Original Investigation

Evaluation of Automated Teleretinal Screening Program for Diabetic Retinopathy

O. Bennett Walton IV, MD, MBA; Robert B. Garoon, MD; Christina Y. Weng, MD, MBA; Jacob Gross, MD; Alex K. Young, MD; Kathryn A. Camero, MD; Haoxing Jin, BS; Petros E. Carvounis, MD, FRCSC; Robert E. Coffee, MD, MPH; Yvonne I. Chu, MD

IMPORTANCE Diabetic retinopathy is a leading cause of blindness, but its detrimental effects are preventable with early detection and treatment. Screening for diabetic retinopathy has the potential to increase the number of cases treated early, especially in populations with limited access to care.

OBJECTIVE To determine the efficacy of an automated algorithm in interpreting screening ophthalmoscopic photographs from patients with diabetes compared with a reading center interpretation.

DESIGN, SETTING, AND PARTICIPANTS Retrospective cohort analysis of 15,015 patients with type 1 or 2 diabetes in the Harris Health System in Harris County, Texas, who had undergone a retinal screening examination and nonmydriatic fundus photography via the Intelligent Retinal Imaging System (IRIS) from June 2013 to April 2014 were included. The IRIS-based interpretations were compared with manual interpretation. The IRIS algorithm population statistics were calculated.

MAIN OUTCOMES AND MEASURES Sensitivity and false-negative rate of the IRIS computer-based algorithm compared with reading center interpretation of the same images.

RESULTS A total of 15,015 consecutive patients (aged 18-98 years); mean 54.3 years with known type 1 or 2 diabetes underwent nonmydriatic fundus photography for a diabetic retinopathy screening examination. The sensitivity of the IRIS algorithm in detecting sight-threatening diabetic eye disease compared with the reading center interpretation was 66.4% (95% CI, 62.8%-69.9%) with a false-negative rate of 2%. The specificity was 72.8% (95% CI, 72.0%-73.5%). In a population where 15.8% of people with diabetes have sight-threatening diabetic eye disease, the IRIS algorithm positive predictive value was 10.8% (95% CI, 9.6%-11.9%) and the negative predictive value was 97.8% (95% CI, 96.8%-98.6%).

CONCLUSIONS AND RELEVANCE In this large urban setting, the IRIS computer algorithm-based screening program had a high sensitivity and a low false-negative rate, suggesting that it may be an effective alternative to conventional reading center image interpretation. The IRIS algorithm shows promise as a screening program, but algorithm refinement is needed to achieve better performance. Further studies of patient safety, cost-effectiveness, and widespread applications of this type of algorithm should be pursued to better understand the role of teleretinal imaging and automated analysis in the global health care system.
Diabetic retinopathy (DR) is the leading cause of newly diagnosed blindness in working-aged adults in the United States (ages 20-74 years). The Centers for Disease Control and Prevention reports that more than 29 million people, or 9.1% of the US population, have diabetes; this number is projected to increase to 44.1 million by 2030. More than 60% of individuals with type 2 diabetes and virtually all patients with type 1 diabetes develop DR within the first 20 years of diagnosis of the disease. Additionally, approximately 4% of people with diabetes in the United States have sight-threatening diabetic eye disease (STDED), defined as severe nonproliferative DR or proliferative DR.

There is evidence that early detection and treatment of DR lead to better visual outcomes in a cost-effective manner. Current guidelines from the American Diabetes Association and the American Academy of Ophthalmology recommend annual DR surveillance examinations; however, only 50% to 65% of patients with diabetes receive these. This gap between screening recommendations and actual screening is caused by multiple factors such as patient awareness, economic constraints, and limited access to health care. The role of the latter will likely grow over the next 15 years, as the population with diabetes is expected to increase by 20% in the developed world, while the number of ophthalmologists is projected to grow by merely 2%. Thus, the development of efficient and accurate screening tools with the capacity to screen large numbers of patients with some degree of automation could be very impactful in reducing the incidence of blindness in patients with diabetes.

The use of single-field nonmydriatic monochromatic digital photography is a sensitive and specific alternative to the traditional reference standard of 7-field stereoscopic color fundus photography and ophthalmoscopy. Studies have demonstrated that trained graders remotely interpreting nonmydriatic images achieved greater sensitivity than ophthalmoscopy and comparable results with 7-field photography. Diabetic retinopathy screening with telemedicine has been shown to be effective in identifying cases of new DR as well as other ophthalmological diseases. Digital screening may be particularly effective in socioeconomically disadvantaged urban populations that have limited access to ophthalmic care. The rate of compliance with annual eye examinations is particularly low among minority populations, with rates of 32% and 49% for African American and Latino American populations, respectively. The implementation of teleretinal imaging for DR screening has been shown to improve the frequency of retinal screening and facilitate specialists’ evaluation of those with potentially blinding disease.

Additionally, teleretinal screening is cost-effective. Medicare studies have demonstrated a cost savings of $36 to $48 per teleretinal-screened patient over unscreened patients who are more likely to present with advanced disease, requiring more invasive and costly treatment. With expanded population models, rural screenings account for an even greater cost savings of $1206 to $1320 per quality-adjusted life-year when factoring in reduction in progression to vision loss among screened vs unscreened patients.

In this study, we evaluated the effectiveness of a newly instituted DR screening protocol using teleretinal screening sites throughout Harris Health System (HHS) centers. Harris Health System is a fully integrated health care system that cares for urban residents of Harris County, Texas, the third most populous county in the United States. It provided more than 1.8 million outpatient clinic visits in 2014 and services a large proportion of the county’s minority population (Table 1). The protocol was started to increase the ophthalmoscopic screening rate of people with diabetes and to reduce the eye clinic patient burden by placing fundus cameras in primary care offices.

The aim of this study was to evaluate the usefulness of an automated computer algorithm–based interpretation of images via the Intelligent Retinal Imaging System (IRIS), using a physician-driven reading center evaluation of the same images as the photographic standards. The Intelligent Retinal Imaging System is the only US Food and Drug Administration–approved system for the storage and management of diagnostic information from computerized diagnostic instruments or systems.

**Methods**

This quality-improvement project was approved by the Quality Improvement Council of HHS; reviewed and exempted by the institutional review boards of Baylor College of Medicine, the University of Texas at Houston, and HHS (thus, patient consent was not required); and complied with the tenets of the Declaration of Helsinki.

Patients carrying a known diagnosis of diabetes (type 1 or 2) in the HHS clinics without a documented eye examination within the last year were identified by their primary care physician and directed to have a retinal screening examination regardless of the presence of symptoms. Harris Health System patients received teleretinal screening in primary care clinics using nonmydriatic fundus photography. Trained personnel at the level of medical assistants followed a standardized protocol to acquire images of each eye using a nonmydriatic, autofocus digital retinography system (CenterVue) set in a darkened room. If image quality was deemed poor because of a small pupillary aperture (<2.5 mm), the eye was dilated with tropicamide, 0.5%, and the image reacquired after pupillary dilatation.
dilation. If the image quality was still insufficient for algorithm analysis, the images were excluded from analysis and patients were referred for evaluation of media opacity. All images were uploaded to a Health Insurance Portability and Accountability Act–compliant, cloud-based server, where they were hosted and processed by IRIS. The IRIS proprietary integrated grading module uses a neural network that has a specific classifier defined and trained by a standard data set of images from the landmark Early Treatment Diabetic Retinopathy Study. The uploaded image is compared with images from the standard data set, and the program searches for the highest probabilistic match. Once this is identified, the patient’s image is classified into a binary “referral” or “observation” category, whereby any patient with severe nonproliferative DR or more advanced disease is classified as a “referral.” A trained optometrist or ophthalmologist located in a reading center in Pensacola, Florida, blinded to the severity grade detected by IRIS programing, subsequently reviewed each image, and the level of retinopathy was manually assigned based on the Early Treatment Diabetic Retinopathy Study classification criteria.25

All patients given a diagnosis of STDED by the reading center were referred to an HHS retina specialist’s clinic within 2 weeks. Those with reading center classifications of moderate, mild, or no nonproliferative DR were referred to follow-up with teleretinal screening within 1 year for continued monitoring of disease progression (Table 2). Personnel within HHS were responsible for contacting the patients and fulfilling the appropriate referrals.

Retinal screening examinations, IRIS-based categorizations, and trained evaluator classification of disease were obtained for all patients with diabetes screened from June 2013 through April 2014. Data for 15 015 patients with diabetes (30,030 eyes) were evaluated. In addition to the main cohort of patients that underwent retinal screening examinations, a retrospective medical record review was performed for 384 patients who were subsequently evaluated in the retina clinic as a result of higher reading center diagnosis.

Study Outcomes
The primary outcome measures were the calculated sensitivity and false-negative rate of the IRIS algorithm used to detect STDED in comparison with the “standard” interpretation of the same images by a trained eye care specialist. Patients who were determined to have STDED were referred to a retina clinic; for this subset of patients, secondary outcome measures included the accuracy of the IRIS algorithm and reading center classification in comparison with physical ophthalmoscopic examination.

Statistical Analysis
Data interpretation and identification of sensitivity and false-negative rates were performed using MedCalc version 15.26

Results
Eighteen thousand twenty-five patients with diabetes had IRIS fundus photographs obtained, and of those, 83.3% (15 015 patients) had fundus photographs acceptable for analysis (Table 1). Based on the reading center diagnosis, the sensitivity of the IRIS algorithm in detecting STDED was 66.4% (95% CI, 62.8%–69.9%) with a false-negative rate of 2% (Table 3). The positive predictive value (PPV) of the IRIS algorithm, defined as the probability that patients with a positive screening test result truly have the disease when compared with the read by specialists in the reading center, was 10.8% (95% CI, 9.6%–11.9%). The negative predictive value of the IRIS algorithm compared with the reading center was 97.8% (95% CI, 96.8%–98.6%; Table 3). For the subset of patients with STDED who were referred to a retinal specialist, when the algorithm was compared with the clinical ophthalmoscopic examination, the algorithm demonstrated 74.8% sensitivity (95% CI, 66.8%–81.8%) with a false-negative rate of 1%. When the reading center interpretation was compared with the clinical ophthalmoscopic examination, the center’s PPV was 62% (95% CI, 55.4%–68.4%). When the IRIS algorithm interpretation was compared with clinical ophthalmoscopic examination, the algorithm’s PPV was 70.3% (95% CI, 65.5%–75.2%; Table 3).

Discussion
In this retrospective study of 15 015 patients with diabetes undergoing teleretinal screening examinations, the IRIS computer-automated algorithm was found to be effective in
identifying patients with STDED. With a sensitivity of 66% and a false-negative rate of 2%, the automated IRIS algorithm shows promise as a screening tool to identify potentially blinding pathology in patients with diabetes who may not have otherwise received ophthalmic care. When a subset of 384 patients with STDED identified by a reading center eye care specialist was evaluated in a retina specialist’s clinic, the ophthalmoscopic examinations correlated with an automated IRIS algorithm sensitivity of nearly 75% and a false-negative rate of only 1% (Table 3). Furthermore, the IRIS algorithm may be superior to the reading center in screening capability, given the impressive negative predictive value of 97.8% and a higher PPV than that of the reading center. This strongly supports the potential utility of automated imaging systems in screening large populations at risk for blinding eye disease.

The IRIS algorithm has comparable sensitivity and predictive values with those of the iGrading, Retmarker, and EyeArt programs, which have been used in the United Kingdom.27,28 While these other programs may detect proliferative disease with more accuracy than the IRIS algorithm, the IRIS program is more capable of excluding patients with no disease, as demonstrated by its referral of a smaller proportion of patients without retinopathy compared with other programs.27,28

To date, most automated teleretinal screening examination studies have not had sample sizes larger than approximately 1000 patients.29 Previous smaller investigations have demonstrated that telemedicine increases the number of patients who are screened for DR and suggest that consecutive screening examinations can be used to safely monitor those with stable DR.30 The results from this study corroborate those of prior studies, albeit on a much larger scale, concluding that there is utility and validity to diabetic teleretinal screening. The IRIS algorithm boasts an impressively low false-negative rate in identifying STDED and therefore very rarely “missed” patients with true vision-threatening disease. While a 66% sensitivity rate for the screening algorithm may initially seem low, it is critical to realize that the test would primarily be used for populations who otherwise would not have received any ophthalmologic care. Screening tests such as this should be considered adjuncts to, rather than substitutes for, the current ocular health provision system.

The noted 10.8% PPV attained when comparing the IRIS algorithm with the reading center interpretation reflects the IRIS algorithm's conservative nature and its tendency to err on the side of “overgrading.” It is important to remember that current screening guidelines recommend that all patients with diabetes be evaluated by an ophthalmologic professional annually. Hence, despite its lower PPV, the screening program still reduces the overall number of patients who need to be seen in the clinic, thereby also reducing the theoretical cost if all patients who should be screened were actually screened in person. Intelligent Retinal Imaging System is actively pursuing algorithm refinement to improve all performance aspects of the program, including its PPV, which will strengthen its screening utility. Furthermore, the IRIS algorithm's PPV is a product of comparing the algorithm interpretation with the “standard” reading center evaluation and thus is affected by reading center errors. When a subset of the reading center images interpreted to have STDED were compared with ophthalmoscopic examination, the PPV of the reading center evaluation was only 62%, while the IRIS algorithm's PPV increased to 70.3%. When the IRIS algorithm and reading center diagnoses are compared, the algorithm's negative predictive value is nearly 98%. A high negative predictive value in a screening program is critical, as it allows patients without STDED to be confidently observed until repeated screening 1 year later. This highlights the program’s potential impact on health care costs, as resources would not be expended on patients who would not necessarily benefit from an in-clinic ophthalmoscopic examination. For HHS, this protocol has been associated with a significant reduction in the patient burden in the eye clinics, opening up appointment slots for patients with more critical diseases.

Given the large number of patients with diabetes and screening sites involved, this study population and setting realistically reflect the variability found in a large urban community with many patients at risk for diabetic eye disease, and adds to the strength of this investigation. The reading center professionals who interpreted images were blinded to algorithm diagnosis as well as patient-identifying data or historical information, allowing for an unbiased interpretation.

Study limitations include image quality. These 2-dimensional images encompassed only a 45° span and lack stereoscopic qualities, making identification of elevation or traction challenging. Even with pupillary dilation, image quality and scope may be insufficient for reliable interpretation, especially in patients with other concurrent pathology or media opacity. If image quality was insufficient for algorithm analysis, the images were excluded from analysis despite the possible presence of DR.

The current US Food and Drug Administration–approved use of the IRIS device is for storing, managing, and displaying data and images for facilitation of ophthalmic care for large populations rather than diagnosing eye disease. Hence, appropriate referral of patients with an STDED diagnosis was dependent on office personnel arranging follow-up with a specialty clinic and the patient actually attending the appointment; both of these actions are susceptible to human error, and thus may skew the data. Last, multiple eye care professionals were involved in the interpretation of screening examinations, and although they were trained in reading images, they were not...
all retinal specialists. Interreader variability and human error may have affected the data.

As presented in this study, more than 15,000 patients with diabetes were screened over the course of 1 year. Validation of patients placed in the “observe” category is needed, and this analysis will take place in subsequent investigations as these eyes are reimaged during their repeated annual teleretinal examination. Further analysis of these patients and their photographs will allow for a better understanding of patients at risk for STDED. Subgroup analysis of patients with certain imaging characteristics or systemic comorbidities may allow for future risk stratification. In addition to physical comorbidities, exploration of patient attitudes and behaviors toward managing their own health care can be investigated as well to identify those at greatest risk for being lost to follow-up and thus presenting later with more severe eye disease.

While the IRIS algorithm is effective at identifying sight-threatening eye disease associated with DR, the program is not currently sensitive enough to detect diabetic macular edema in the absence of retinopathy. Developers of the IRIS are in the process of integrating optical coherence tomography interpretation capability into their algorithm so that patients could be screened for both DR and diabetic macular edema. Future investigations will be aimed at analysis of these data. Although beyond the scope of this study, the IRIS program also successfully recognized the presence of other possible ocular pathology, including retinal vein occlusions, age-related macular degeneration, epiretinal membrane, macular hole, glaucoma, and cataract; these patients were all given a “refer” result. Algorithms to screen for other common pathologies are presently being more carefully evaluated.

Perhaps the most significant future analysis of the data from the teleretinal screening program involves quantitatively estimating how much vision loss was prevented as a result of the screening intervention. While many patients screened in the general population did not have STDED, those patients verified to have severe nonproliferative DR or proliferative DR were treated appropriately with panretinal laser photocoagulation or pars plana vitrectomy as deemed necessary by the supervising retina specialist (Figure). Future studies will compare these patients who underwent intervention to salvage sight from potentially blinding disease with patients who presented with end-stage disease. We expect that the benefits of earlier intervention will prove to be striking.

Finally, the economic benefits of teleretinal screening compared with conventional screening have been well documented. We plan to perform a large-scale cost-benefit analysis of this screening program. The initial trends from this large sample of patients will assist in the identification of the benefits not only in terms of vision, but also in terms of socioeconomic factors, quality-adjusted life-years, and cost to society. Currently, the IRIS cameras are being used at less than 50% of system capacity. As supported by findings of this study, the IRIS program could potentially serve as a direct referral source to the ophthalmology clinic, bypassing the reading center and thereby conveying significant cost savings.

Conclusions

To our knowledge, this is one of the largest automated diabetic teleretinal screening studies in the literature to date. Its high sensitivity and low false-negative rate suggest that the IRIS computer algorithm–based diabetic teleretinal screening program may be an effective alternative to conventional manual image interpretation in screening for one of the world’s leading causes of preventable blindness. In addition to the benefits of having the capability of evaluating large numbers of patients, it has also been well documented that teleretinal screening for DR is cost-efficient and convenient for patients and physicians alike. With the growing number of people with diabetes requiring eye examinations and the relative shortage of eye care professionals who are being trained to provide health care to these patients, this study provides compelling evidence that automated teleretinal screening may be an effective and cost-efficient means of evaluating large populations and identifying those at highest risk for STDED. Further investigations into patient safety, cost-effectiveness, and widespread applications of this type of algorithm should be pursued in the future to better understand the role of teleretinal imaging and automated analysis in the global health care system.
Automated Teleretinal Screening Program for Diabetic Retinopathy

Original Investigation Research

Diabetes Care


Previous Presentation: Data from this article were presented at the 2015 Association for Research in Vision and Ophthalmology Annual Meeting; May 4, 2015; Denver, Colorado.

REFERENCES