Diabetic Retinopathy Assessment Variability Among Eye Care Providers in an Urban Teleophthalmology Program

Yao Liu, MD,[1] Victoria P. Rajamanickam, MS,[2]
Ravi S. Parikh, MD, MPH,[3] Stephanie J. Loomis, MS,[4]
Carolyn E. Kloek, MD,[4] Leo A. Kim, MD, PhD,[4]
Dorothy L. Hitchmoth, OD,[5] Brian J. Song, MD,[4]
Dean C. Xerras, MD,[6] and Louis R. Pasquale, MD[4,7]

Departments of [1]Ophthalmology and Visual Sciences and
[2]Biostatistics and Medical Informatics, University of Wisconsin
School of Medicine and Public Health, Madison, Wisconsin.
[3]Department of Ophthalmology and Visual Science, Yale School
of Medicine, New Haven, Connecticut.
[4]Massachusetts Eye and Ear Infirmary, Harvard Medical School,
Boston, Massachusetts.
[5]White River Junction Veteran's Affairs Medical Center, Veteran's
Health Administration, White River Junction, Vermont.
[6]Harvard Medical School, Massachusetts General Hospital
Chelsea HealthCare Center, Boston, Massachusetts.
[7]Channing Division of Network Medicine, Brigham and Women's
Hospital, Boston, Massachusetts.

Abstract

Background: Teleophthalmology is an evidence-based method for diabetic eye screening. It is unclear whether the type of eye care provider performing teleophthalmology interpretation produces significant variability.

Introduction: We assessed grading variability between an optometrist, general ophthalmologist, and retinal specialist using images from an urban, diabetic retinopathy teleophthalmology program.

Methods: Three readers evaluated digital retinal images in 100 cases (178 eyes from 90 patients with type 2 diabetes). Fisher's exact test, percent agreement, and the observed proportion of positive (P_{pos}) or negative agreement (P_{neg}) were used to assess variability.

Results: Among cases deemed gradable by all three readers ($n = 65$), there was substantial agreement on absence of any retinopathy ($88\% \pm 4.6\%$, $P_{neg} = 0.91-0.95$), presence of moderate nonproliferative or worse retinopathy ($87\% \pm 3.9\%$, $P_{pos} = 0.67-1.00$), and presence of macular edema ($99\% \pm 0.9\%$, $P_{pos} = 0.67-1.00$). There was limited agreement regarding presence of referable nondiabetic eye pathology ($61\% \pm 11\%$, $P_{pos} = 0.21-0.59$) and early, nonroutine referral for a follow-up clinical eye exam ($66\% \pm 8.1\%$, $P_{pos} = 0.19-0.54$). Among all cases ($n = 100$), there was acceptable agreement regarding which had gradable images ($77\% \pm 5.0\%$, $P_{pos} = 0.50-0.90$).

Discussion: Inclusion of multiple types of eye care providers as teleophthalmology readers is unlikely to produce significant variability in the assessment of diabetic retinopathy among high-quality images. Greater variability was found regarding image gradability, nondiabetic eye pathology, and recommended clinical referral times.

Conclusions: Our results suggest that more extensive training and uniform referral standards are needed to improve consensus on image gradability, referable nondiabetic eye pathology, and recommended clinical referral times.

Keywords: telemedicine, ophthalmology, teleophthalmology, e-Health

Introduction

An estimated 29.1 million American adults have diabetes mellitus and among them, approximately one quarter have diabetic retinopathy.[1] Early detection and treatment dramatically reduce the risk of severe vision loss, yet diabetic retinopathy remains the leading cause of blindness among working-age U.S. adults, largely due to low screening rates.[2] However, fewer than 50% of Americans with diabetes mellitus receive yearly diabetic eye screening as recommended by the American Diabetes Association (ADA), the American Academy of Ophthalmology (AAO), and the American Optometric Association (AOA).[2–4]

An urgent need exists for improved access to diabetic eye screening. The rapidly increasing prevalence of diabetes—expected to more than double by the year 2050—has created a

growing demand for screening that cannot be met by the current eye care provider workforce.[5,6] As an alternative to traditional screening methods, where an eye care provider performs a dilated eye examination, teleophthalmology programs using store-and-forward retinal photographs have been well validated to provide high-quality diabetic eye screening, to improve access to care, and to reduce blindness.[7–10] Teleophthalmology is an evidence-based, cost-effective approach to diabetic eye screening that represents one of many dramatic advances in telemedicine on the verge of significant widespread adoption in the United States.[11,12] These retinal photographs can be obtained in the primary care setting where 90% of patients with known diabetes obtain medical care.[13] This technology is particularly well suited for preventing blindness in urban minority populations such as Latinos, who are at greater risk for vision loss due to a higher prevalence of more severe diabetic retinopathy and lower screening rates.[14]

While teleophthalmology has become increasingly accepted for diabetic eye screening in primary care settings, the availability of adequately trained, qualified readers may limit its widespread implementation.[15,16] Prior studies have compared many different types of readers, including general practitioners, nurses, optometrists, and ophthalmologists.[15] Most U.S. teleophthalmology programs currently use eye care providers as readers, but training backgrounds among these readers can vary significantly. It remains unclear as to whether the type of eye care provider performing teleophthalmology interpretation produces significant variability. This study is the first to directly assess variability between an optometrist, a general ophthalmologist, and a fellowship-trained retina specialist for teleophthalmology evaluation of diabetic retinopathy.

Methods

TELEOPHTHALMOLOGY IMAGE ACQUISITION AND PRESENTATION

Three readers assessed a set of masked telemedicine cases from 90 patients (178 eyes) with type 2 diabetes mellitus, who participated in a clinical teleophthalmology program for diabetic eye screening. All images were acquired by a trained ophthalmic photographer from 2008 to 2010 at a single urban primary care clinic affiliated with Massachusetts General Hospital located in Chelsea, Massachusetts, using a nonmydriatic camera (Topcon TRC NW-6S, Topcon Medical Systems, Inc., Tokyo, Japan). Four nonmydriatic, nonstereoscopic digital images were acquired from each eye following the validated Joslin 3-field technique (i.e., 45-degree views of the posterior pole, area nasal to the disc, and area superotemporal to the disc) plus an external photograph.[17] Images

from both eyes were included with the exception of 2 cases with monocular patients in whom their only eye was imaged. To assess intrareader reliability, 10 cases were duplicated and all 3 readers independently assessed 100 masked case vignettes from 90 patients. In addition to retinal photographs, each masked telemedicine case included a brief synopsis of the patient's demographic information and medical history, including age, gender, self-reported ethnicity, diabetes type, diabetes medications, last hemoglobin A1c, and past medical history. Readers viewed all masked telemedicine cases within Microsoft PowerPoint (Microsoft, Inc., Redmond, WA).

READER BACKGROUNDS AND GRADING INSTRUCTIONS

The readers were a Veteran's Affairs (VA) Medical Center optometrist (D.H.), an academic general ophthalmologist (C.K.), and an academic fellowship-trained retina specialist (L.K.). All readers completed the standard teleophthalmology diabetic retinopathy reader training program (an electronic independent study course) instituted by the VA healthcare system before initiation of the study. All were active readers in clinical teleophthalmology programs, but the optometrist had several years more of experience from her prior participation in the VA teleophthalmology program.

Readers had not previously evaluated any of the teleophthalmology images in this study and did not have access to the patients' prior clinical assessments for diabetic retinopathy. All readers were oriented to the navigation of the telemedicine case vignettes and the electronic grading sheet by the same investigator (Y.L.) using a series of five sample cases that were not included in the data analysis. Readers were instructed on how to adjust images as needed to maximize gradability, including adjustment for brightness and contrast, as well as for increasing magnification over any area of interest. Readers were informed that the intent of the study was to assess variability in telemedicine diabetic retinopathy assessments.

The readers were asked to determine whether or not images from each case were of sufficient quality to be gradable for diabetic retinopathy, whether there was any diabetic retinopathy present, the severity of diabetic retinopathy if present, and whether there was referable nondiabetic eye pathology (i.e., glaucomatous optic nerve, advanced macular degeneration, etc.). For each case, readers selected one of the three following options for recommended referral time to a clinical eye examination: (1) urgent/within 1 month, (2) nonurgent/within 2–6 months, or (3) routine/within 1 year. Diabetic retinopathy was evaluated on a three-level severity scale: (1) none, (2) mild nonproliferative, or (3) moderate nonproliferative or worse based on the International

Classification of Diabetic Retinopathy.[18] If images were considered ungradable, then all remaining assessments regarding the absence or presence of retinopathy, macular edema, or nondiabetic eye pathology were categorized as "unable to determine." There were no synchronous, clinical dilated eye examinations or dilated seven-field ETDRS retinal photographs to serve as reference standards because study images were previously acquired in an active clinical teleophthalmology program.[19]

STATISTICAL ANALYSIS

Comparisons of inter-reader and intrareader variability were assessed using Fisher's exact test and percent agreement. In addition, a statistical measure known as the observed proportion of positive (P_{pos}) or negative agreement (P_{neg}) was used in place of kappa (κ) to confirm agreement due to the "kappa paradox," wherein kappa values are artificially low due to the infrequency of retinopathy in this community-based sample.[20,21] Values of 70–90% agreement are generally considered acceptable for inter-reader reliability.[22,23] Percent agreement has been used in similar studies and can be safely used since readers were well-trained eye care providers who were unlikely to be guessing in their assessments.[8,24] Agreement was considered limited if <70%, acceptable if 70% or greater, and substantial if 80% or greater. All statistical calculations were made using Stata software (StataCorp LP, College Station, TX).

ETHICS/IRB APPROVAL

This study was approved by the Human Studies Committee/Institutional Review Board at the Massachusetts Eye and Ear Infirmary and complied with the tenets of the Declaration of Helsinki.

Results

Masked telemedicine cases in this study included patients ($n = 90$) who were predominantly Latino (73.4%) and female (60%) with an average age of 54 years (range: 28–83 years) (*Table 1*). All had a diagnosis of type 2 diabetes mellitus with an average hemoglobin A1c of 7.5% (range: 5.2–12.1%).

Among all three readers in this study, the average percentage of all telemedicine cases ($n = 100$) found to have clinically significant diabetic retinopathy (i.e., moderate nonproliferative or worse) was $3.7\% \pm 1.5\%$ (range: 2.0–5.0%) and the average percentage of cases found to have macular edema was $2.7\% \pm 0.6\%$ (range 2.0–3.0%) (*Table 2*). These values were not significantly different ($p > 0.05$) among the three readers. The frequency of cases with images considered ungradable for diabetic retinopathy was variable—27% by the general ophthal-

Table 1. Patient Demographics ($n = 90$)

Age	54 years (range: 28–83 years, median 52.5 years)
Female	54 (60%)
Ethnicity, %	
Latino/Hispanic	73.4
Caucasian, non-Latino/Hispanic	14.4
African American	8.9
Middle Eastern	2.2
Asian	1.1
Diagnosis of type 2 diabetes mellitus	100%
Average hemoglobin A1c	7.5% (range: 5.2–12.1), 58 mmol/mol (range: 33–109)

mologist, 17% by the optometrist, and 3% by the retina specialist ($p < 0.005$). However, inter-reader agreement regarding gradable images for diabetic retinopathy among all cases ($n = 100$) was acceptable ($77\% \pm 5.0\%$, $P_{pos} = 0.50$–0.90).

To assess inter-reader variability in diabetic retinopathy evaluation, we examined the subset of teleophthalmology cases ($n = 65$) with high-quality images considered gradable for retinopathy by all three readers (*Table 3*). In this subset, there was substantial agreement on: (1) the absence of any retinopathy ($88\% \pm 4.6\%$, $P_{neg} = 0.91$–0.95), (2) the presence of moderate nonproliferative or worse retinopathy (i.e., severity of retinopathy) ($87\% \pm 3.9\%$, $P_{pos} = 0.67$–1.00), and (3) the presence of macular edema ($99\% \pm 0.9\%$, $P_{pos} = 0.67$–1.00). Of note, all cases ($n = 7$) found to have clinically significant diabetic eye disease (i.e., macular edema and/or moderate nonproliferative or worse retinopathy) by the retina specialist were also identified as such by the other two readers. Referral recommendations were compared as follows: agreement on early, nonroutine (i.e., urgent or nonurgent) referral (P_{pos}) and agreement on routine referral (P_{neg}). There was limited inter-reader agreement regarding the presence of referable nondiabetic eye pathology ($61\% \pm 11\%$, $P_{pos} = 0.21$–0.59) and earlier, nonroutine referral for a follow-up clinical eye examination ($66\% \pm 8.1\%$, $P_{pos} = 0.19$–0.54).

Intrareader agreement was more variable in each of the diagnostic and disposition categories (*Table 4*). Overall, there was greater intrareader agreement (80–90%, $P_{pos} = 0.67$–1.00) for the optometrist and retina specialist than for the general ophthalmologist (50–70%, $P_{pos} = 0.60$–0.77) on gradability of images for diabetic retinopathy, severity of

Table 2. Image Gradability, Diabetic Retinopathy Severity, and Macular Edema in All Cases (n=100)

	OPTOMETRIST, %	GENERAL OPHTHALMOLOGIST, %	RETINA SPECIALIST, %	AVERAGE±STANDARD DEVIATION, %	p–values[a]
Ungradable images	17	27	3.0	16±12	**<0.005**
No retinopathy	70	56	89	72±17	**<0.005**
Any retinopathy	13	17	8.0	12±4.5	0.17
Mild retinopathy	11	12	4.0	9.0±4.4	0.09
Moderate or worse retinopathy	2.0	5.0	4.0	3.7±1.5	0.64
Macular edema	2.0	3.0	3.0	2.7±0.6	1.0

[a]Fisher's exact test.

Values shown in bold are statisically significant.

retinopathy, and the presence of macular edema. The optometrist and general ophthalmologist had lesser intrareader agreement (60%, $P_{pos}=0.40–0.71$) than the retina specialist (90%, $P_{pos}=0.86–1.00$) regarding the presence of referable nondiabetic eye disease and earlier, nonroutine referral to a clinical examination.

Discussion

Our study demonstrated substantial inter-reader agreement in diabetic retinopathy assessment among the subset of teleophthalmology cases (n=65) with high-quality images considered gradable by all three readers: an optometrist, a general ophthalmologist, and a fellowship-trained retina specialist. Prior studies have assessed variability among other types of teleophthalmology readers for diabetic retinopathy. Ruamvi-

boonsuk et al. evaluated inter-reader agreement between three retina specialists, three general ophthalmologists, three ophthalmic nurses, and three ophthalmic photographers in single-field digital retinal photographs.[15] Notably, their study did not include optometrists, who form the majority of readers in the VA healthcare system—one of the largest U.S. tele-ophthalmology programs.[11] A 2-day instruction course was provided to all nonphysician readers in that study, but no standardized training was provided to general ophthalmologists or retina specialists. The authors found only fair agreement between and within groups, except for retina specialists who showed moderate to substantial intragroup agreement. Vujosevic et al. demonstrated similar findings among retina specialists.[25] Longer training may be needed for nonphysician readers, who showed moderate to substantial agreement with

Table 3. Inter–Reader Agreement in Cases Gradable for Diabetic Retinopathy by All Readers (n=65)

	OPTOMETRIST VS. GENERAL OPHTHALMOLOGIST		OPTOMETRIST VS. RETINAL SPECIALIST		GENERAL OPHTHALMOLOGIST VS. RETINA SPECIALIST		AVERAGE±STANDARD DEVIATION
	PERCENT AGREEMENT	P_{POS} P_{NEG}	PERCENT AGREEMENT	P_{POS} P_{NEG}	PERCENT AGREEMENT	P_{POS} P_{NEG}	PERCENT AGREEMENT
Any retinopathy	92	$P_{pos}=0.76$ $P_{neg}=0.95$	88	$P_{pos}=0.20$ $P_{neg}=0.93$	83	$P_{pos}=0.15$ $P_{neg}=0.91$	88±4.6
Severity of retinopathy	91	$P_{pos}=0.67$ $P_{neg}=0.99$	88	$P_{pos}=1.00$ $P_{neg}=1.00$	83	$P_{pos}=0.67$ $P_{neg}=0.99$	87±3.9
Macular edema	99	$P_{pos}=0.67$ $P_{neg}=0.99$	99	$P_{pos}=0.67$ $P_{neg}=0.99$	100	$P_{pos}=1.00$ $P_{neg}=1.00$	99±0.9
Referable nondiabetic eye disease	63	$P_{pos}=0.59$ $P_{neg}=0.67$	51	$P_{pos}=0.21$ $P_{neg}=0.66$	68	$P_{pos}=0.22$ $P_{neg}=0.80$	61±11
Recommended referral time	51	$P_{pos}=0.54$ $P_{neg}=0.62$	45	$P_{pos}=0.19$ $P_{neg}=0.60$	74	$P_{pos}=0.33$ $P_{neg}=0.85$	66±8.1

Table 4. Intrareader Agreement in Duplicate Cases ($n = 10$)

	OPTOMETRIST		GENERAL OPHTHALMOLOGIST		RETINA SPECIALIST		AVERAGE ± STANDARD DEVIATION
	PERCENT AGREEMENT	P_{POS} P_{NEG}	PERCENT AGREEMENT	P_{POS} P_{NEG}	PERCENT AGREEMENT	P_{POS} P_{NEG}	PERCENT AGREEMENT
Gradable images	90	$P_{pos} = 0.93$ $P_{neg} = 0.80$	60	$P_{pos} = 0.60$ $P_{neg} = 0.60$	80	$P_{pos} = 0.89$ $P_{neg} = 0.00$	77 ± 15
Any retinopathy	90	$P_{pos} = 1.00$ $P_{neg} = 1.00$	60	$P_{pos} = 0.77$ $P_{neg} = 0.57$	90	$P_{pos} = 0.80$ $P_{neg} = 0.93$	80 ± 17
Severity of retinopathy	80	$P_{pos} = 1.00$ $P_{neg} = 1.00$	50	$P_{pos} = 0.67$ $P_{neg} = 0.94$	90	$P_{pos} = 1.00$ $P_{neg} = 1.00$	73 ± 21
Macular edema	90	$P_{pos} = 0.86$ $P_{neg} = 0.92$	70	$P_{pos} = 0.73$ $P_{neg} = 0.67$	90	$P_{pos} = 0.67$ $P_{neg} = 0.94$	83 ± 12
Referable nondiabetic eye disease	60	$P_{pos} = 0.67$ $P_{neg} = 0.50$	60	$P_{pos} = 0.40$ $P_{neg} = 0.80$	90	$P_{pos} = 0.86$ $P_{neg} = 0.92$	70 ± 17
Recommended referral time	60	$P_{pos} = 0.71$ $P_{neg} = 0.33$	60	$P_{pos} = 0.60$ $P_{neg} = 0.60$	90	$P_{pos} = 1.00$ $P_{neg} = 1.00$	70 ± 17

retina specialists after 1 year of rigorous training in a study by Bhargava et al.[16] While using nonphysician readers may expand reader availability, investment in a year-long training program may not be practical in many U.S. healthcare settings. In this study, all readers were eye care providers who underwent a brief electronic independent self-study training program instituted by the VA healthcare system.

Ruamviboonsuk et al. concluded that retinal specialists were generally the most reliable teleophthalmology readers.[15] It is therefore notable that in our study, all cases identified by the retina specialist as having clinically significant diabetic eye disease (i.e., those with macular edema or moderate nonproliferative or worse retinopathy) were also identified as such by the other two readers. Therefore, if we were to use the retina specialist's interpretation as a reference standard, then the general ophthalmologist and optometrist both had 100% sensitivity in identifying vision-threatening diabetic eye disease for early clinical referral, although they referred additional patients (e.g., 70% and 81% specificity, respectively) often as a result of noting more ungradable images.[18] A previous study comparing the results of in-person clinical diabetic eye examinations between optometrists and general ophthalmologists found that optometrists had a referral sensitivity of 77% and 92% for moderate to severe macular edema and diabetic retinopathy, respectively.[26] Our study demonstrates that telemedicine evaluation by non-retina specialist, eye care providers is highly sensitive for detecting clinically significant, vision-threatening forms of diabetic eye disease

when a standardized training program is used and furthermore, that this appears to exceed clinical thresholds currently accepted for in-person dilated eye examinations.

Since most patients in our study did not have retinopathy, it was unsurprising that there was greater agreement on the absence of any retinopathy (P_{neg}) versus its presence (P_{pos}) and greater agreement on absent or mild retinopathy (P_{neg}, severity of retinopathy) versus moderate or worse retinopathy (P_{pos}, severity of retinopathy) due to the difference in the frequencies of these findings in the community. Similarly, because of the lower frequency of referable nondiabetic eye disease and earlier, nonroutine referral, there was greater agreement on the absence of referable nondiabetic eye disease (P_{neg}) versus its presence (P_{pos}) and greater agreement on routine (P_{neg}, recommended referral time) versus earlier, nonroutine referral (P_{pos}, recommended referral time).

In addition to inter-reader agreement, we also evaluated intrareader agreement by duplicating 10 cases in our dataset. The general ophthalmologist had the lowest intrareader agreement, in part, from having a high percentage (27%) of ungradable image assessments, which led to more frequent assessments of "unable to determine" in each category. The higher frequency of ungradable image assessments by the general ophthalmologist may reflect less experience compared to the retina specialist in evaluating diabetic retinopathy under suboptimal conditions (e.g., through a postoperative gas bubble). Nor did the general ophthalmologist have as much experience with teleophthalmology as the VA optometrist. These data support the existence of a

learning curve in teleophthalmology grading among eye care providers.[15] As a result, teleophthalmology programs should consider having more experienced readers supervise newer readers with protocols for on-going quality assurance, wherein more than one reader assesses a subset of images ("over-reads").

The large range of ungradable rates (3–27%) found among our readers is similar to that of other published studies (4–35%).[8,17,27] Since the absence of retinopathy is more common in our clinical image dataset, the fact that the optometrist and general ophthalmologist noted more ungradable images than did the retina specialist also resulted in their noting fewer cases with no diabetic retinopathy. The wide range of gradability found by readers in our study suggests that improved training modules to standardize criteria for gradability should be developed. This is especially important because unnecessarily high ungradable rates could strain downstream eye care services (e.g., patients with ungradable images are typically referred for clinical evaluation because 50% have significant ocular pathology).[28] In addition, there was limited inter-reader agreement on referable nondiabetic eye disease and recommended clinical referral times. Owsley et al. have found that 44.2% of patients participating in teleophthalmology screening for diabetic retinopathy have nondiabetic eye pathology.[29] Given the high frequency of nondiabetic eye pathology in teleophthalmology programs and the need to regulate the volume of downstream eye clinic referrals, training modules to further standardize referral criteria in these areas are also needed.

One proposed solution to inter-reader variability is automated grading using objective software algorithms. Prior studies have shown that automated grading systems have high sensitivity, but low specificity.[30] A more recent study by Gulshan et al. showed both high sensitivity and specificity for detecting moderate or worse diabetic retinopathy using a deep learning algorithm.[31] However, costs for licensing such software, professional liability, and billing models remain to be determined. Also, these algorithms may be limited in their ability to correctly assess the relative significance of a wide variety of nondiabetic eye pathology.[32] Thus, automated grading systems may be best used in combination with human readers. Future studies should evaluate the optimal implementation of automated grading systems in clinical teleophthalmology programs.

A major strength of our study is that the telemedicine cases came directly from a clinical teleophthalmology program for diabetic retinopathy screening in an urban, primary care clinic. In addition, our study population was unique given the large proportion (73.4%) of Latino patients. This group has been found to have an increased prevalence, incidence, and likelihood of progression of diabetic retinopathy as well as lower rates of diabetic eye screening compared to other ethnic groups.[14,33] Latinos and other high-risk groups with historically low screening rates have the greatest potential to benefit from increased access to diabetic eye screening using teleophthalmology. Thus, it is important for these groups to be well represented in teleophthalmology research studies. Interestingly, the frequency of diabetic retinopathy was substantially lower in our study than that reported previously in a Los Angeles Latino population.[34] This may have resulted, in part, from a lack of patients with type 1 diabetes in our study. In addition, patients with type 2 diabetes obtaining regular care at this community-based primary care clinic may have relatively better long-term glycemic control than those in other parts of the country with less access to care.

Limitations of our study included the absence of a reference standard for the telemedicine cases (e.g., synchronous in-person examination or ETDRS 7-field images) since our cases came directly from a real-world teleophthalmology program.[19] Furthermore, a single individual represented each type of eye care provider in our study. Some of our results may be attributable to unique aspects of each reader and larger studies with more readers may help to determine the generalizability of our findings. Although all readers underwent standardized training, there were no interim assessments to provide readers with additional feedback. Such feedback may be critical for ongoing quality assurance, and the grading of a subset of reads by at least one other experienced reader should be used for quality control purposes in teleophthalmology programs. Another limitation was the small number of duplicates (10%) used to assess intrareader variability. However, this exceeded the proportion of duplicate reads (5%) used as quality control standards for retinal image assessments in large-scale research studies.[35]

In summary, substantial inter-reader agreement was found among three telemedicine readers: an optometrist, general ophthalmologist, and retinal specialist in diabetic retinopathy assessment among cases with high-quality images. This is the first study directly comparing variability in diabetic teleophthalmology evaluation between readers with these training backgrounds. Our study suggests that the inclusion of multiple types of eye care providers as teleophthalmology readers is unlikely to produce significant variability in diabetic retinopathy grading when a standardized training program is used. Training modules to further standardize criteria for image gradability, referable nondiabetic eye pathology, and recommended clinical referral times should be developed. Continued advances in image assessment are

needed to expand access to diabetic eye screening using teleophthalmology to meet the growing demand in communities worldwide.

Disclosure Statement

No competing financial interests exist.

REFERENCES

1. National Diabetes Statistics Report. *Estimates of diabetes and its burden in the United States, 2014.* Atlanta, GA: U.S. Department of Health and Human Services; **2014**.

2. Lee PP, Feldman ZW, Ostermann J, Brown DS, Sloan FA. Longitudinal rates of annual eye examinations of persons with diabetes and chronic eye diseases. *Ophthalmology* **2003;**110:1952–1959.

3. Garg S, Davis RM. Diabetic Retinopathy screening update. *Clin Diabetes* **2009;** 27:140–145.

4. American Optometric Association. Evidence-based clinical practice guideline. In: Evidence-based Optometry Guideline Development Group. *Eye care of the patients with diabetes mellitus.* St. Louis, MO: American Optometric Association, **2014:**2–83.

5. Etzioni DA, Liu JH, Maggard MA, Ko CY. The aging population and its impact on the surgery workforce. *Ann Surg* **2003;**238:170–177.

6. Boyle JP, Thompson TJ, Gregg EW, Barker LE, Williamson DF. Projection of the year 2050 burden of diabetes in the US adult population: Dynamic modeling of incidence, mortality, and prediabetes prevalence. *Popul Health Metr* **2010;**8:29.

7. Chow SP, Aiello LM, Cavallerano JD, et al. Comparison of nonmydriatic digital retinal imaging versus dilated ophthalmic examination for nondiabetic eye disease in persons with diabetes. *Ophthalmology* **2006;** 113:833–840.

8. Ahmed J, Ward TP, Bursell SE, Aiello LM, Cavallerano JD, Vigersky RA. The sensitivity and specificity of nonmydriatic digital stereoscopic retinal imaging in detecting diabetic retinopathy. *Diabetes Care* **2006;**29:2205–2209.

9. Mansberger SL, Sheppler 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.

10. Scanlon PH. The English National Screening Programme for diabetic retinopathy 2003–2016. *Acta Diabetol* **2017;**54:515–525.

11. Kirkizlar E, Serban N, Sisson JA, Swann JL, Barnes CS, Williams MD. Evaluation of telemedicine for screening of diabetic retinopathy in the Veterans Health Administration. *Ophthalmology* **2013;**120:2604–2610.

12. Dorsey ER, Topol EJ. State of telehealth. *N Engl J Med* **2016;**375:154–161.

13. Davidson JA. The increasing role of primary care physicians in caring for patients with type 2 diabetes mellitus. *Mayo Clin Proc* **2010;**85(12 Suppl):S3–S4.

14. Varma R, Choudhury F, Klein R, et al. Four-year incidence and progression of diabetic retinopathy and macular edema: The Los Angeles Latino Eye Study. *Am J Ophthalmol* **2010;**149:752–761.

15. Ruamviboonsuk P, Teerasuwanajak K, Tiensuwan M, Yuttitham K. Interobserver agreement in the interpretation of single-field digital fundus images for diabetic retinopathy screening. *Ophthalmology* **2006;**113:826–832.

16. Bhargava M, Cheung CY, Sabanayagam C, et al. Accuracy of diabetic retinopathy screening by trained nonphysician graders using nonmydriatic fundus camera. *Singapore Med J* **2012;**53:715–719.

17. Cavallerano AA, Cavallerano JD, Katalinic P, et al. Use of Joslin Vision Network digital-video nonmydriatic retinal imaging to assess diabetic retinopathy in a clinical program. *Retina* **2003;**23:215–223.

18. Li HK, Horton M, Bursell SE, et al. Telehealth practice recommendations for diabetic retinopathy, second edition. *Telemed J E Health* **2011;**17: 814–837.

19. Li HK, Hubbard LD, Danis RP, Esquivel A, Florez-Arango JF, Krupinski EA. Monoscopic versus stereoscopic retinal photography for grading diabetic retinopathy severity. *Invest Ophthalmol Vis Sci* **2010;**51:3184–3192.

20. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol* **1990;**43:543–549.

21. Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol* **1990;**43:551–558.

22. Stemler SE. A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Pract Assess Res Eval* **2004;**9: 1–11.

23. Hartmann DP. Considerations in the choice of interobserver reliability estimates. *J Appl Behav Anal* **1977;**10:103–116.

24. McHugh ML. Interrater reliability: The kappa statistic. *Biochem Med (Zagreb)* **2012;**22:276–282.

25. Vujosevic S, Benetti E, Massignan F, et al. Screening for diabetic retinopathy: 1 and 3 nonmydriatic 45-degree digital fundus photographs vs 7 standard early treatment diabetic retinopathy study fields. *Am J Ophthalmol* **2009;**148: 111–118.

26. Hammond CJ, Shackleton J, Flanagan DW, Herrtage J, Wade J. Comparison between an ophthalmic optician and an ophthalmologist in screening for diabetic retinopathy. *Eye* **1996;**10 (Pt 1):107–112.

27. Cavallerano AA, Cavallerano JD, Katalinic P, et al. A telemedicine program for diabetic retinopathy in a Veterans Affairs Medical Center— The Joslin Vision Network Eye Health Care Model. *Am J Ophthalmol* **2005;** 139:597–604.

28. Conlin PR, Asefzadeh B, Pasquale LR, Selvin G, Lamkin R, Cavallerano AA. Accuracy of a technology-assisted eye exam in evaluation of referable diabetic retinopathy and concomitant ocular diseases. *Br J Ophthalmol* **2015;**99:1622–1627.

29. Owsley C, McGwin G, Jr., Lee DJ, et al. Diabetes eye screening in urban settings serving minority populations: Detection of diabetic retinopathy and other ocular findings using telemedicine. *JAMA Ophthalmol* **2015;**133:174–181.

30. Abramoff MD, Folk JC, Han DP, et al. Automated analysis of retinal images for detection of referable diabetic retinopathy. *JAMA Ophthalmol* **2013;**131:351–357.

31. 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.

32. Wong TY, Bressler NM. Artificial intelligence with deep learning technology looks into diabetic retinopathy screening. *JAMA* **2016;**316:2366–2367.

33. Paz SH, Varma R, Klein R, Wu J, Azen SP; Los Angeles Latino Eye Study Group. Noncompliance with vision care guidelines in Latinos with type 2 diabetes mellitus: The Los Angeles Latino Eye Study. *Ophthalmology* **2006;** 113:1372–1377.

34. Varma R, Torres M, Pena F, Klein R, Azen SP; Los Angeles Latino Eye Study Group. Prevalence of diabetic retinopathy in adult Latinos: The Los Angeles Latino eye study. *Ophthalmology* **2004;**111:1298–1306.

35. Blodi BA, Domalpally A, Scott IU, et al. Standard Care vs Corticosteroid for Retinal Vein Occlusion (SCORE) Study system for evaluation of stereoscopic color fundus photographs and fluorescein angiograms: SCORE Study Report 9. *Arch Ophthalmol* **2010;**128:1140–1145.

Address correspondence to:
Yao Liu, MD
Department of Ophthalmology and Visual Sciences
University of Wisconsin School of Medicine
and Public Health
2870 University Avenue, Street 206
Madison, WI 53705

E-mail: liu463@wisc.edu