

Original Paper

Implementation of Artificial Intelligence–Based Diabetic Retinopathy Screening in a Tertiary Care Hospital in Quebec: Prospective Validation Study

Fares Antaki^{1,2,3,4,5*}, MDCM; Imane Hammana^{6*}, PhD; Marie-Catherine Tessier², MSc; Andrée Boucher⁷, MD; Maud Laurence David Jetté⁸, MPA; Catherine Beauchemin⁹, PhD; Karim Hammamji^{2,3}, MD; Ariel Yuhan Ong^{1,5,10}, MBChB; Marc-André Rhéaume^{2,3}, MD; Danny Gauthier^{2,3}, MD; Mona Harissi-Dagher^{2,3}, MD; Pearse A Keane^{1,5,11}, MD; Alfons Pomp^{6,12}, MD

¹Institute of Ophthalmology, University College London, London, United Kingdom

²Department of Ophthalmology, Centre Hospitalier de l'Université de Montréal, Montreal, QC, Canada

³Department of Ophthalmology, Université de Montréal, Montreal, QC, Canada

⁴The CHUM School of Artificial Intelligence in Healthcare, Centre Hospitalier de l'Université de Montréal, Montreal, QC, Canada

⁵Moorfields Eye Hospital NHS Foundation Trust, London, United Kingdom

⁶Health Technology Assessment Unit, Centre Hospitalier de l'Université de Montréal, Montreal, QC, Canada

⁷Division of Endocrinology, Department of Medicine, Centre Hospitalier de l'Université de Montréal, Montreal, QC, Canada

⁸Direction du soutien à la transformation, Centre Hospitalier de l'Université de Montréal, Montreal, QC, Canada

⁹Faculty of Pharmacy, University of Montreal, Montreal, QC, Canada

¹⁰Oxford Eye Hospital, Oxford University Hospitals NHS Foundation Trust, Oxford, United Kingdom

¹¹NIHR Moorfields Biomedical Research Centre, London, United Kingdom

¹²Department of Surgery, University of Montréal, Montreal, QC, Canada

* these authors contributed equally

Corresponding Author:

Fares Antaki, MDCM
Institute of Ophthalmology
University College London
11-43 Bath St
London, EC1V 9EL
United Kingdom
Phone: 44 20 7608 6800
Email: f.antaki@ucl.ac.uk

Abstract

Background: Diabetic retinopathy (DR) affects about 25% of people with diabetes in Canada. Early detection of DR is essential for preventing vision loss.

Objective: We evaluated the real-world performance of an artificial intelligence (AI) system that analyzes fundus images for DR screening in a Quebec tertiary care center.

Methods: We prospectively recruited adult patients with diabetes at the Centre hospitalier de l'Université de Montréal (CHUM) in Montreal, Quebec, Canada. Patients underwent dual-pathway screening: first by the Computer Assisted Retinal Analysis (CARA) AI system (index test), then by standard ophthalmological examination (reference standard). We measured the AI system's sensitivity and specificity for detecting referable disease at the patient level, along with its performance for detecting any retinopathy and diabetic macular edema (DME) at the eye level, and potential cost savings.

Results: This study included 115 patients. CARA demonstrated a sensitivity of 87.5% (95% CI 71.9-95.0) and specificity of 66.2% (95% CI 54.3-76.3) for detecting referable disease at the patient level. For any retinopathy detection at the eye level, CARA showed 88.2% sensitivity (95% CI 76.6-94.5) and 71.4% specificity (95% CI 63.7-78.1). For DME detection, CARA had 100% sensitivity (95% CI 64.6-100) and 81.9% specificity (95% CI 75.6-86.8). Potential yearly savings from implementing CARA at the CHUM were estimated at CAD \$245,635 (US \$177,643.23, as of July 26, 2024) considering 5000 patients with diabetes.

Conclusions: Our study indicates that integrating a semiautomated AI system for DR screening demonstrates high sensitivity for detecting referable disease in a real-world setting. This system has the potential to improve screening efficiency and reduce costs at the CHUM, but more work is needed to validate it.

(*JMIR Diabetes* 2024;9:e59867) doi: [10.2196/59867](https://doi.org/10.2196/59867)

KEYWORDS

artificial intelligence; diabetic retinopathy; screening; clinical validation; diabetic; diabetes; screening; tertiary care hospital; validation study; Quebec; Canada; vision; vision loss; ophthalmological; AI; detection; eye

Introduction

Diabetes mellitus is a prevalent metabolic disease affecting 5.7 million Canadians, or 14% of the population, in 2022. This number is expected to increase to 7.3 million by 2032 [1]. Diabetic retinopathy (DR) is a common complication of the disease, affecting up to 25% of patients in Canada [2]. DR is also the leading cause of vision loss in people of working age and is associated with increased mortality [3]. Detection of DR at an early stage, when it can be treated with the best prognosis, is crucial in preventing vision loss [4].

Diabetes Canada recommends annual to biennial DR screening by an ophthalmologist or an optometrist for all people with diabetes [3]. Despite this recommendation, up to a third of diabetic patients go unscreened in Ontario, for example [5]. Demographic and socioeconomic factors, such as low income and immigration, are the main risk factors for being unscreened [5]. Numerous Canadian provinces have implemented more inclusive DR screening programs through DR telescreening (DRTS) [6-10].

In Quebec, significant progress has been made in developing DRTS pathways for patients with diabetes in distant regions [11]. However, in large urban centers, only a minority of patients benefit from such access [6]. For most patients, screening is carried out in optometry clinics, resulting in out-of-pocket expenses, as these services are not covered by the provincial health insurance program [12]. Many patients are referred to consult with an ophthalmologist at a hospital. However, this care process is usually inefficient, with numerous obstacles to effective screening [6].

The recent demonstrations of artificial intelligence (AI)-based grading systems for DR have sparked interest into their integration in pre-established DRTS care pathways in Quebec

[11,13]. These systems can integrate existing DRTS pathways in 2 ways: a semiautomated manner, where they replace the preliminary triage currently performed by level 1 trained graders [14]. Alternatively, they can operate in a fully autonomous way, which would not require any human oversight [15].

In this study, we share findings from incorporating a semiautomated AI model into the care strategy for diabetic patients at a major tertiary care center in Quebec. We report on the prospective clinical validation of this AI system, showcasing its real-world performance at both the patient and eye levels for detecting retinopathy and diabetic macular edema (DME). We also perform an economic analysis to estimate the potential cost savings achievable with the implementation of the AI system.

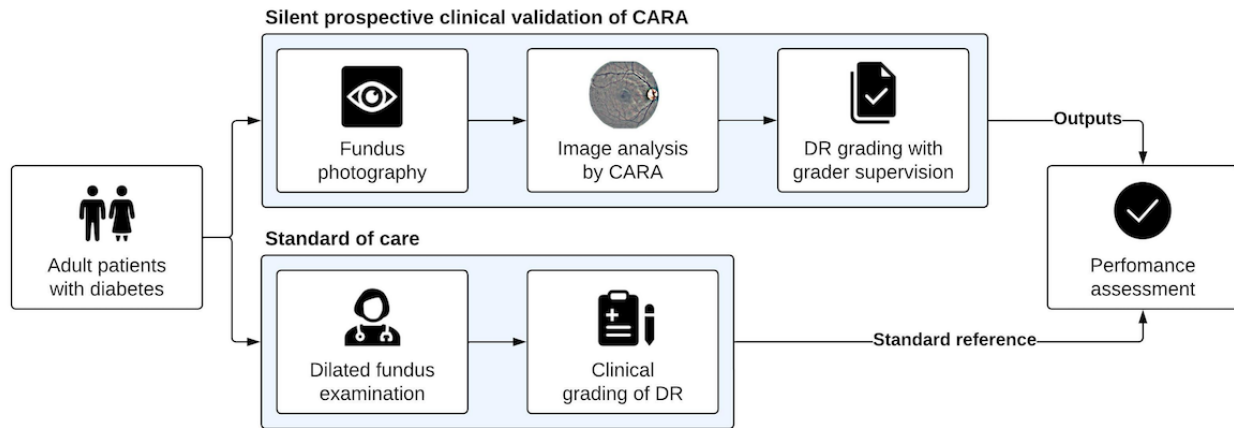
Methods

Study Design

The Centre hospitalier de l'Université de Montréal (CHUM) is a tertiary and quaternary care teaching hospital and one of the two major health care networks in Montreal, Quebec, Canada. It serves 500,000 patients annually. At the CHUM, patients with diabetes are managed by the endocrinology service in specialized clinics, with DR screening conducted either externally or within the ophthalmology department, depending on patient and physician choice.

To improve the patient experience and increase screening uptake, we implemented an AI-based DRTS pathway within the CHUM. To evaluate our approach, we designed a silent prospective clinical validation trial. All recruited patients were screened through 2 parallel pathways: screening by an AI-based system called Computer Assisted Retinal Analysis (CARA), followed by standard of care screening by a CHUM ophthalmologist who was masked to the output of CARA. The study pathway is described in [Figure 1](#).

Figure 1. Overview of the silent prospective clinical validation study. Adult patients with diabetes were recruited from CHUM's endocrinology service. After consenting, fundus photography was obtained on the same day. The fundus images were then transferred through the cloud to the prospective validation pipeline. The images were automatically analyzed by the CARA algorithm. Before generating the report, a trained grader employed by the service provider would grade the CARA-processed images. Within a month, with the outputs of CARA masked, the patients were referred to receive standard of care screening with an ophthalmologist at the CHUM. To determine the discriminative performance of CARA, the outputs from the silent prospective validation were compared to the ophthalmologist exam (standard reference). CARA: Computer Assisted Retinal Analysis; CHUM: Centre hospitalier de l'Université de Montréal; DR: diabetic retinopathy.



Study Population

Patients who attended the diabetes clinic of the endocrinology service between September 2018 and January 2019 were recruited. All patients were new patients to the department of ophthalmology. Previous DR diagnosis, past retinal disease, or intraocular surgery was not exclusionary.

CARA: A Semiautomated AI Model for DR Detection

The CARA system, a semiautomated AI developed by Diagnos Inc. in Montreal, is a clinically inspired machine learning (ML) algorithm designed to detect retinopathy and maculopathy. As a traditional ML model, CARA required the handcrafting of feature extractors that transform pixel values of a retinal image into feature vectors for an ML classifier [16,17]. This is different from deep learning approaches which became prominent after the development of CARA [17,18]. CARA identifies specific lesions in retinal images, including dark lesions like microaneurysms and hemorrhages, as well as bright lesions such as hard exudates and cotton wool spots [19,20].

The algorithm uses a multistep process involving image enhancement, laterality determination, vascular network, optic nerve and fovea detection, and lesion identification. A final output is reached by a weighted combination of image quality, the highest probability of bright and dark lesions, and image conformity. This output is reviewed by a Diagnos Inc senior grader before being sent to the CHUM, marking the process as semiautomated. The senior grader, an ophthalmologist with 12 years of grading experience, was employed by Diagnos Inc. The CHUM was not informed of any formal auditing or quality assessment methods used by Diagnos Inc.

CARA received approval from the US Food and Drug Administration as a class II medical device in 2011 through the 510(k) pathway. It was intended as a software platform to visualize, store, and enhance color fundus images through computerized networks, but was not certified for diagnosing

DR at that time. Prior to this current study, CARA had shown a 93% sensitivity and 71% specificity in retinopathy detection across an international data set of 509 eyes (personal communication with Diagnos Inc). Our study served as a first prospective clinical validation study for the CARA service.

Fundus Photography

On recruitment day, in the endocrinology clinic, 45-degree, nonstereoscopic color fovea-centered fundus photographs were obtained by a technician, without pupillary mydriasis, using the Centervue DRS camera (Hillrom, Chicago, United States). A single image was obtained per eye as CARA was designed to make classifications from single-field images. At the time of the study's design, there were no Canadian recommendations in that regard, but this approach seemed reasonable as single-field images were being used in the Scottish DR screening services, for example [21]. The acquired images were uploaded to the CARA platform for them to be processed by the ML algorithm. The generation of AI outputs was carried out asynchronously and the patient, the referring endocrinologist, and all eye care providers were masked to the output of the AI system.

Ophthalmologist Examination (Standard Reference)

Within a month of recruitment, patients were referred to undergo screening by an ophthalmologist using slit lamp examination, typically with a 78 or 90D lens. Fundus photography and optical coherence tomography were not routinely carried out. Each patient's grading was performed by a single attending ophthalmologist, with the study including a total of 28 ophthalmologists with diverse subspecialties and levels of experience (Multimedia Appendix 1). No specific grading training was administered; however, all participating ophthalmologists were board-certified by the Royal College of Physicians and Surgeons of Canada. They were expected to have the competence to diagnose and grade DR, as it is considered an objective (#3.1.2.4.2.2) of ophthalmology training

by the Royal College [22]. Using standardized case report forms, the ophthalmologists graded each eye using the Scottish Diabetic Retinopathy Grading Scheme [23,24]. Standardized imaging examples were available as needed. This represented the standard reference to which the AI model was compared to.

Disease Definition

In a DR screening program, individuals with no or mild DR undergo annual monitoring, while those with more than mild (mtmDR), and those with DME are referred for evaluation [25]. Accordingly, numerous pivotal trials of AI-based DR screening models have focused on their ability to detect mtmDR [13,15]. The intended use of CARA, however, was as a tool to detect any retinopathy (including mild disease) and DME. This included, according to the Scottish Diabetic Retinopathy Grading Scheme: R1, R2, R3, and R4 for retinopathy grading, and M1 and M2 for DME grading [23]. This is equivalent to level 1 triaging responsibility as defined by the CR2N Tele-Retina Steering Committee [14].

For each patient, we determined if they were referable by mapping the retinopathy and DME labels to “retinopathy or DME present” (referable) or “retinopathy and DME absent” (not referable), taking the worst of the 2 eyes to correspond to the outputs of the AI system at the patient level. Patients without referral outputs for both eyes were considered “inconclusive.”

Economic Analysis

We conducted an economic evaluation to estimate the costs of adopting the CARA system for screening 5000 annual patients with diabetes at the CHUM, analyzed from a health care system perspective. The direct costs of screening, and costs for inconclusive outputs and false positive referrals were considered in the analysis. All values are presented in Canadian dollars.

Outcomes and Statistical Analysis

The primary outcomes were the sensitivity and specificity of the AI system to detect referable disease at the patient level. The secondary outcomes were first, the sensitivity and specificity of the AI system for the detection of retinopathy and DME at the eye level, and second, the cost savings in Canadian dollars should the system be implemented.

Sensitivity, specificity, positive predictive value, and negative predictive value for each of retinopathy and DME detection were computed from the confusion matrices. We used the Wilson score interval method to calculate 95% CI. The study follows the Standards for Reporting of Diagnostic Accuracy (STARD) checklist [26].

Ethical Considerations

Patients were enrolled only if they understood the study and provided informed consent. This research complied with the Declaration of Helsinki. Departmental approval was obtained, and an ethical waiver was granted by the CHUM given the silent, noninterventive nature of the study. All data was deidentified.

Results

Patient Flow

Between September 2018 and January 2019, the endocrinology service referred 133 patients to the study, of which 115 (230 eyes) underwent fundus imaging and received CARA outputs. [Multimedia Appendix 2](#) presents demographics of the 18 excluded patients, showing no significant differences from those included. [Figures 2 and 3](#) illustrate recruitment at the patient and eye levels for retinopathy and DME, respectively.

Figure 2. STARD (Standards for Reporting of Diagnostic Accuracy) flowchart at the patient level. Referable disease was defined as “retinopathy or DME present” and not referable was defined as “retinopathy and DME absent”, taking the worst of the two eyes to correspond to the outputs of the AI system at the patient level. CARA: Computer Assisted Retinal Analysis; DME: diabetic macular edema.

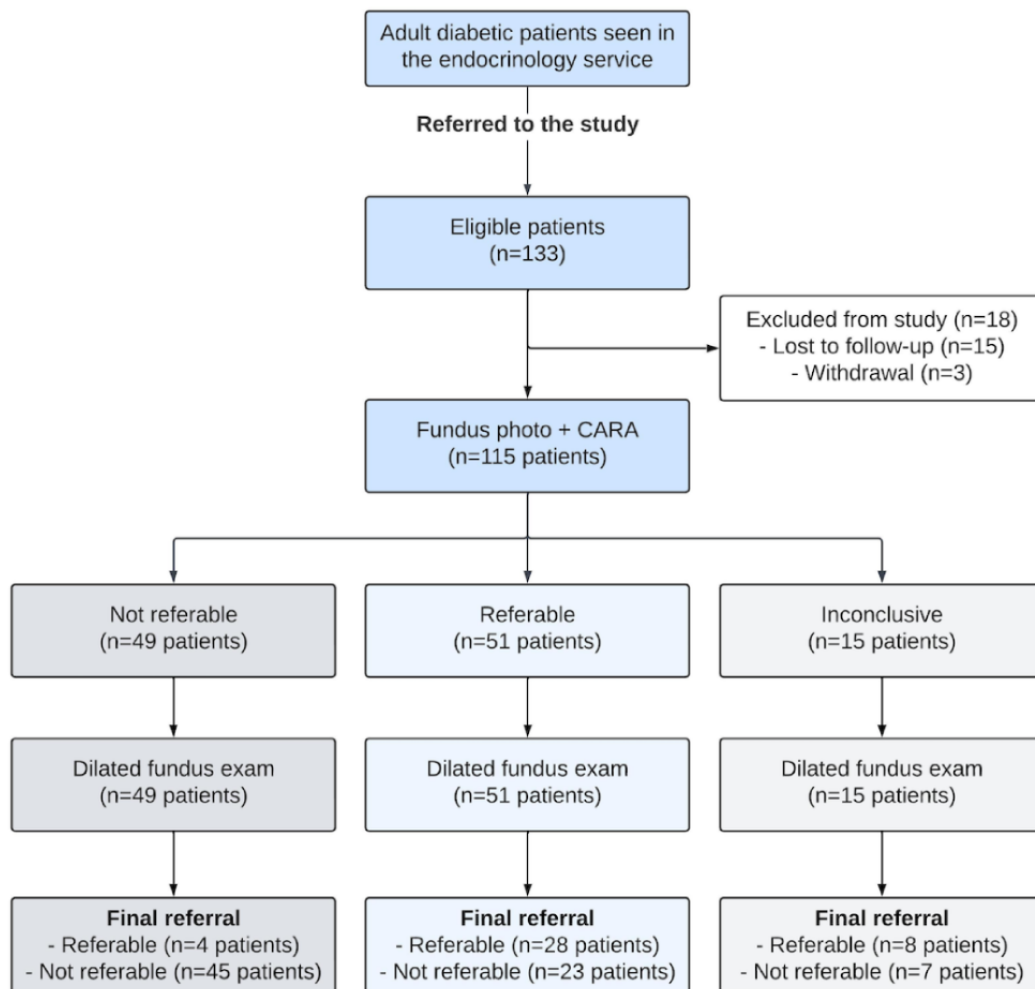
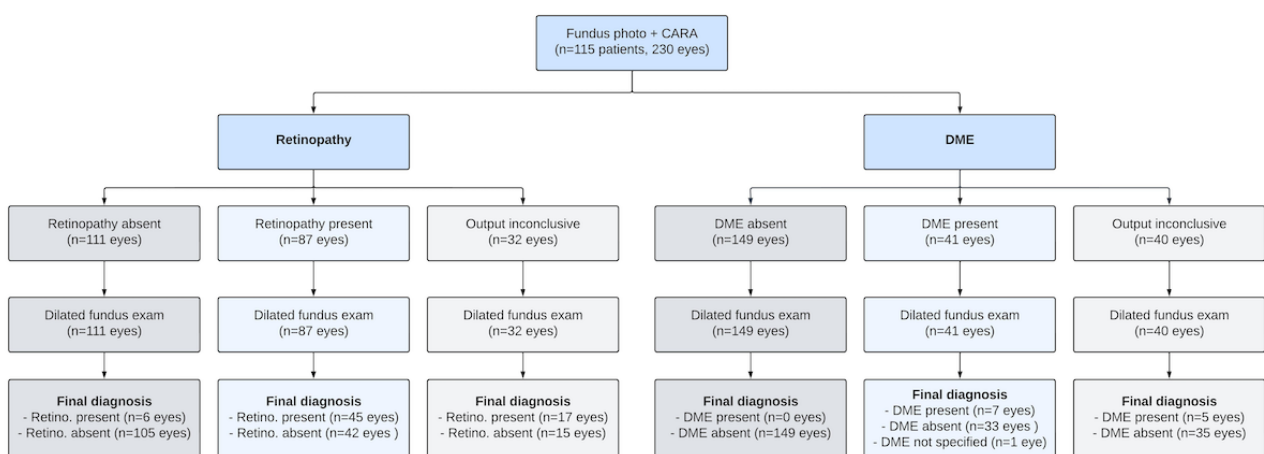


Figure 3. STARD (Standards for Reporting of Diagnostic Accuracy) flowchart at the eye level. Using the Scottish Diabetic Retinopathy Grading Scheme, retinopathy was defined as the presence of any retinopathy (corresponding to R1, R2, R3, and R4). DME was defined as the presence of M1 or M2. CARA: Computer Assisted Retinal Analysis; DME: diabetic macular edema.



Study Cohort

Patient demographics and characteristics are summarized in Table 1. The majority of patients were male (57.4%) with a

mean age of 55.4 (SD 15.6) years. Most patients were White (60.9%). While the type of diabetes was the most often unspecified (43.5%), type 2 diabetes (36.5%) was more common than type 1 (20.0%).

Table 1. Patient demographics for the patients included in the study.

Demographic	Value (N=115)
Sex, n (%)	
Male	66 (57.4)
Female	49 (42.6)
Age (years)	
Mean (SD)	55.4 (15.6)
Range	20-90
Diabetes type, n (%)	
Type 1	23 (20)
Type 2	42 (36.5)
Unspecified ^a	50 (43.5)
Ethnicity, n (%)	
White	70 (60.9)
Middle-Eastern	17 (14.8)
Hispanic	13 (11.3)
Black	11 (9.6)
Unknown	4 (3.5)

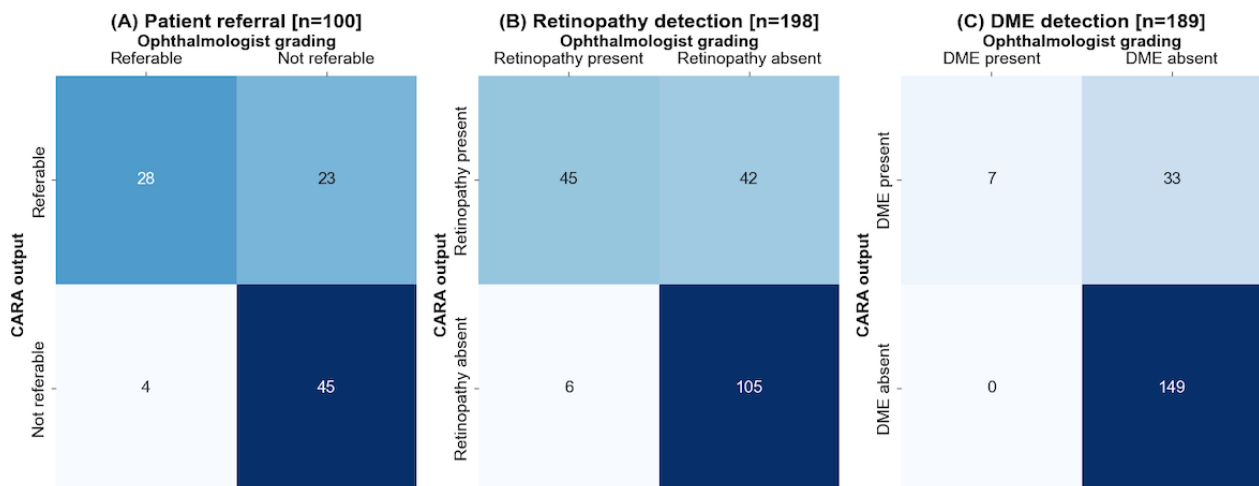
^aDiabetes subtype not specified on the study referral form.

Performance for Detecting Referable Patients

The primary outcome was model performance at the patient level. For the 100 patients with analyzable data, the confusion matrix demonstrating the performance of CARA for referable patient detection is shown in Figure 4A. CARA had a sensitivity of 87.5% (95% CI 71.9-95.0) and a specificity of 66.2% (95% CI 54.3-76.3). There were 4 false negatives with mild background retinopathy (R1) not requiring treatment. No cases of vision-threatening retinopathy were missed.

Inconclusive outputs occurred in 13% (15/115) patients. Reasons varied, including imageability issues like small pupil (n=5), uncooperative patient (n=1), and media opacity (n=2), along with ungradable model outputs due to processing errors or unmet decision thresholds (n=7). Multimedia Appendix 3 describes the demographics of this patient cohort, noting older age in patients with inconclusive outputs (69.5 vs 53.3, $P<.001$). In this group, 8 patients (53.3%) had referable disease, including 3 cases of DME but no cases of severe nonproliferative or proliferative retinopathy.

Figure 4. Confusion matrices showing the discriminative performance of CARA (Computer Assisted Retinal Analysis) at the patient level (referral) and eye level (retinopathy and DME [diabetic macular edema] detection) for patients with analyzable outputs. At the patient level (n=115), 100 had analyzable outputs, and 15 were inconclusive. At the eye level, from 230 eyes, 32 and 40 eyes had inconclusive AI outputs for retinopathy and DME, respectively, and 1 eye lacked an ophthalmologist's grading (DME). This resulted in 198 and 189 eyes with analyzable outcomes for retinopathy and DME, respectively.



Performance for Detecting Disease at the Eye Level

The confusion matrix demonstrating the performance of CARA for retinopathy detection is shown in [Figure 4B](#). CARA had a sensitivity of 88.2% (95% CI 76.6-94.5) and a specificity of 71.4% (95% CI 63.7-78.1). Among the 32 eyes with an inconclusive CARA output for retinopathy, 17 eyes had retinopathy (53.1%) and 15 did not (46.9%).

The confusion matrix demonstrating the performance of CARA for DME detection is shown in [Figure 4C](#). CARA had a sensitivity of 100% (95% CI 64.6-100) and a specificity of 81.9% (95% CI 75.6-86.8). The remaining performance metrics are summarized in [Table 2](#). Among the 40 eyes with an inconclusive CARA output for DME, 5 eyes had DME (12.5%) and 35 did not (87.5%).

Table 2. Summary of the discriminative performance of CARA^a.

Metric	Patient level	Eye level	
	Referable disease	Retinopathy	DME ^b
Sensitivity (95% CI)	87.5 (71.9-95.0)	88.2 (76.6-94.5)	100.0 (64.6-100)
Specificity (95% CI)	66.2 (54.3-76.3)	71.4 (63.7-78.1)	81.9 (75.6-86.8)
PPV ^c (95% CI)	54.9 (41.4-67.7)	51.7 (41.4-61.9)	17.5 (8.8-32.0)
NPV ^d (95% CI)	91.8 (80.8-96.8)	94.6 (88.7-97.5)	100.0 (97.5-100)

^aCARA: Computer Assisted Retinal Analysis

^bDME: Diabetic macular edema.

^cPPV: Positive predictive value (precision).

^dNPV: Negative predictive value.

Economic Analysis

A detailed description of the assumptions and calculations performed for the economic analysis is described in [Multimedia Appendix 4](#). Implementing the AI system for screening could result in a yearly savings of CAD \$245,635 (US \$177,643.23, as of July 26, 2024) or CAD \$49 (US \$35.44) per patient, as shown in [Table 3](#).

These estimates are based on 5000 patients followed annually at the diabetes clinic of the CHUM. The calculations are based on specific costs per patient, the prevalence of diabetic retinopathy among analyzable cases, the percentage of inconclusive outputs by the AI system, and the AI system's specificity. The final row highlights the total cost savings achievable when employing the AI system for screening, demonstrating a significant economic advantage over the current standard of care.

Table 3. Cost comparison between standard of care and AI^a system screening for DR^b screening.

Cost component (CAD \$ ^c)	Standard of care	AI system
Direct screening cost	590,500	150,000
Cost for inconclusive outputs	N/A ^d	76,765
Cost for false positive referrals	N/A	118,100
Total cost	590,500	344,865
Cost savings	N/A	245,635
Average cost per patient	118	69
Average cost savings per patient	N/A	49

^aAI: Artificial intelligence

^bDR: Diabetic retinopathy

^cAs of July 26, 2024, CAD \$1=US \$0.7232)

^dN/A: not applicable.

Discussion

Principal Findings

AI-based telescreening is a promising approach that has the potential to improve community-based screening of DR [27,28]. When effective, AI-based DRTS systems can help reduce

unnecessary examinations of patients without DR, allowing resource allocation to those who need more active management [28]. In this work, we describe the outcomes of real-world implementation of an AI-based DRTS for the screening of DR in a tertiary care hospital in Montreal.

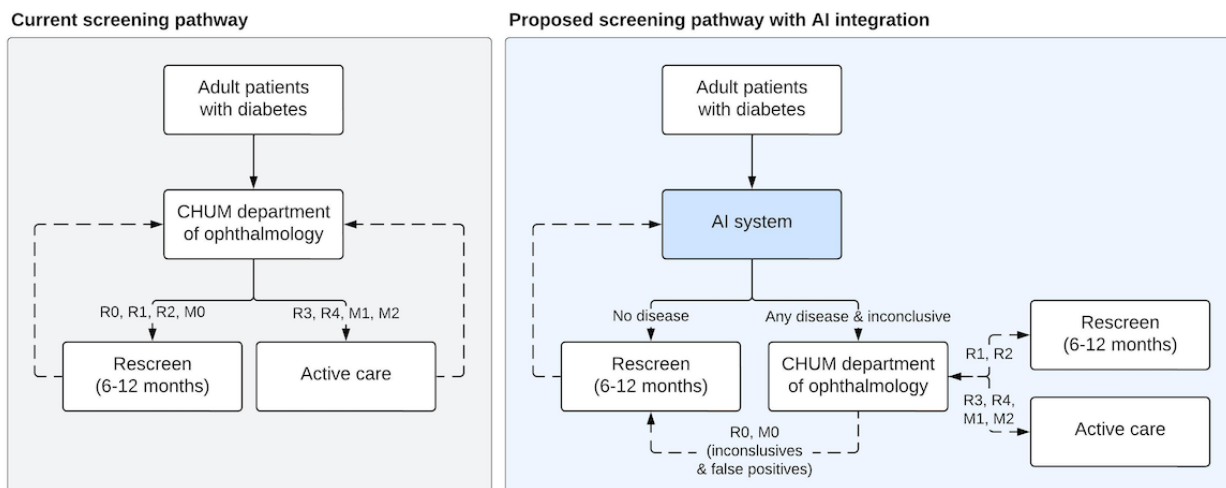
We carried out a silent prospective clinical validation study to assess the performance of the CARA system prior to its implementation in our local population [29]. Our cohort featured an equitable sex distribution and a strong representation of several visible minority groups in Montreal (Black, Middle-Eastern, and Hispanic), which makes our findings generalizable to the local population [30]. The retention rate in this study was notably high, with 86.5% completing the study. While this hints at positive patient acceptance, further studies will be needed for a conclusive understanding of patient acceptability of this technology. The literature suggests a general preference among patients for automated DR screening [31,32], especially when clinicians are involved in supervising the AI system's decisions [33].

CARA identified patients for referral with 87.5% sensitivity and 66.2% specificity. It detected retinopathy with 88.2% sensitivity and 71.4% specificity, and DME with 100% sensitivity and 81.9% specificity. These results meet the 85% sensitivity benchmark obtained by several similar AI systems designed for DR screening [34]. In the real-world setting, a recent meta-analysis showed a pooled sensitivity of 91% (95% CI 87-94) and specificity of 92% (95% CI 88-94) for detecting any retinopathy at the eye level [35]. For vision-threatening DR, which includes DME, sensitivity was 99% (95% CI 95-100) and specificity 92% (95% CI 74-98) [35]. CARA's performance

in detecting retinopathy and DME is on par with other systems in sensitivity but not in specificity, suggesting it effectively identifies disease needing referral but may result in higher false positive rates than desired. The performance gap between CARA and the current state of the art may be partly due to the inherent limitations of traditional feature-based models like CARA [36,37]. In contrast, novel neural network-based models have seen significant improvements in recent years, driven by enhanced architectures and the availability of extensive training data sets [38].

Despite the limited specificity, CARA could significantly reduce the workload of the ophthalmology department at the CHUM by filtering out patients without DR (Figure 5). While false positives can create patient stress and costs [39], since ophthalmologists evaluate these cases, they are unlikely to lead to unnecessary treatments. A similar approach has been successful in the Scottish DRTS program, where all images undergo a first pass through an "autograder" algorithm [40]. With similar specificity to CARA, the Scottish system's automated grading has been shown to be safe [41], and has reduced the burden of manual grading by up to one-third [42]. The current CARA service, in its semiautomated setup, however, still requires grader oversight. Despite that, we expect it could also reduce the burden of manual grading.

Figure 5. Proposed clinical pathway for the artificial intelligence (AI) system implementation. In the current screening pathway, all patients are evaluated by the department of ophthalmology. With AI system integration as a level 1 triage agent, disease-free patients would be filtered out, reducing unnecessary referrals. Only those with referable disease or an inconclusive AI output would be directed to ophthalmology. False positives would re-enter the AI screening pathway, while patients with any disease receive follow-up care in ophthalmology. CHUM: Centre hospitalier de l'Université de Montréal.



Implementing CARA in our screening pathway could yield an estimated first-year saving of approximately CAD \$245,635 (US \$177,643.23) for 5000 patients (CAD \$49 [US \$35.44] per patient), considering direct screening costs and those related to false positive and inconclusive outputs. Of note, the potential costs associated with false negatives (missed cases) were not considered in the analysis due to the lack of such data in our context. The impact of the CARA system on quality-adjusted life years also was not assessed. In the future, we will focus on modeling the screening pathway with decision trees to calculate the incremental cost-effectiveness of our approach, also accounting for false negative cases.

Our study has several limitations and design flaws that have become identifiable in hindsight. Firstly, the provision of the CARA AI system as a service provided to the CHUM introduced several constraints to our study. We lacked control over the model both before and during its implementation. The final outputs were sent to us post grader assessment, without insights into model failures and ungradable images. The semiautomated nature of the pipeline meant that the performance metrics could have been influenced by grader performance, which we are now unable to measure retrospectively. Secondly, we employed dilated fundus examinations as the reference standard instead of using image-assisted evaluations with fundus photographs

and optical coherence tomographic scans. This screening method, common in Quebec, might not capture the full human diagnostic accuracy achievable through a multimodal approach. Due to these limitations, our findings should not be used as a basis to support the approval of CARA as a software medical device. Instead, this study should be viewed as a preliminary trial that helped develop pathways for future, more robust studies involving AI in the DR screening process at CHUM.

Building on those learnings, we have designed a new clinical trial to evaluate the real-world performance of a fully automated deep learning system called NeoRetina (Diagnos Inc) [43]. This new model uses neural networks, which represent the current state of the art in AI. To have a more robust standard reference, in addition to the routine ophthalmological evaluation of DR and DME, masked grading of the same retinal photographs used by NeoRetina will be performed. These images will be assessed for quality and then graded by at least 2 fellowship-trained retina

specialists, with a predefined arbitration process. The new trial will leverage the screening pathways developed in this study and will aim to recruit 630 patients by December 2026.

Conclusions

In conclusion, we report findings from the first real-world implementation study of an AI-based DRTS system in the province of Quebec and possibly in Canada [35]. Large-scale implementation of CARA at CHUM would be expected to result in cost savings and reduced waiting times. However, we are currently investigating more advanced models that we aim to validate more robustly. Once deployed, any model will require routine audits to ensure model performance is maintained, especially with changes in population demographics and disease patterns over time [44,45]. Similarly, ensuring strict information governance policies will be crucial for protecting patient data and responsibly leveraging the benefits of AI systems.

Acknowledgments

We extend our sincere gratitude to Dave Bourgault (Direction du budget et de la performance économique) and Lucie Fuzeau (Centre ambulatoire de médecine) for their instrumental support in the establishment and running of this project. We thank the CHUM ophthalmologists for their expertise in examining patients, and the endocrinologists for their referrals. We are appreciative of the patients who agreed to participate in this study, making this work possible. We thank Martine Henrichon and Martine Loiselle (Department of Ophthalmology of the CHUM) for their invaluable clerical support.

Authors' Contributions

IH, MCT, AB, and AP conceptualized the project. MLDJ provided logistical support. KH, MAR, and DG provided critical expertise on study design. MCT collected the data. IH, MCT, and FA analyzed the results. FA wrote the manuscript, designed the figures, and performed the statistical analyses. CB performed the economic analysis. AYO validated the results. MHD, PAK, and AP provided supervision. All authors reviewed and edited the manuscript. All authors reviewed and approved the final manuscript.

Conflicts of Interest

FA is supported by a Moorfields Eye Charity Springboard Award and a postdoctoral bursary from the Fonds de recherche du Québec - Santé (FRQS). PAK has acted as a consultant for Google, DeepMind, Roche, Novartis, Apellis, and Bitfount, and is an equity owner in Big Picture Medical. He has received speaker fees from Heidelberg Engineering, Topcon, Allergan, and Bayer. He is supported by a Moorfields Eye Charity Career Development Award (R190028A) and a UK Research & Innovation Future Leaders Fellowship (MR/T019050/1). These potential conflicts of interest are unrelated to the present work. Other authors have no conflicts to declare.

Multimedia Appendix 1

Subspecialties, years of experience of the grading ophthalmologists, and numbers of images graded.

[\[PDF File \(Adobe PDF File\), 17 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Patient demographics of the 18 patients who were excluded from the study. After recruitment, 15 patients were lost to follow-up and 3 patients withdrew. There were no statistically significant differences in the patient demographics. *The comparison between the study cohort (n=115) and excluded group (n=18) was performed using Mann-Whitney U test for all continuous variables. For categorical variables, we used the Chi-squared test. **Diabetes subtype not specified on the study referral form.

[\[PDF File \(Adobe PDF File\), 20 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Patient demographics of the 15 patients with inconclusive Computer Assisted Retinal Analysis (CARA) outputs. The cohort of patients with inconclusive CARA outputs was significantly older than the one with analysable outputs ($P < 0.001$). *The comparison between the cohort with analysable outputs (n=100) and the group with inconclusive outputs (n=15) was performed using

Mann-Whitney U test for all continuous variables. For categorical variables, we used the chi-squared test. **Diabetes subtype not specified on the study referral form.

[[PDF File \(Adobe PDF File\), 19 KB-Multimedia Appendix 3](#)]

Multimedia Appendix 4

Cost parameters, assumptions and calculations for the economic analysis. AI: artificial intelligence; FPR: false positive rate: pt: patient.

[[PDF File \(Adobe PDF File\), 22 KB-Multimedia Appendix 4](#)]

References

1. Diabetes in Canada. URL: <https://www.diabetes.ca/advocacy---policies/advocacy-reports/national-and-provincial-backgrounders/diabetes-in-canada> [accessed 2023-11-21]
2. Thomas RL, Halim S, Gurudas S, Sivaprasad S, Owens D. IDF diabetes atlas: a review of studies utilising retinal photography on the global prevalence of diabetes related retinopathy between 2015 and 2018. *Diabetes Res Clin Pract.* 2019;157:107840. [doi: [10.1016/j.diabres.2019.107840](https://doi.org/10.1016/j.diabres.2019.107840)] [Medline: [31733978](https://pubmed.ncbi.nlm.nih.gov/31733978/)]
3. Diabetes Canada Clinical Practice Guidelines Expert Committee, Altomare F, Kherani A, Lovshin J. Retinopathy. *Can J Diabetes.* 2018;42(Suppl 1):S210-S216. [doi: [10.1016/j.jcjd.2017.10.027](https://doi.org/10.1016/j.jcjd.2017.10.027)] [Medline: [29650099](https://pubmed.ncbi.nlm.nih.gov/29650099/)]
4. Liew G, Michaelides M, Bunce C. A comparison of the causes of blindness certifications in England and Wales in working age adults (16-64 years), 1999-2000 with 2009-2010. *BMJ Open.* 2014;4(2):e004015. [FREE Full text] [doi: [10.1136/bmjopen-2013-004015](https://doi.org/10.1136/bmjopen-2013-004015)] [Medline: [24525390](https://pubmed.ncbi.nlm.nih.gov/24525390/)]
5. Felfeli T, Katsnelson G, Kiss A, Plumtre L, Paterson JM, Ballios BG, et al. Prevalence and predictors for being unscreened for diabetic retinopathy: a population-based study over a decade. *Can J Ophthalmol.* 2023;58(4):278-286. [doi: [10.1016/j.jcjo.2022.04.002](https://doi.org/10.1016/j.jcjo.2022.04.002)] [Medline: [35577027](https://pubmed.ncbi.nlm.nih.gov/35577027/)]
6. Boucher MC, El Yamani MEM. Urban diabetic retinopathy teleophthalmology screening: results and impact at the service corridor. *Can J Ophthalmol.* Jun 2019;54(3):359-366. [doi: [10.1016/j.jcjo.2018.06.008](https://doi.org/10.1016/j.jcjo.2018.06.008)] [Medline: [31109477](https://pubmed.ncbi.nlm.nih.gov/31109477/)]
7. Boucher MC, Chahdi HO, El Yamani MEM. Compliance to follow-up care after urban diabetic retinopathy tele-screening. *Can J Ophthalmol.* 2020;55(1 Suppl 1):2-7. [doi: [10.1016/j.jcjo.2019.01.001](https://doi.org/10.1016/j.jcjo.2019.01.001)] [Medline: [31712019](https://pubmed.ncbi.nlm.nih.gov/31712019/)]
8. Felfeli T, Alon R, Merritt R, Brent MH. Toronto tele-retinal screening program for detection of diabetic retinopathy and macular edema. *Can J Ophthalmol.* 2019;54(2):203-211. [doi: [10.1016/j.jcjo.2018.07.004](https://doi.org/10.1016/j.jcjo.2018.07.004)] [Medline: [30975344](https://pubmed.ncbi.nlm.nih.gov/30975344/)]
9. Cao J, Felfeli T, Merritt R, Brent MH. Sociodemographics associated with risk of diabetic retinopathy detected by tele-ophthalmology: 5-year results of the Toronto tele-retinal screening program. *Can J Diabetes.* 2022;46(1):26-31. [doi: [10.1016/j.jcjd.2021.05.001](https://doi.org/10.1016/j.jcjd.2021.05.001)] [Medline: [34144907](https://pubmed.ncbi.nlm.nih.gov/34144907/)]
10. Kanjee R, Dookeran RI. Tele-ophthalmology for diabetic retinopathy in Canada—meeting the needs of a growing epidemic. *Can J Ophthalmol.* 2016;51(3):133-134. [doi: [10.1016/j.jcjo.2016.04.004](https://doi.org/10.1016/j.jcjo.2016.04.004)] [Medline: [27316256](https://pubmed.ncbi.nlm.nih.gov/27316256/)]
11. Artificial intelligence-assisted diabetic retinopathy tele-screening. INESSS. URL: <https://tinyurl.com/4mp8xvha> [accessed 2023-11-21]
12. Optometric services. URL: <https://www.ramq.gouv.qc.ca/en/citizens/health-insurance/optometric-services> [accessed 2023-11-21]
13. Ipp E, Liljenquist D, Bode B, Shah VN, Silverstein S, Regillo CD, et al. EyeArt Study Group. Pivotal evaluation of an artificial intelligence system for autonomous detection of referable and vision-threatening diabetic retinopathy. *JAMA Netw Open.* 2021;4(11):e2134254. [FREE Full text] [doi: [10.1001/jamanetworkopen.2021.34254](https://doi.org/10.1001/jamanetworkopen.2021.34254)] [Medline: [34779843](https://pubmed.ncbi.nlm.nih.gov/34779843/)]
14. Boucher MC, Qian J, Brent MH, Wong DT, Sheidow T, Duval R, et al. Steering Committee for Tele-Ophthalmology Screening, Canadian Retina Research Network. Evidence-based Canadian guidelines for tele-retina screening for diabetic retinopathy: recommendations from the Canadian retina research network (CR2N) tele-retina steering committee. *Can J Ophthalmol.* 2020;55(1 Suppl 1):14-24. [doi: [10.1016/j.jcjo.2020.01.001](https://doi.org/10.1016/j.jcjo.2020.01.001)] [Medline: [32089161](https://pubmed.ncbi.nlm.nih.gov/32089161/)]
15. 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. [FREE Full text] [doi: [10.1038/s41746-018-0040-6](https://doi.org/10.1038/s41746-018-0040-6)] [Medline: [31304320](https://pubmed.ncbi.nlm.nih.gov/31304320/)]
16. Schmidt-Erfurth U, Sadeghipour A, Gerendas BS, Waldstein SM, Bogunović H. Artificial intelligence in retina. *Prog Retin Eye Res.* 2018;67:1-29. [FREE Full text] [doi: [10.1016/j.preteyeres.2018.07.004](https://doi.org/10.1016/j.preteyeres.2018.07.004)] [Medline: [30076935](https://pubmed.ncbi.nlm.nih.gov/30076935/)]
17. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521(7553):436-444. [doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539)] [Medline: [26017442](https://pubmed.ncbi.nlm.nih.gov/26017442/)]
18. Caixinha M, Nunes S. Machine learning techniques in clinical vision sciences. *Curr Eye Res.* 2017;42(1):1-15. [doi: [10.1080/02713683.2016.1175019](https://doi.org/10.1080/02713683.2016.1175019)] [Medline: [27362387](https://pubmed.ncbi.nlm.nih.gov/27362387/)]
19. Akram UM, Khan SA. Automated detection of dark and bright lesions in retinal images for early detection of diabetic retinopathy. *J Med Syst.* 2012;36(5):3151-3162. [doi: [10.1007/s10916-011-9802-2](https://doi.org/10.1007/s10916-011-9802-2)] [Medline: [22090037](https://pubmed.ncbi.nlm.nih.gov/22090037/)]

20. Phan TV, Seoud L, Chakor H, Cheriet F. Automatic screening and grading of age-related macular degeneration from texture analysis of fundus images. *J Ophthalmol*. 2016;2016:5893601. [FREE Full text] [doi: [10.1155/2016/5893601](https://doi.org/10.1155/2016/5893601)] [Medline: [27190636](https://pubmed.ncbi.nlm.nih.gov/27190636/)]
21. Huemer J, Wagner SK, Sim DA. The evolution of diabetic retinopathy screening programmes: a chronology of retinal photography from 35 mm slides to artificial intelligence. *Clin Ophthalmol*. 2020;14:2021-2035. [FREE Full text] [doi: [10.2147/OPTH.S261629](https://doi.org/10.2147/OPTH.S261629)] [Medline: [32764868](https://pubmed.ncbi.nlm.nih.gov/32764868/)]
22. Objectives of training in the specialty of ophthalmology. URL: <https://www.royalcollege.ca/content/dam/documents/ibd/ophthalmology/ophthalmology-otr-e.html> [accessed 2024-06-17]
23. Zachariah S, Wykes W, Yorston D. Grading diabetic retinopathy (DR) using the scottish grading protocol. *Community Eye Health*. 2015;28(92):72-73. [FREE Full text] [Medline: [27418727](https://pubmed.ncbi.nlm.nih.gov/27418727/)]
24. Philip S, Cowie LM, Olson JA. The impact of the health technology board for Scotland's grading model on referrals to ophthalmology services. *Br J Ophthalmol*. 2005;89(7):891-896. [FREE Full text] [doi: [10.1136/bjo.2004.051334](https://doi.org/10.1136/bjo.2004.051334)] [Medline: [15965173](https://pubmed.ncbi.nlm.nih.gov/15965173/)]
25. Flaxel CJ, Adelman RA, Bailey ST, Fawzi A, Lim JI, Vemulakonda GA, et al. Diabetic retinopathy preferred practice pattern®. *Ophthalmology*. 2020;127(1):P66-P145. [doi: [10.1016/j.ophtha.2019.09.025](https://doi.org/10.1016/j.ophtha.2019.09.025)] [Medline: [31757498](https://pubmed.ncbi.nlm.nih.gov/31757498/)]
26. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Standards for Reporting of Diagnostic Accuracy. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ*. 2003;326(7379):41-44. [FREE Full text] [doi: [10.1136/bmj.326.7379.41](https://doi.org/10.1136/bmj.326.7379.41)] [Medline: [12511463](https://pubmed.ncbi.nlm.nih.gov/12511463/)]
27. Ruamviboonsuk P, Tiwari R, Sayres R, Nganthavee V, Hemarat K, Kongprayoon A, et al. Real-time diabetic retinopathy screening by deep learning in a multisite national screening programme: a prospective interventional cohort study. *Lancet Digit Health*. 2022;4(4):e235-e244. [FREE Full text] [doi: [10.1016/S2589-7500\(22\)00017-6](https://doi.org/10.1016/S2589-7500(22)00017-6)] [Medline: [35272972](https://pubmed.ncbi.nlm.nih.gov/35272972/)]
28. Poschkamp B, Stahl A. Application of deep learning algorithms for diabetic retinopathy screening. *Ann Transl Med*. 2022;10(24):1298. [FREE Full text] [doi: [10.21037/atm-2022-73](https://doi.org/10.21037/atm-2022-73)] [Medline: [36660730](https://pubmed.ncbi.nlm.nih.gov/36660730/)]
29. Larson DB, Harvey H, Rubin DL, Irani N, Tse JR, Langlotz CP. Regulatory frameworks for development and evaluation of artificial intelligence-based diagnostic imaging algorithms: summary and recommendations. *J Am Coll Radiol*. 2021;18(3 Pt A):413-424. [FREE Full text] [doi: [10.1016/j.jacr.2020.09.060](https://doi.org/10.1016/j.jacr.2020.09.060)] [Medline: [33096088](https://pubmed.ncbi.nlm.nih.gov/33096088/)]
30. Government of Canada, Canada Statistics. Profile table, Census Profile, 2021 Census of Population - Montréal, Ville (V) Census subdivision, Quebec. URL: <https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/details/page.cfm?Lang=E&SearchText=Montreal&DGUIDlist=2021A00052466023&GENDERlist=1,2,3&STATISTIClist=1&HEADERlist=0> [accessed 2024-02-23]
31. Dow ER, Chen KM, Zhao CS, Knapp AN, Phadke A, Weng K, et al. Artificial intelligence improves patient follow-up in a diabetic retinopathy screening program. *Clin Ophthalmol*. 2023;17:3323-3330. [FREE Full text] [doi: [10.2147/OPTH.S422513](https://doi.org/10.2147/OPTH.S422513)] [Medline: [38026608](https://pubmed.ncbi.nlm.nih.gov/38026608/)]
32. Keel S, Lee PY, Scheetz J, Li Z, Kotowicz MA, MacIsaac RJ, et al. Feasibility and patient acceptability of a novel artificial intelligence-based screening model for diabetic retinopathy at endocrinology outpatient services: a pilot study. *Sci Rep*. 2018;8(1):4330. [FREE Full text] [doi: [10.1038/s41598-018-22612-2](https://doi.org/10.1038/s41598-018-22612-2)] [Medline: [29531299](https://pubmed.ncbi.nlm.nih.gov/29531299/)]
33. Yap A, Wilkinson B, Chen E, Han L, Vaghefi E, Galloway C, et al. Patients perceptions of artificial intelligence in diabetic eye screening. *Asia Pac J Ophthalmol (Phila)*. 2022;11(3):287-293. [FREE Full text] [doi: [10.1097/APO.0000000000000525](https://doi.org/10.1097/APO.0000000000000525)] [Medline: [35772087](https://pubmed.ncbi.nlm.nih.gov/35772087/)]
34. Zhelev Z, Peters J, Rogers M, Allen M, Kijauskaite G, Seedat F, et al. Test accuracy of artificial intelligence-based grading of fundus images in diabetic retinopathy screening: a systematic review. *J Med Screen*. 2023;30(3):97-112. [FREE Full text] [doi: [10.1177/09691413221144382](https://doi.org/10.1177/09691413221144382)] [Medline: [36617971](https://pubmed.ncbi.nlm.nih.gov/36617971/)]
35. Joseph S, Selvaraj J, Mani I, Kumaragurupari T, Shang X, Mudgil P, et al. Diagnostic accuracy of artificial intelligence-based automated diabetic retinopathy screening in real-world settings: a systematic review and meta-analysis. *Am J Ophthalmol*. 2024;263:214-230. [doi: [10.1016/j.ajo.2024.02.012](https://doi.org/10.1016/j.ajo.2024.02.012)] [Medline: [38438095](https://pubmed.ncbi.nlm.nih.gov/38438095/)]
36. Shekar S, Satpute N, Gupta A. Review on diabetic retinopathy with deep learning methods. *J Med Imaging (Bellingham)*. 2021;8(6):060901. [FREE Full text] [doi: [10.1117/1.JMI.8.6.060901](https://doi.org/10.1117/1.JMI.8.6.060901)] [Medline: [34859116](https://pubmed.ncbi.nlm.nih.gov/34859116/)]
37. Hassan B, Raja H, Hassan T, Akram MU, Raja H, Abd-alrazaq AA, et al. A comprehensive review of artificial intelligence models for screening major retinal diseases. *Artif Intell Rev*. 2024;57(5):111. [doi: [10.1007/s10462-024-10736-z](https://doi.org/10.1007/s10462-024-10736-z)]
38. Zhou Y, Chia MA, Wagner SK, Ayhan MS, Williamson DJ, Struyven RR, et al. A foundation model for generalizable disease detection from retinal images. *Nature*. 2023;622(7981):156-163. [FREE Full text] [doi: [10.1038/s41586-023-06555-x](https://doi.org/10.1038/s41586-023-06555-x)] [Medline: [37704728](https://pubmed.ncbi.nlm.nih.gov/37704728/)]
39. Wong RL, Tsang CW, Wong DS, McGhee S, Lam CH, Lian J, et al. Are we making good use of our public resources? The false-positive rate of screening by fundus photography for diabetic macular oedema. *Hong Kong Med J*. 2017;23(4):356-364. [FREE Full text] [doi: [10.12809/hkmj166078](https://doi.org/10.12809/hkmj166078)] [Medline: [28684650](https://pubmed.ncbi.nlm.nih.gov/28684650/)]
40. Styles CJ. Introducing automated diabetic retinopathy systems: it's not just about sensitivity and specificity. *Eye (Lond)*. 2019;33(9):1357-1358. [FREE Full text] [doi: [10.1038/s41433-019-0535-7](https://doi.org/10.1038/s41433-019-0535-7)] [Medline: [31358924](https://pubmed.ncbi.nlm.nih.gov/31358924/)]

41. Philip S, Fleming AD, Goatman KA, Fonseca S, McNamee P, Scotland GS, et al. The efficacy of automated "disease/no disease" grading for diabetic retinopathy in a systematic screening programme. *Br J Ophthalmol*. 2007;91(11):1512-1517. [FREE Full text] [doi: [10.1136/bjo.2007.119453](https://doi.org/10.1136/bjo.2007.119453)] [Medline: [17504851](https://pubmed.ncbi.nlm.nih.gov/17504851/)]
42. Fleming AD, Goatman KA, Philip S, Prescott GJ, Sharp PF, Olson JA. Automated grading for diabetic retinopathy: a large-scale audit using arbitration by clinical experts. *Br J Ophthalmol*. 2010;94(12):1606-1610. [doi: [10.1136/bjo.2009.176784](https://doi.org/10.1136/bjo.2009.176784)] [Medline: [20858722](https://pubmed.ncbi.nlm.nih.gov/20858722/)]
43. Evaluation of NeoRetina Artificial Intelligence Algorithm for the Screening of Diabetic Retinopathy at the CHUM (DR-NeoRetina). URL: <https://classic.clinicaltrials.gov/ct2/show/NCT04699864> [accessed 2024-06-17]
44. Wong CYT, O'Byrne C, Taribagil P, Liu T, Antaki F, Keane PA. Comparing code-free and bespoke deep learning approaches in ophthalmology. *Graefes Arch Clin Exp Ophthalmol*. 2024. [doi: [10.1007/s00417-024-06432-x](https://doi.org/10.1007/s00417-024-06432-x)] [Medline: [38446200](https://pubmed.ncbi.nlm.nih.gov/38446200/)]
45. Youssef A, Pencina M, Thakur A, Zhu T, Clifton D, Shah NH. External validation of AI models in health should be replaced with recurring local validation. *Nat Med*. 2023;29(11):2686-2687. [doi: [10.1038/s41591-023-02540-z](https://doi.org/10.1038/s41591-023-02540-z)] [Medline: [37853136](https://pubmed.ncbi.nlm.nih.gov/37853136/)]

Abbreviations

AI: artificial intelligence
CARA: Computer Assisted Retinal Analysis
CHUM: Centre hospitalier de l'Université de Montréal
DME: diabetic macular edema
DR: diabetic retinopathy
DRTS: DR telescreening
ML: machine learning
STARD: Standards for Reporting of Diagnostic Accuracy

Edited by L Quinlan; submitted 24.04.24; peer-reviewed by FK Malerbi, J Huemer; comments to author 23.05.24; revised version received 28.06.24; accepted 06.07.24; published 03.09.24

Please cite as:

Antaki F, Hammana I, Tessier M-C, Boucher A, David Jetté ML, Beauchemin C, Hammamji K, Ong AY, Rhéaume M-A, Gauthier D, Harissi-Dagher M, Keane PA, Pomp A

Implementation of Artificial Intelligence-Based Diabetic Retinopathy Screening in a Tertiary Care Hospital in Quebec: Prospective Validation Study

JMIR Diabetes 2024;9:e59867

URL: <https://diabetes.jmir.org/2024/1/e59867>

doi: [10.2196/59867](https://doi.org/10.2196/59867)

PMID: [39226095](https://pubmed.ncbi.nlm.nih.gov/39226095/)

©Fares Antaki, Imane Hammana, Marie-Catherine Tessier, Andrée Boucher, Maud Laurence David Jetté, Catherine Beauchemin, Karim Hammamji, Ariel Yuhan Ong, Marc-André Rhéaume, Danny Gauthier, Mona Harissi-Dagher, Pearse A Keane, Alfons Pomp. Originally published in *JMIR Diabetes* (<https://diabetes.jmir.org>), 03.09.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Diabetes*, is properly cited. The complete bibliographic information, a link to the original publication on <https://diabetes.jmir.org/>, as well as this copyright and license information must be included.