Accuracy of trained rural ophthalmologists versus non-medical image graders in the diagnosis of diabetic retinopathy in rural China.

McKenna, M., Chen, T., McAneney, H., Angel Vazquez Membrillo, M., Jin, L., Xiao, W., Peto, T., He, M., Hogg, R., & Congdon, N. (2018). Accuracy of trained rural ophthalmologists versus non-medical image graders in the diagnosis of diabetic retinopathy in rural China. *British Journal of Ophthalmology*, *102*(11), 1471-1476. https://doi.org/10.1136/bjophthalmol-2018-312440

Published in:
British Journal of Ophthalmology

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Title Page

Title: Accuracy of trained rural ophthalmologists versus non-medical image graders in the diagnosis of diabetic retinopathy in rural China.

Authors:

Martha McKenna MPH[1,2*], Tingting Chen MD[1*], Helen McAneney PhD[2,3], Miguel Angel Vázquez Membrillo MD[1,4], Ling Jin MS[1], Wei Xiao MD[1], Tunde Peto MD[2], Mingguang He MD[1,5], Ruth Hogg MD[2], Nathan Congdon MD[1,2,6#]

*Co-first authors, #Corresponding Author

Affiliations:

[1] State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, People's Republic of China

[2] Centre for Public Health, Queen's University Belfast, Belfast BT12 6BJ, Northern Ireland

[3] Northern Ireland Methodology Hub, Centre for Public Health, Queen's University Belfast, Belfast BT12 6BJ, Northern Ireland

[4] Retina and Vitreous Department of Instituto Mexicano de Oftalmologia, Queretaro, México

[5] Centre for Eye Research Australia, University of Melbourne, Royal Victorian Eye and Ear Hospital,

Melbourne, Australia

[6] Orbis International, New York, USA

Correspondence to:

Prof Nathan Congdon,

State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Guangzhou, People's Republic

of China, and Centre for Public Health, Queen's University Belfast, Belfast BT12 6BJ, Northern Ireland;

Email: ncongdon1@gmail.com

Telephone: +44 (0)28 9097 8929

Fax: +44 (0)28 9097

Word count- 3,890

Precis

Non-medical personnel in rural China graded images obtained by nurses on patients with high cataract prevalence, meeting standards of the United Kingdom's National Diabetic Eye Screening Programme, and surpassing the performance of trained local ophthalmologists.

Abstract

Background/aims To determine the diagnostic accuracy of trained rural ophthalmologists and non-medical image graders in the assessment of diabetic retinopathy (DR) in rural China.

Methods Consecutive patients with diabetes mellitus were examined from January 2014 to December 2015 at ten county-level facilities in rural Southern China. Trained rural ophthalmologists performed a complete eye examination, recording diagnoses using the United Kingdom National Diabetic Eye Screening Programme (NDESP) classification system. Two field, mydriatic, 45° digital photographs were made by nurses using NDESP protocols, and graded by trained graders with no medical background using the NDESP system. A fellowship-trained retina specialist graded all images in masked fashion and served as reference standard.

Results Altogether, 375 participants (mean age 60 +/-10 years, 48% men) were examined and 1277 images graded. Grader sensitivity (0·82-0·94 (median 0·88)) and specificity (0·91-0·99 (median 0·98)), reached or exceeded NDESP standards (sensitivity 80%, specificity 95%) in all domains except specificity detecting any DR. Rural ophthalmologists' sensitivity was 0·65-0·95 (median 0·66) and specificity 0·59-0·95 (median 0·91). There was strong agreement between graders and the reference standard (kappa=0·84-0·87, P<0·001) and weak-moderate agreement between rural doctors and the reference (kappa= 0·48-0·64, P<0·001).

Conclusion This is the first study of diagnostic accuracy in DR grading among non-medical

graders or ophthalmologists in LMICs. Non-medical graders can achieve high levels of accuracy,

whereas accuracy of trained rural ophthalmologists is not optimal.

INTRODUCTION

China faces the world's largest diabetes mellitus (DM) epidemic, with the prevalence in Chinese adults increasing from <1% in 1980,[1] to 11.6% in 2010,[2]. This represents an estimated 114 million affected persons, the largest number of any country in the world, placing an enormous public health burden on Chinese health services,[3]. Of note, half of the Chinese adult population suffers from pre-diabetes (50.1%), suggesting that DM prevalence will continue to rise,[2].

Diabetic Retinopathy (DR), together with diabetic macular edema (DME) is a major microvascular complication of DM. It is the leading cause of blindness and visual impairment in the working-age population worldwide, except in relatively few countries such as Iceland and the UK, where established screening programmes have led to DR no longer being the most common cause of blindness in this age group,[4]. Approximately 75% of those with DM in China will experience DR/DME within 15 to 20 years of diagnosis,[5]. Early diagnosis and prompt laser treatment prevents blindness in more than 90% of cases, therefore regular and accurate DR screening is critical,[6]. DR is a particular concern in rural China, as a recent population study showed that only 10% of patients with diabetic eye disease are diagnosed and treated in such settings,[7]. Moreover, 43% of those with DM in rural China have DR/DME, 6.3% with vision-threatening disease, both higher than for most other reported populations,[7].

In rural China, 81% of patients with DM do not receive an annual eye examination,[8] as recommended by Chinese guidelines adopted from the American Academy of Ophthalmology,[9] and 69% have never received an eye examination,[8]. This is largely due to uneven distribution of providers, with 80% of Chinese ophthalmologists located in urban areas,[10]. Therefore, it is necessary to explore alternate DR screening approaches in order to reduce demands on the small cadre of rural Chinese ophthalmologists, so capacity is freed for seeing those requiring treatment. To achieve this, critical evaluation of the accuracy of different models of rural DR screening is required.

The aim of the present study is to compare the accuracy of two methods of DR screening: trained rural doctors performing complete eye examinations with grading of DR severity, and trained non-medical graders remotely examining retinal photographs taken by ophthalmic nurses.

METHODS

Study Setting

This study was undertaken in Guangdong Province, Southern China. Data were collected from the Comprehensive Rural Eye Service Training (CREST) programme, an ongoing, five year collaboration

between Zhongshan Ophthalmic Center (ZOC) and ORBIS International,[11]. The ten hospitals involved were chosen for geographic coverage of the province, and their ability to perform independent cataract surgery, indicating readiness for training in management of more complex diseases. The CREST programme aims to create a rural model of comprehensive eye care, including the specific goal of developing capacity to diagnose and treat DR, including DME. From the outset, the project was designed to compare the accuracy and efficiency of two screening methods in the assessment of DR,[12], trained rural ophthalmologists and non-medical graders using digital images.

The protocol for the current study was approved in full by the Ethics Committee at Zhongshan Ophthalmic Center (ZOC), Sun Yat-sen University, Guangzhou, China, and patients, providers and graders taking part provided oral informed consent. The tenets of the Declaration of Helsinki were followed throughout.

Study Design and Participants

This was a retrospective analysis of consecutive patients screened for DR in the CREST programme between 1 January 2014 and 31 December 2015. Eligibility criteria included:

- Age ≥18 years.

- Diagnosis of DM according to the International Diabetes Classification.[13]

- Receipt of a complete eye examination in a rural clinic in the CREST network.

- Existence of a record in the project's Electronic Medical Record (EMR) system.

- Retinal photography completed in at least one eye.

Training and Diagnostic Methods

Rural ophthalmologists

At each CREST facility, at least two ophthalmologists underwent training in identifying and grading both glaucoma and DR involving two phases: two months at ZOC (the project coordinating center), consisting of didactic lectures, observation of clinics and surgery and practical sessions on Goldmann applanation tonometry, slitlamp biomicroscopy, indirect ophthalmoscopy and gonioscopy, proctored by senior sub-specialists; and a hands-on training phase, during which sub-specialists in glaucoma and DR from ZOC travelled to the ten CREST hospitals on a rotating basis to supervise pan-retinal photocoagulation (PRP), focal laser and trabeculectomy surgery. Hands-on training also included a graded checklist review of all phases of a complete eye examination. Areas graded as weak on one training visit would receive additional emphasis on the next; a total of 65 training visits were completed from 1 January 2014 to 31 December 2015.

Rural ophthalmologists' assessment of severity of DR/DME status was determined by complete bilateral

eye examinations on all patients aged 40 years and above, including all persons known to have DM.

Examinations included evaluation of the following: visual acuity, intraocular pressure (IOP), pupillary

response including presence of afferent pupillary defect, gonioscopy, anterior segment , and examination

of the fundus and optic nerve with 90D lens and indirect ophthalmoscopy after pharmacological dilation

of the pupil. Ophthalmologists recorded their diagnoses in the CREST Electronic Medical Record (EMR)

system using the United Kingdom National Diabetic Eye Screening Programme (UK NDESP) grading

framework,[14].

Trained non-medical graders

All retinal images were acquired (3Nethra Classic, Forus, Inc., Bangalore, India) by 17 ophthalmic nurses

at the ten centers after having successfully completed a one-week training provided by technicians from

Forus. Photographers captured two 45° fields per eye at 1024x768 resolution, one image centered on the

macula and the other on the optic nerve, in accordance with UK NDESP DR screening protocol.[15]

Nurses were trained to identify and repeat inadequate images as a quality control measure. Images were

uploaded to a software system created specifically for the project (CSIRO, Perth, Australia) and

transferred to ZOC for grading. Retinal photographs were taken on the same day as the rural

ophthalmologist's diagnosis. Rural doctors were not shown these images and were masked to the results

of DR grading at the time of their clinical assessment.

Four graders based at ZOC with secretarial or data entry backgrounds, and without any history of medical

training, received one-month of intensive grader's training, supervised by a trained ophthalmologist who

completed his grading training at the University of Wisconsin-Madison Fundus Photograph Reading

Center in the United States. Graders' training focused on ocular anatomy, retinal disease, DR signs and

severity based on the UK NDESP DR grading protocols, with a marked examination at the end of the

training course. This was further supplemented by routine quality assurance monitoring exams every three

months for the first year, followed by one to two routine exams per year thereafter, all requiring the

grading of standard image sets downloaded from the Reading Centre at Moorfields Eye Hospital NHS

Foundation Trust's website,[16].

Two graders assessed each photograph independently and graded it according to the UK NDESP National

Grading system. All disagreements, and an additional 10% of all images, were sent for re-grading and

internal quality assurance. These images were referred to an arbitrator, the ophthalmologist who

supervised grader training, and his decision was taken as final. The arbitrator was not the same

ophthalmologist as the reference standard (see below), and evaluated the images in a fully masked,

independent way.

Image quality was rated "adequate" or "inadequate" for grading in accordance with NDESP guidelines.[15]

An image was deemed adequate if it met the following criteria:

- Macular image: Center of fovea >2DD (disc diameters) from edge of image, and vessels

    visible within 1DD of center of fovea.

- Disc image: Complete optic disc >2DD from edge of image, and fine vessels visible on

    surface of disc.

Reference Standard

A reference standard was used to assess the accuracy of assessments by the non-medical graders and local

ophthalmologists. This reference standard was a senior retina specialist with full medical and surgical

retina training to the fellowship level, and was not included in the training of the non-medical graders or

the rural ophthalmologists. DR grading of all retinal images using the grading system and protocols

outlined above was carried out by the reference standard, masked to any prior diagnoses or grades from

the rural physicians or graders.

Statistical Methods

Data were entered into an Excel spreadsheet and data were analyzed using IBM SPSS Statistics v23

software (IBM Corporation, Armonk, NY). Sensitivity and specificity (with 95% Confidence Intervals),

positive and negative predictive values, area under the receiver operating characteristic (ROC) curve and

kappa statistic comparing both rural ophthalmologists and graders with the reference standard were

calculated for the following:

- *Any Diabetic Retinopathy (Any DR)* - defined as grades R1 (mild nonproliferative phase), R2

  (moderate or severe nonproliferative phase), R3a (active proliferative phase) and R3s (stable

  proliferative retinopathy). A grade of R0 was defined as no diabetic retinopathy.

- *Diabetic Retinopathy (DR) requiring treatment* - defined as R3a grade. Grades of R0, R1, R2 or

  R3s were defined as DR not requiring treatment.

- *Diabetic Macular Edema (DME) requiring referral* - defined as grade of M1 (any macular

  edema). A grade of M0 was defined as no diabetic macular edema.

In addition, the non-medical graders' diagnostic accuracy was assessed separately for those cases where

there was no intervention by the arbitrator, and when intervention by the arbitrator was mandated by

protocol, to investigate possible effects of the arbitrator's involvement.

Logistic regression models were used to examine the association between various characteristics of the rural physicians and correct diagnosis of DR requiring treatment, defined as agreement with the reference standard. The Generalised Estimating Equations procedure was used to adjust for the correlation between two eyes of the same patient, and for repeated measures. Variables found to be significant at the 0.05 level in simple univariate models were entered into the multiple regression models, after investigation of potential co-linearity of variables, and further developed through removal of non-significant variables by backward elimination.

RESULTS

Participant Characteristics

Rural Ophthalmologists, Graders and Reference Standard

A total of 28 rural ophthalmologists were involved in the study (mean age 37.0 years, standard deviation (SD) 7.3 years, 77% male), who had a median 10.5 years of experience in ophthalmic practice (inter-quartile range 5.25-19.0) and a median 5.5 years providing cataract surgery independently (inter-quartile range 0.0-15.8). The four graders had a mean age of 28 years (SD 3.37) and 75% were female. The reference standard was a 35-year-old male who had completed retina fellowship training 24 months prior to taking part in the study.

Patients

Between January 2014 and December 2015, 6824 patients were recorded in the CREST EMR system, of

which 537 (7.87%) were identified as having DM. A total of 375 persons (69.8%, mean age 60.3 [10.4]

years, 47.7% male) among this cohort successfully completed retinal photography in at least one eye, and

were eligible for the study. These patients had a median 6-year known duration of DM (inter-quartile

range 2-10 years), and 240 (64.0%) were identified by the rural ophthalmologist as having cataract on

dilated slitlamp examination.

Images, diagnosis and grading

Among 375 participants, 357 (95.2%) had at least one image in each eye, and 18 (4.8%) had images for

only one eye, for a total of 732 eyes and 1277 images, which serve as the basis for all subsequent

analyses. Grading by the reference standard showed 394 (55%) eyes had no DR, 323 (45%) had any DR,

including the 68 (9%) eyes with DR requiring treatment. DME requiring referral was seen in 172 (25%)

of eyes (Table 1). Rural doctors failed to record DR and DME status in the EMR system in 11/750 (1.47

%) and 65/750 (8.67 %) eyes respectively. Graders failed to assign grades for DR and DME status in

24/732 (3.28%) and 56/732 (7.65%) eyes respectively, and identified 312/732 eyes (42.6%) as having

images of inadequate quality (though the majority of these were still gradable for certain characteristics,

see below). Of these, 177/312 (56.7%) were due to fewer than two images being available for an eye,

contrary to NDESP image quality criteria,[15]. Retinal images were unavailable to the reference standard

for 17 (2.3%) eyes. The reference standard failed to assign grades for DR and DME status in 33/750

(4.4%) and 55/750 (7.3%) eyes respectively, though he identified a total of 245/733 (33.4%) images as

not being of adequate quality, of which, of which 103/245 (42%) were due to there being fewer than two

images available for an eye.

Table 1. Diagnostic results of rural doctors, non-medical graders and the reference standard.

	No. (%) of eyes		
	Rural Doctors	**Non-medical Graders**	**Reference Standard**
DR grade			
R0	255 (35)	375 (53)	394 (55)
R1	237 (32)	160 (22)	148 (21)
R2	134 (18)	105 (15)	102 (14)
R3s	6 (1)	4 (1)	5 (1)
R3a	107 (14)	64 (9)	68 (9)
DME grade			
M0	554 (81)	527 (78)	523 (75)
M1	131 (19)	149 (22)	172 (25)

The sensitivity and specificity of rural ophthalmologists compared to the reference standard for the various domains ranged from 0.65-0.95 and 0.59-0.95 respectively, while the non-medical graders had sensitivity and specificity values ranging from 0.82-0.94 and 0.91-0.99 respectively (Table 2). There was weak-moderate agreement between rural doctors and the reference standard in all domains (kappa= 0.48-0.64, *P*<.001), whereas the agreement for non-medical graders with the reference standard was strong (kappa=0.84-0.87, *P*<.001) (Table 2). Positive predictive value (PPV) and Negative predictive value (NPV) of rural ophthalmologists for the various domains ranged from 0.66-0.78 and 0.90-0.96 respectively, whereas those of the non-medical graders ranged from 0.89-0.94 and 0.94-0.99 respectively (Table 2).

Table 2. Accuracy of diagnosis among trained rural ophthalmologists and non-medical graders

compared to the reference standard

	Non-medical Graders	Rural Ophthalmologists
Any DR [a]		
Sensitivity (95% CI)	0.94 (0.92 to 0.97)	0.95 (0.93 to 0.97)
Specificity (95% CI)	0.91 (0.88 to 0.94)	0.59 (0.54 to 0.64)
AUC ROC (95% CI)	0.93 (0.90 to 0.95)	0.77 (0.74 to 0.81)
Kappa (*P* value)	0.85 (<.001)	0.52 (<.001)
DR requiring treatment [b]		
Sensitivity (95% CI)	0.88 (0.80 to 0.96)	0.66 (0.55 to 0.77)
Specificity (95% CI)	0.99 (0.98 to 1.00)	0.91 (0.89 to 0.94)
AUC ROC (95% CI)	0.93 (0.87 to 0.98)	0.79 (0.72 to 0.86)
Kappa (*P* value)	0.87 (<.001)	0.48 (<.001)
DME requiring referral [c]		
Sensitivity (95% CI)	0.82 (0.76 to 0.88)	0.65 (0.57 to 0.72)
Specificity (95% CI)	0.98 (0.97 to 0.99)	0.95 (0.93 to 0.97)
AUC ROC (95% CI)	0.90 (0.87 to 0.94)	0.80 (0.75 to 0.85)
Kappa (*P* value)	0.84 (<.001)	0.64 (<.001)

Abbreviations:

DR=Diabetic Retinopathy, DME= Diabetic Macular Edema, CI= Confidence Intervals, AUC= Area

Under Curve, ROC= Receiver Operating Characteristic

[a]*Any DR,* defined as grades R1, R2, R3a and R3s.

[b]DR requiring treatment, defined as R3a grade.

[c]DME requiring referral, defined as a grade of M1.

In 206 (54.9%) of participants (412/732 = 56.3% of eyes, 716/1277 =56.1% of images), the arbitrator was involved in the grading decision per protocol, either as part of the random 10% quality control sample or due to disagreement between graders (178/643 = 27.7% of eyes for DR and 64/617 =10.4% of eyes for DME status). In 222/357 (62.2%) and 105/332 (31.6%) of arbitrated eyes for DR and DME status respectively, the grade assigned by one or more grader was changed by the arbitrator, and in 135/357 (37.8%) and 227/332 (68.3%), no change was made to DR and DME status respectively. When images where a change in grading was made by the arbitrator were removed from consideration, grader performance was generally lower: sensitivity was 0.75-0.88 (median 0.80), specificity was 0.97-0.99 (median 0.98), AUC ROC was 0.86-0.93 (median 0.88) and kappa was 0.78-0.88 ($P < 0.001$).

In the univariate analysis, younger age, fewer number of years working in ophthalmic practice, shorter length of time providing cataract surgery independently, receiving didactic training in DR management and having fewer outpatient visits annually all were significantly associated with rural ophthalmologists correctly identifying patients with DR requiring treatment (Table 3). Age of the doctor, the length of ophthalmic practice and years providing cataract surgery were highly correlated (Spearman's rho> 0·75),

and so only one of these variables were retained in the multiple regression model to avoid co-linearity.

After consideration of the three possible multiple regression models, age of the doctor provided a slightly

better Quasi Likelihood under Independence Model Criterion (QIC) goodness of fit. In the final multiple

regression models, younger age and receiving didactic training in DR through the programme remained

significantly associated with diagnostic accuracy (Table 3).

Table 3. Logistic regression model of potential predictors of correct diagnosis by rural doctors of diabetic retinopathy requiring treatment

	Simple regression		Multiple regression Model A		Multiple regression Model A reduced	
	Odds Ratio (95% CI)	*P* value	Odds Ratio (95% CI)	*P* value	Odds Ratio (95% CI)	*P* value
Doctor characteristics						
Age (y)	0.939 (0.908 to 0.971)	<.001	0.959 (0.931 to 0.988)	.006	0.949 (0.918 to 0.980)	.002
Male sex	2.062 (0.940 to 4.521)	.071				
Years working in ophthalmic practice	0.959 (0.922 to 0.997)	.035				
Years providing cataract surgery independently	0.961 (0.926 to 0.997)	.034				
Received didactic training in DR through programme	2.884 (1.393 to 5.968)	.004	2.495 (1.193 to 5.215)	.015	2.506 (1.189 to 5.279)	.016
Number of hands-on training sessions	1.008	0.767				

received in managing DR	(0.958 to 1.061)			
Annual number of outpatient visits	0.999	0.011	0.999	0.087
	(0.998 to 1.000)		(0.999 to 1.000)	

Abbreviations: DR=Diabetic Retinopathy CI= Confidence Interval

DISCUSSION

The sensitivity (0.82-0.94 for various domains) and specificity (0.91-0.99) of non-medical graders in this

setting exceeded the DR screening standards of the British Diabetic Association (Diabetes UK)

(sensitivity of 0.80, specificity of 0.95),[17,18] in all domains except "any DR" (specificity, 0.91). In

contrast, rural doctors lacked diagnostic accuracy. Doctors who had received didactic training in DR

management performed significantly better than their counterparts.

Screening by centrally-located, non-medical graders offers many advantages in low resource settings over

reliance on rural doctors. These include allowing the time of the few available trained practitioners to be

focused more fully on provision of care; ease of quality assurance at a central location; and the potential

to improve DR screening quality by reducing missed diagnoses and un-necessary referrals. The explosive

recent growth of DM in China,[2] is likely to make DR screening models which rely on scarce, trained

rural ophthalmologists increasingly untenable.

Shortcomings of image-grading telemedicine models in this low-resource setting must also be

acknowledged, principally the costs of equipment, training and on-going quality assurance. Reliable

internet connections with adequate bandwidth for image transfer may also be a practical limiting factor.

Cost-effectiveness analyses comparing direct clinical detection of DR with telemedicine models are

needed, and will be provided for our programme in future manuscripts. Another barrier to grading by non-medical personnel, in high-resource as well as low-resource settings, is resistance from eye care practitioners themselves. For example, a recent report on DR screening in India found that 79% of programmes used ophthalmologists or optometrists as primary screeners,[19]. However recent guidance from the Royal College of Ophthalmologists has advocated training of non-medical personnel for extended roles to increase capacity of care,[20]. Additional work is needed to explore the accuracy of non-medical personnel in other settings, given the potential advantages outlined above. Our own results show that the inclusion of trained medical expert arbitrators did improve the performance of non-medical graders, as it is designed to. It is expected that non-medical graders would operate under the supervision of trained medical professionals in most settings (at they do for example in the NHS), and thus we do not feel that the expected modest improvement associated with arbitration in any way undercuts our conclusions about the usefulness of non-medical graders in this and potentially other settings.

These results have significant implications for DR screening in low and middle income countries (LMICs). Both VISION 2020,[21] and China's current 5-Year Blindness Prevention Plan,[22] emphasized the need to develop accurate models to detect DR in settings where there are insufficient trained ophthalmologists to provide comprehensive screening. Non-medical graders, which have not been widely used in LMICs, may provide an important part of the solution, if accuracy levels in other settings

are similar to those we observed. Conversely, our results point to very significant limitations in the diagnostic accuracy of rural ophthalmologists in this setting, even after provision of extensive didactic and practical training.

In June 2016, we carried out a review of publications appearing in English with no time limits in Medline, Embase and Web of Science databases. The review terms included rural health or low income settings and DR or retinal screening. Accuracy of grading performed by a number of different types of medical practitioners has been examined in both low and high-resource settings, with variable results. In India, both general practitioners (kappa=0.84) and optometrists (kappa=0.72) showed good agreement with ophthalmologists in detecting 'absent DR', 'non sight-threatening diabetic retinopathy' or 'sight-threatening diabetic retinopathy,' after 25 hours of training over 5 weeks,[23].

In Singapore, a cross-sectional study in a primary care setting against retina specialists as a reference standard showed substantial agreement (kappa=0.66) for non-physician graders and fair agreement (kappa=0.40) for family physicians. Non-physician graders were validated by the University of Melbourne, and underwent a year of rigorous training and annual audits,[24]. In the UK, optometrists using slit-lamp biomicroscopy following accreditation and assessment at a retinal clinic showed slightly lower sensitivity (0.75 vs. 0.80) and identical specificity (0.98) compared to digital photography graded

by an ophthalmologist. The reference standard was an experienced ophthalmologist who examined and

graded the referrals from both screening methods using slit-lamp biomicroscopy,[25]. A study in Scotland

compared the appropriateness of DR referrals among junior doctors who had received formal training in a

DR clinic over 10-12 weeks to junior doctors who were comparable in terms of overall experience, but

had not received specialist ophthalmic instruction. The reference standard used was the initial assessment

of the examining ophthalmologist. Junior doctors with ophthalmic training had a significantly higher ratio

of appropriate to inappropriate referrals,[26]. Orthoptists receiving no additional DR training in an

Australian study performed well in detecting any DR, with a mean sensitivity of 0.86 and specificity of

0.91,[27]. In Norway, 74 optometrists each screening 14 single-field retinal images for the presence or

absence of DR, had sensitivity and specificity of 0.67 and 0.84 respectively, with independent

assessments by two ophthalmologists acting as the reference standard. Investigators concluded that only

5% of optometrists met the existing standards for DR screening programmes (sensitivity=0.80,

specificity=0.95),[28]. Our review found little information on the performance of non-medical graders in

detection of DR,[24] or on grading in LMICs,[23], and no prior publications on non-medical graders in

LMICs.

Strengths of the current study include independent, masked, simultaneous diagnosis and grading of DR on

a large number of consecutive patients with DM by two different groups relevant to screening in LMICs:

rural ophthalmologists and non-medical graders, both compared to a fully-trained retina specialist. The study took place in an on-going programme, so that high-quality training and equipment was available for both ophthalmologists and graders. Images were acquired by local nurses (rather than photographers) on rural patients without prior experience of retinal photography and with high prevalence of cataract, all typical for programmes in under-served areas. A simple, well-documented and widely-used image-grading system, that of the UK NDESP, was employed, and reference made to well-known standards for acceptable grading accuracy, from Diabetes UK.

Limitations of the study must also be acknowledged. It was not practical for the reference standard to examine patients clinically, which precluded direct comparison with the rural ophthalmologists. Had such direct comparison been possible, performance of the rural ophthalmologists may have appeared better. However, grading of high-quality images by a certified clinical specialist is widely used as the gold standard in screening programmes. Thus, we felt that clinical examination of patients by the reference standard would not clearly have added relevant information, while comparison with the assessment of local ophthalmologists, currently the most widely-used screening approach in rural China, is highly relevant. The current study was not planned to provide information on cost-effectiveness of the two approaches, which is important in evaluating appropriate methods of screening for low-resource settings. This will be the focus of future publications. The widely-used research standard for the detection and

classification of DR, stereoscopic color fundus photographs in 7 standard fields as defined by the Early

Treatment Diabetic Retinopathy Study (ETDRS) group,[29], was not employed in the current study.

Although accurate and reproducible, this approach is time-consuming and uncomfortable for

patients,[30]. Two-field retinal photography is more commonly-used in DR screening programmes in

high as well as low-resource settings,[31], and its proven accuracy, relatively low cost and greater

acceptability,[32] made it more suitable for the CREST programme and the current study.

Finally, although the CREST programme included ten county-level hospitals spread across one of China's

larger provinces, application to other settings may be made only with caution. Such county-level

facilities are the major provider of care for rural populations in every region of the country, though quality

of medical services may vary widely. As participants in an NGO programme, these facilities may have

had access to more resources than typical hospitals in rural China. Nonetheless, the study is an important

proof of principle: non-medical graders with modest training could produce highly-accurate results on

images taken by recently-trained nurses of patients who had little experience with fundus photography

and a high prevalence of lens opacity, which would be expected to affect image quality. On the other

hand, the relatively poor performance of local ophthalmologists in a comparatively optimal setting

suggests that their accuracy in more typical rural Chinese settings may also be problematic.

CONCLUSIONS

This is the first study to examine accuracy in the detection of DR among non-medical graders or trained

rural ophthalmologists in LMICs, and it adds to very limited evidence,[24] on the accuracy of non-

medical graders in any setting. Given the very rapid current and projected growth of DM in LMICs, and

the lack of proven models for DR screening in such areas, these results are of great importance for

programme planners. Our findings of very good accuracy among non-medical graders and relatively poor

performance by local ophthalmologists require replication in other settings.

Financial Disclosures: No financial disclosures.

REFERENCES

1. National Diabetes Research Group. Diabetes mellitus survey of 300,000 in fourteen provinces and cities of China. *Chin Med J* 1981;20(11):678–683.

2. Xu Y, Wang L, Jiang H, et al. Prevalence and control of diabetes in Chinese adults. *JAMA* 2013;310(9):948–959. doi: 10.1001/jama.2013.168118.

3. International Diabetes Federation. 2015 China. Available at http://www.idf.org/membership/wp/china. Accessed 23rd August 2016.

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:e004015. doi: 10.1136/bmjopen-2013-004015.

5. Tu Y, Xu L, Wei W, Wang S, Wang YX, Jonas JB. Progression of diabetic retinopathy: the Beijing Eye Study. *Chin Med J (Engl)* 2011;124:3635–3640.

6. Early Treatment Diabetic Retinopathy Study Research Group. Early photocoagulation for diabetic retinopathy: ETDRS report number 9. *Ophthalmology* 1991;98(5):766–785.

7. Wang F, Liang Y, Zhang F, et al. Prevalence of diabetic retinopathy in rural China: the Handan Eye Study. *Ophthalmology* 2009;116(3):461–467. doi: 10.1016/j.ophtha.2008.10.003.

8. Wang D, Ding X, He M, et al. Use of eye care services among diabetic patients in urban and rural China. *Ophthalmology* 2010;117(9):1755–1762. doi: 10.1016/j.ophtha.2010.01.019.

9. American Academy of Ophthalmology Retina/Vitreous Panel. Preferred Practice Pattern Guidelines. Diabetic Retinopathy. San Francisco, CA: American Academy of Ophthalmology; 2016. Available at http://www.aao.org/preferred-practice-pattern/diabetic-retinopathy-ppp-updated-2016 . Accessed 8th November 2016.

10. Yin Q, Hu A, Liang Y, et al. A two-site, population-based study of barriers to cataract surgery in rural China. *Invest Ophthalmol Vis Sci* 2009;50(3):1069–1075. doi: 10.1167/iovs.08-2783.

11. World Diabetes Federation. Chinese rural diabetic eyecare network WDF12-705. Available at http://www.worlddiabetesfoundation.org/projects/china-wdf12-705 . Accessed 20th August 2016.

12. Yan X, Liu T, Gruber L, He M, Congdon N. Attitudes of physicians, patients, and village health workers toward glaucoma and diabetic retinopathy in rural China: A focus group study. *Arch Ophthalmol* 2012;130(6):761–770. doi: 10.1001/archophthalmol.2012.145.

13. International Diabetes Federation. IDF Diabetes Atlas, 7th edn. Brussels, Belgium: International Diabetes Federation, 2015 Available at http://www.diabetesatlas.org . Accessed 10th November 2016.

14. National Health System Diabetic Eye Screening Information for Health Professionals. Available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/256492/22_nhs_diabetic_eye.pdf . Accessed 24th February 2017.

15. NHS Screening Programmes. Diabetic eye screening revised grading definitions (Version 1.3, November 2012). Available at

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/402294/Revised_Grading_Definitions_V1_3_1Nov12_SSG.pdf . Accessed 10th November 2016.

16. Moorfields Eye Hospital, NHS. The Reading Centre. Available at

http://www.readingcentre.org/Training/DR_Course/grading.aspx . Accessed 8th November 2016.

17. Johnston S. Screening for Diabetic Retinopathy [internet]. 2014 Aug 13; Diapedia 7105192813 rev. no. 16. Available from: http://dx.doi.org/10.14496/dia.7105192813.16 . Accessed: 10th November 2016.

18. Ong G, Ripley L, Newsom R, Casswell A. Assessment of colour vision as a screening test for sight threatening diabetic retinopathy before loss of vision. *Br J Ophthalmol* 2003;87:747–752. doi:10.1136/bjo.87.6.747.

19. Indian Institute of Public Health, Public Health Foundation of India, London School of Hygiene & Tropical Medicine (2014) National Diabetic Retinopathy Summit, Available at:

https://blogs.lshtm.ac.uk/iceh/files/2015/01/DR-Summit-Report.pdf. Accessed 8th December 2016.

20. The Royal College of Ophthalmologists. Improving eye health and reducing sight loss – A 'Call to Action' Response. Available at https://www.rcophth.ac.uk/wp-content/uploads/2016/10/RCOphth-Call-to-Action-10th-Sept-2014-Final.pdf . Accessed 8th November 2016.

21. World Health Organisation. VISION 2020: The right to sight. Global initiative for the elimination of

avoidable blindness: Action plan 2006–2011. Available at

http://www.who.int/blindness/Vision2020_report.pdf . Accessed 10th November 2016.

22. China National Blindness Prevention and Treatment. Notice on Distributing "National Plan for the

Prevention and Treatment of Blindness (2012–2015)". Available at

http://www.moheyes.com/News/Details/e2c673729ae247118e173a07550fd4c3 . Accessed 8th November

2016.

23. Verma L, Prakash G, Tewari H, et al. Screening for diabetic retinopathy by non-ophthalmologists: an

effective public health tool. *Acta Ophthalmol* 2003;81:373–377. doi: 10.1034/j.1600-0420.2003.00004.x.

24. Bhargava M, Cheung CY, Sabanayagam C, et al. Accuracy of diabetic retinopathy screening by

trained non-physician graders using non-mydriatic fundus camera. *Singapore Med J* 2012;53:715.

25. Tu KL, Palimar P, Sen S, et al. Comparison of optometry vs digital photography screening for

diabetic retinopathy in a single district. *Eye* 2004;18:3–8. doi:10.1038/sj.eye.6700497.

26. Bibby K, Barrie T, Patterson KR, MAcCuish AC. Benefits of training junior physicians to detect

diabetic retinopathy–the Glasgow experience. *J R Soc Med* 1992;85:326–328.

27. Georgievski Z, Koklanis K, Fenton A, Koukouras I. Victorian orthoptists' performance in the photo

evaluation of diabetic retinopathy. *Clin Experiment Ophthalmol* 2007;35:733–738. doi: 10.1111/j.1442-

9071.2007.01576.x.

28. Sundling V, Gulbrandsen P, Straand J. Sensitivity and specificity of Norwegian optometrists'

evaluation of diabetic retinopathy in single-field retinal images – a cross-sectional experimental study.

BMC Health Serv Res 2013;13:17. doi: 10.1186/1472-6963-13-17.

29. Early Treatment Diabetic Retinopathy Study Research Group. Grading diabetic retinopathy from

stereoscopic color fundus photographs – an extension of the modified Airlie House classification. ETDRS

report no. 10. *Ophthalmology* 1991;98(5):786–806. doi: http://dx.doi.org/10.1016/S0161-6420(13)38012–

9.

30. American Academy of Ophthalmology. Screening for diabetic retinopathy-2014. Available at

http://www.aao.org/clinical-statement/screening-diabetic-retinopathy. Accessed 10th November 2016.

31. Pieczynski J, Grzybowski A. Review of Diabetic Retinopathy Screening Methods and Programmes

Adopted in Different Parts of the World. *European Ophthalmic Review* 2015;9(1):49–55. doi:

http://doi.org/10.17925/EOR.2015.09.01.49

32. Stellingwerf C, Hardus PL, Hooymans JM. Two-field photography can identify patients with vision-

threatening diabetic retinopathy: a screening approach in the primary care setting. *Diabetes Care*

2001;24(12):2086–2090. doi: https://doi.org/10.2337/diacare.24.12.2086