported a prior diabetic eye exam and 95% agreed to proceed with POC screening. When considering the broader population health, the higher ACS with implementation of AI indicates that the majority of DR cases in the entire at-risk population will be identified (33). Thus, implementing this system in the diabetes care setting is likely to increase screening rates and adherence to screening guidelines and reach value-based care measures (Healthcare Effectiveness Data and Information Set). By providing real-time results, especially if the result is positive for DR, the patient is more likely to follow through with the recommended eye care provider exam (34,35). Additionally, with the increasing

**Table 1—Characteristics of the study population**

Variable	Entire study population	Did not meet screening guidelines	Met screening guidelines	P value
Patients, n (%)	310	161 (51.9)	149 (48.1)	
Age (years), mean ± SD	12.2 ± 3.6	10.1 ± 3.3	14.5 ± 2.3	<b>0.000</b>
Age at diagnosis (years), mean ± SD	8.9 ± 4.4	8.5 ± 4.2	9.3 ± 4.4	0.119
Duration of diabetes (years), median (IQR)	2.52 (0.75, 5.33)	1.61 (0.47, 2.79)	5.04 (2.33, 7.21)	<b>0.000</b>
Female sex, n (%)	166 (53.5)	93 (57.8)	73 (49)	0.122
Race/ethnicity, n (%)				
White	177 (57.1)	102 (63.3)	75 (50.3)	<b>0.003</b>
Black	100 (32.3)	39 (24.2)	61 (40.9)	
Hispanic or Latino	13 (4.2)	5 (3.1)	8 (5.4)	
Other	20 (6.5)	15 (9.3)	5 (3.4)	
Type of diabetes, n (%)				
T1D	254 (81.9)	154 (95.7)	100 (67.1)	<b>0.000</b>
T2D	48 (15.5)	0 (0)	48 (32.2)	
CFRD/other	8 (2.6)	7 (4.3)	1 (0.7)	
HbA <sub>1c</sub> , n (%)				
<7.5%	81 (26.1)	45 (27.8)	36 (24.3)	0.547
≥7.5%	218 (70.3)	109 (67.7)	109 (73.2)	
Unknown	11 (3.6)	7 (4.3)	4 (2.7)	
Insulin form, n (%)				
Injection	147 (47.4)	87 (53.7)	60 (40.5)	<b>0.000</b>
Insulin pump	141 (45.5)	73 (45.1)	68 (46)	
Not taking insulin	22 (7.1)	2 (1.2)	20 (13.5)	
CGM, n (%)*	159 (62.6)	97 (63)	62 (62)	0.874
Metformin, n (%)**	39 (81.3)	0 (0)	39 (81.3)	

The t test was used to compare continuous variables, and  $\chi^2$  test was used to analyze categorical data. Duration of diabetes was nonparametric and thus is reported as median with IQR, and Wilcoxon rank sum test was used to compare subgroups. Boldface type indicates significance level at  $P < 0.05$ . CFRD, cystic fibrosis–related diabetes; CGM, continuous glucose monitor; IQR, interquartile range. \*Patients with T1D. \*\*Patients with T2D.

incidence of T2D in adolescents, POC screening is even more critical because T2D disproportionately affects minorities and youth from lower socioeconomic backgrounds who are otherwise less likely to be screened (7,30).

The IDx-DR biomarker-based algorithm used in our study is lesion based rather than a “black-box” approach to detect DR (18). It uses explicit detectors for characteristic DR lesions that are racially invariant, such as hemorrhages and microaneurysms (36). The algorithm uses population encoding of multiple lesion detectors that overlap in the high-dimensional feature space, rather than being statistically independent, built using machine learning from known lesions

samples and mathematical models of the disease process forming the lesion (18,37,38). In our study, we showed successful use, efficacy, and equity of this AI system across age, sex, race, and ethnicity, suggesting unbiased and equitable performance in the pediatric diabetes population.

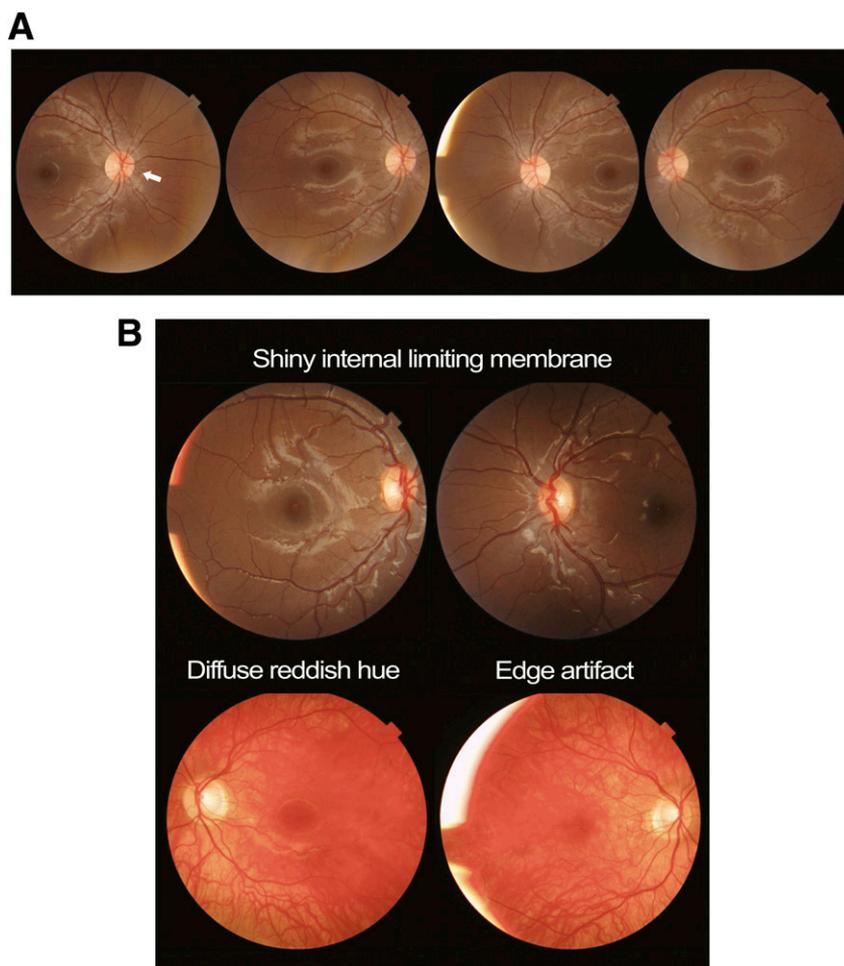
We had a high rate of false-positive results from AI in our patient population compared with AI performance in adults (18,19). The predominant reason for the false-positive results was a more prominent ILM in pediatrics, leading to either the appearance of a yellowish sheen or shiny dots that are more easily mistaken as exudates by the algorithm. It is possible that if an optical coherence tomography

(OCT) scan rather than retina specialist grades had been used to determine the “ground truth,” some of these false positives could have been true positives (i.e., if the OCT scan had shown edema in areas that looked like possible exudates on the images). Despite the high rate of false positives, it is important to note that sensitivity and specificity are still quite high at 85.7% (95% CI 42.1–99.6%) and 79.3% (74.3–83.8%). In our sample, only 2.3% of patients had moderate DR, and none had proliferative DR, making it highly unlikely that a case of vision-threatening DR would be missed by implementing the autonomous AI system for screening of DR in pediatric diabetes clinics. In fact, when we looked at ACS for detection of DR in our population, we found that having POC AI leads to many more patients meeting the screening guidelines for the diabetic eye exam. The benefits and cost-effectiveness of DR screening is well established by prior studies (39,40). Screening criteria have been established by the ADA and American Academy of Ophthalmology, and here we compare the performance of a newer modality (AI based) for the

**Table 2—Results of reference standard grading and AI output**

AI output	DR (per reference standard)			Total
	Negative or mild DR	mtmDR	Undiagnosable	
Negative or mild DR	234	1	0	235
mtmDR	61	6	0	67
Undiagnosable	6	0	2	8
Total	301	7	2	310

Data include all study participants.



**Figure 2**—A: Fundus photos from a patient with false-negative results. Disc- and macula-centered fundus photos from both eyes of one patient who was interpreted as having no DR by the AI system but noted to have moderate DR by the reference standard. Review of the fundus photos shows a small hemorrhage nasal to the disc in the right eye (arrow), with no other lesions suggestive of DR. B: Fundus photos from patients with false-positive results. Fundus photographs from patients who were interpreted as positive for DR by the AI system but were noted to have no DR per the reference standard. The top fundus photos are from the right and left eye of a patient. These photos depict the very shiny ILM surface noted in our patient cohort. The shiny ILM reflections appear as broad yellow bands or as focal yellow dots, which can be mistaken for exudates. The bottom fundus photos show diffuse reddish hue and edge artifact.

diabetic eye exam with more traditional screening methods (i.e., review of photos by retina specialists). While the number of false negatives is much lower than the standard-of-care diabetic eye exam by a clinician (with a sensitivity of ~30–40%) (16), the number of false positives is higher compared with a clinician reference standard, and every false positive requires evaluation by a clinician. For the first time ever, in the IDx-DR FDA clinical trial for adults, the FDA set sensitivity and specificity at the population level at 80% (18). While those end points needed to be met in comparison with a surrogate outcome, rather than with clinicians, the performance in this pediatric study was compared with clinician grading

and shows noninferiority to the FDA's thresholds. There is also the advantage of an immediate result associated with AI in addition to time and cost savings to the patient, which we recently demonstrated in a cost-savings analysis (41).

Our results are limited by the low prevalence of DR in the pediatric population. Studies with larger samples and more sensitive reference standards, such as those using OCT scans, are needed to further optimize an AI system for the pediatric population. The few true positives also limit conclusions on the specificity. However, the low prevalence of disease makes it less likely that we would miss vision-threatening retinopathy. Larger

multicenter studies are being planned to overcome the limitation of a single-center study. Future studies assigning patients to AI-based versus traditional screening methods are needed to accurately determine the role of AI in improving adherence.

In conclusion, use of a nonmydriatic fundus camera with an autonomous AI system was safe and effective for the diabetic eye exam in our study. However, given that this is a single-center study, the results may not be generalizable to the entire pediatric diabetes population. As the prevalence of diabetes increases in youth and adherence to screening guidelines remains suboptimal, POC AI can help to improve screening rates in this population.

**Acknowledgments.** The authors are grateful to the diabetes providers and patients at the Johns Hopkins Pediatric Diabetes Center for participation.

**Funding.** This investigator-initiated study was funded by a Johns Hopkins Children's Center Innovation Award to R.M.W. and R.C. This work was supported in part by a Research to Prevent Blindness unrestricted grant to the Department of Ophthalmology and Visual Sciences, University of Wisconsin.

**Duality of Interest.** R.M.W. reports financial support from Boehringer Ingelheim not relevant to this article. M.D.A. reports being an employee, investor, and executive chairman of IDx. M.D.A. has patents and patent applications assigned to the University of Iowa and IDx that are relevant to the subject matter of this article. No other potential conflicts of interest relevant to this article were reported.

**Author Contributions.** R.M.W., C.T., and K.S. acquired the data. R.C., T.Y.A.L., I.Z.G., and M.D.A. interpreted the images. R.M.W., M.D.A., and R.C. conceived of the study. R.M.W. and R.C. designed the study. L.P. performed the analysis. All authors contributed to the writing and critical review of the manuscript. R.M.W. and R.C. are guarantors of this work and, as such, 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.

**Prior Presentation.** Parts of this study were presented in abstract form at the 79th Scientific Sessions of the American Diabetes Association, San Francisco, CA, 7–11 June 2019.

## References

- Centers for Disease Control and Prevention. *National Diabetes Statistics Report, 2020*. Atlanta, GA, Centers for Disease Control and Prevention, US Department of Health and Human Services, 2020
- Mayer-Davis EJ, Lawrence JM, Dabelea D, et al.; SEARCH for Diabetes in Youth Study. Incidence trends of type 1 and type 2 diabetes among youths, 2002–2012. *N Engl J Med* 2017; 376:1419–1429

3. American Diabetes Association. Standards of medical care in diabetes—2014. *Diabetes Care* 2014;37(Suppl. 1):S14–S80
4. Nathan DM, Genuth S, Lachin J, et al.; Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med* 1993;329:977–986
5. Foster NC, Beck RW, Miller KM, et al. State of type 1 diabetes management and outcomes from the T1D Exchange in 2016–2018. *Diabetes Technol Ther* 2019;21:66–72
6. American Diabetes Association. 13. Children and adolescents: *Standards of Medical Care in Diabetes—2019*. *Diabetes Care* 2019;42(Suppl. 1):S148–S164
7. Wang SY, Andrews CA, Gardner TW, Wood M, Singer K, Stein JD. Ophthalmic screening patterns among youths with diabetes enrolled in a large US managed care network. *JAMA Ophthalmol* 2017;135:432–438
8. Dabelea D, Mayer-Davis EJ, Saydah S, et al.; SEARCH for Diabetes in Youth Study. Prevalence of type 1 and type 2 diabetes among children and adolescents from 2001 to 2009. *JAMA* 2014;311:1778–1786
9. Mayer-Davis EJ, Davis C, Saadine J, et al.; SEARCH for Diabetes in Youth Study Group. Diabetic retinopathy in the SEARCH for Diabetes in Youth Cohort: a pilot study. *Diabet Med* 2012;29:1148–1152
10. Kolomeyer AM, Nayak NV, Simon MA, et al. Feasibility of retinal screening in a pediatric population with type 1 diabetes mellitus. *J Pediatr Ophthalmol Strabismus* 2014;51:299–306
11. Roser P, Kalscheuer H, Groener JB, et al. Diabetic retinopathy screening ratio is improved when using a digital, nonmydriatic fundus camera onsite in a diabetes outpatient clinic. *J Diabetes Res* 2016;2016:4101890
12. Stillman JK, Gole GA, Wootton R, et al. Telepaediatrics and diabetic retinopathy screening of young people with diabetes in Queensland. *J Telemed Telecare* 2004;10(Suppl. 1):92–94
13. Tapley JL, McGwin G Jr., Ashraf AP, et al. Feasibility and efficacy of diabetic retinopathy screening among youth with diabetes in a pediatric endocrinology clinic: a cross-sectional study. *Diabetol Metab Syndr* 2015;7:56
14. Scarpa G, Urban F, Vujosevic S, et al. The nonmydriatic fundus camera in diabetic retinopathy screening: a cost-effective study with evaluation for future large-scale application. *J Ophthalmol* 2016;2016:4625096
15. Aptel F, Denis P, Rouberol F, Thivolet C. Screening of diabetic retinopathy: effect of field number and mydriasis on sensitivity and specificity of digital fundus photography. *Diabetes Metab* 2008;34:290–293
16. Lin DY, Blumenkranz MS, Brothers RJ, Grosvenor DM. The sensitivity and specificity of single-field nonmydriatic monochromatic digital fundus photography with remote image interpretation for diabetic retinopathy screening: a comparison with ophthalmoscopy and standardized mydriatic color photography. *Am J Ophthalmol* 2002;134:204–213
17. Liu Y, Rajamanickam VP, Parikh RS, et al. Diabetic retinopathy assessment variability among eye care providers in an urban teleophthalmology program. *Telemed J E Health* 2019;25:301–308
18. Abramoff 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
19. Verbraak FD, Abramoff MD, Bausch GCF, et al. Diagnostic accuracy of a device for the automated detection of diabetic retinopathy in a primary care setting. *Diabetes Care* 2019;42:651–656
20. Bhaskaranand M, Ramachandra C, Bhat S, et al. 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;21:635–643
21. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 2016;316:2402–2410
22. Schnell O, Crocker JB, Weng J. Impact of HbA<sub>1c</sub> testing at point of care on diabetes management. *J Diabetes Sci Technol* 2017;11:611–617
23. American Diabetes Association. 11. Microvascular complications and foot care: *Standards of Medical Care in Diabetes—2020*. *Diabetes Care* 2020;43(Suppl. 1):S135–S151
24. Wilkinson CP, Ferris FL III, Klein RE, et al.; Global Diabetic Retinopathy Project Group. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology* 2003;110:1677–1682
25. Channa R, Wolf R, Abramoff MD. Autonomous artificial intelligence in diabetic retinopathy: from algorithm to clinical application. *J Diabetes Sci Technol*. 4 March 2020 [Epub ahead of print]. DOI: 10.1177/1932296820909900
26. Abramoff MD, Tobey D, Char DS. Lessons learned about autonomous AI: finding a safe, efficacious, and ethical path through the development process. *Am J Ophthalmol* 2020;214:134–142
27. Sackett DL. Bias in analytic research. *J Chronic Dis* 1979;32:51–63
28. Early Treatment Diabetic Retinopathy Study Research Group. Fundus photographic risk factors for progression of diabetic retinopathy. ETDRS report number 12. *Ophthalmology* 1991;98(Suppl.):823–833
29. Dabelea D, Stafford JM, Mayer-Davis EJ, et al.; SEARCH for Diabetes in Youth Research Group. Association of type 1 diabetes vs type 2 diabetes diagnosed during childhood and adolescence with complications during teenage years and young adulthood. *JAMA* 2017;317:825–835
30. TODAY Study Group. Retinopathy in youth with type 2 diabetes participating in the TODAY clinical trial. *Diabetes Care* 2013;36:1772–1774
31. Eppens MC, Craig ME, Cusumano J, et al. Prevalence of diabetes complications in adolescents with type 2 compared with type 1 diabetes. *Diabetes Care* 2006;29:1300–1306
32. Porter M, Channa R, Wagner J, Prichett L, Liu TYA, Wolf RM. Prevalence of diabetic retinopathy in children and adolescents at an urban tertiary eye care center. *Pediatr Diabetes* 2020;21:856–862
33. Wong TY, Bressler NM. Artificial intelligence with deep learning technology looks into diabetic retinopathy screening. *JAMA* 2016;316:2366–2367
34. Mansberger SL, Shepler C, Barker G, et al. Long-term comparative effectiveness of telemedicine in providing diabetic retinopathy screening examinations: a randomized clinical trial. *JAMA Ophthalmol* 2015;133:518–525
35. Crossland L, Askew D, Ware R, et al. Diabetic retinopathy screening and monitoring of early stage disease in Australian general practice: tackling preventable blindness within a chronic care model. *J Diabetes Res* 2016;2016:8405395
36. Hansen MB, Abramoff MD, Folk JC, Mathenge W, Bastawrous A, Peto T. Results of automated retinal image analysis for detection of diabetic retinopathy from the Nakuru Study, Kenya. *PLoS One* 2015;10:e0139148
37. Abramoff MD, Alward WL, Greenlee EC, et al. Automated segmentation of the optic disc from stereo color photographs using physiologically plausible features. *Invest Ophthalmol Vis Sci* 2007;48:1665–1673
38. Finlayson SG, Bowers JD, Ito J, Zittrain JL, Beam AL, Kohane IS. Adversarial attacks on medical machine learning. *Science* 2019;363:1287–1289
39. Javitt JC, Aiello LP. Cost-effectiveness of detecting and treating diabetic retinopathy. *Ann Intern Med* 1996;124:164–169
40. Xie Y, Nguyen QD, Hamzah H, et al. Artificial intelligence for teleophthalmology-based diabetic retinopathy screening in a national programme: an economic analysis modelling study. *Lancet Digit Health* 2020;2:e240–e249
41. Wolf RM, Channa R, Abramoff MD, Lehmann HP. Cost-effectiveness of autonomous point-of-care diabetic retinopathy screening for pediatric patients with diabetes. *JAMA Ophthalmol* 2020;138:1063–1069