Article published online: 2022-07-07

Original Article

Inter-grader Agreement in the Diabetic Retinopathy Screening Program in Palestine

Riyad George Banayot[1]

[1]St. John Eye Hospital, Jerusalem, Palestine, Israel

Abstract

Aims: This audit aims to assess the quality and accuracy of primary graders in the diabetic retinopathy screening program in the occupied Palestinian territories. **Subjects and Methods:** A retrospective audit of 198 image sets from eight fully trained primary graders in the diabetic retinopathy screening program in the occupied Palestinian territories was performed. An expert grader regraded all images and audited their quality. The interobserver agreement between primary graders and the expert grader and the corresponding Kappa coefficient were determined for overall grading, referable, nonreferable disease, and ungradable disease. The audit standard was set at 80% for interobserver agreement with a Kappa coefficient of 0.7. **Results:** The interobserver agreement was 80% or better for overall outcome, referable, and nonreferable disease. The Kappa coefficient was 0.70 (substantial) for the overall grading results, 0.72 (substantial) for referable disease, 0.86 (almost perfect) for nonreferable disease, and was 0.21 (fair) for ungradable disease. About 82% of pictures showed two positions, and 75% of pictures showed good and adequate quality. **Conclusions:** The audit demonstrates an adequate level of quality and accuracy for primary grading in the diabetic retinopathy screening program in the occupied Palestinian territories.

Keywords: Diabetic screening, grader, interobserver agreement, kappa, Palestine

Introduction

From an epidemiological perspective, and from data collected from the Diabetic Retinopathy Study[1] and the Early Treatment Diabetic Retinopathy Study,[2] it was found that diabetic patients are usually symptom-free and retinopathy may be well advanced before visual deterioration is noted. Diabetic retinopathy study showed that photocoagulation reduced the 2-year incidence of severe visual loss by more than half in the eyes with proliferative diabetic retinopathy, both with and without high-risk characteristics. The focal treatment was carried out in those eyes with macular edema. Early treatment diabetic retinopathy study recommended that scatter photocoagulation is not recommended for eyes with mild or moderate nonproliferative retinopathy. When retinopathy is more severe, scatter photocoagulation should be considered. Results demonstrated that, for eyes with macular edema, focal photocoagulation is effective in reducing the risk of moderate visual loss. The argument that patients are generally symptom-free when they should receive preventive treatment is a strong argument for establishing a screening program.

Access this article online

Quick Response Code:

Website: www.ijmbs.org

DOI: 10.4103/ijmbs.ijmbs_67_18

The British Diabetic Association in 1995 established an understanding that a screening test for diabetic retinopathy should have a minimum specificity of 80% and a specificity of 95%. Retinal photography has proved highly effective, achieving sensitivities, and specificities of 89% and 86%, respectively.[3]

Studies have shown that screening programs using digital retinal images taken with or without dilation may enable early detection of diabetic retinopathy along with an appropriate referral.[4] Consequently, screening for diabetic retinopathy should be an effective technique in the prevention of vision loss and thus represents good clinical practice.

The Irish National Diabetic Retinal Screening Program[5] recommends for internal quality assurance that 10% of disease-negative cases (normal) and all disease-positive cases

Address for correspondence: Dr. Riyad George Banayot, St. John Eye Hospital, Sheikh Jarrah, P.O. Box 19960, Jerusalem 91198, Palestine, Israel.
E-mail: riyadbanayot@gmail.com

How to cite this article: Banayot RG. Inter-grader agreement in the diabetic retinopathy screening program in Palestine. Ibnosina J Med Biomed Sci 2018;10:193-7.

should be regraded independently by at least two graders. Intergrader agreement reports based on these results provide proof of the accuracy of grading.

Subjects and Methods

In 2012, St. John of Jerusalem Eye Hospital and in partnership with United Nations Relief and Works Agency (UNRWA) initiated a 3-year screening, treatment, and management program for diabetic retinopathy among diabetic patients in East Jerusalem and the West Bank including the refugee population of the Southern districts of the West Bank (Bethlehem and Hebron). The program consisted of training UNRWA health professionals (nurses) to use the Canon CR-2 retinal camera to capture images (2 per eye), acquired in meiosis, and grade those images according to Diabetic Retinopathy Screening Program in the Occupied Palestinian Territories standard become primary graders. "Diabetic retinopathy screening-training program" for UNRWA nurses consisted of a 3-week course (theoretical: 15 h and practical: 29 h). Three days were dedicated for screening patients (simulated screening). All participants started a scheduled supervised screening immediately after the course working in pairs. This schedule was adhered to for a minimum of 4 months. A unified DRS-OPT screening form are used by all graders. Grading levels were as follows: no retinopathy (R0), background retinopathy (R1), preproliferative retinopathy (R2), proliferative retinopathy (R3), no maculopathy (M0), and maculopathy (M1). Any retinopathy (preproliferative, proliferative, and other findings, e.g., vitreous hemorrhage and end-stage disease) detected by a primary grader (R1, R2, R3, and M1) was sent for secondary grading performed by another grader. If there was any disagreement between the primary and secondary grader, the images were sent to arbitration, which was performed by an ophthalmologist. The form consisted of two parts; one part is filled by the primary grader/screener and the ophthalmologist fills the second part when patients are referred for further management. The program was directed toward reducing the proliferation of preventable diabetic retinopathy in the occupied Palestinian territories. It was estimated that 40,000 patients would be screened over the 3 full years of the project (all diabetic patients registered at UNRWA clinics in East Jerusalem and the West Bank). The primary objective of the screening component of the project is to detect the maximum number of cases of sight-threatening retinopathy and refer them for further examination and management by an ophthalmologist while retaining those with nonsight-threatening disease under periodic review. St. John Eye Hospital took the role of training, implementing, and quality control of screening, while UNRWA implemented the program in their clinics. At the end of the program, UNRWA will own the program and continue screening their patients at their clinics.

This report was conducted by the lead ophthalmologist at SJEH on behalf of UNRWA as part of their internal quality control aiming to reduce error probability, ensuring that errors are dealt with competence and provide help to screeners, and UNRWA to improve service delivery. The work done by Patra *et al*. represented a simple and straightforward procedure that can be repeated every year and it did fit UNRWA's requirements. We used their work as the framework for this audit and applied their procedures to UNRWA's data.[6]

Two hundred screening sheets and their corresponding fundus photos of UNRWA patients, screened as part of the DRS-OPT program between April and October 2014, were chosen randomly. Two image sets were discarded; one due to duplication and the other due to lack of images. A retrospective audit of 198 image sets of 198 patients graded by eight fully trained primary graders in the DRS-OPT was carried out. Four image sets contained images of one eye only, resulting in 392 eyes included in the audit. Ethical approval and permission to conduct the audit were obtained from UNRWA Ethics Committee. Written informed consent was not obtained since the confidentiality of the study was maintained by concealing the names of patients on archived data sheets. Primary grading was performed for each eye according to the DRS-OPT grading standard. The grading referral outcome was recorded for each image set consisting of four images, two of each eye.

Three hundred and ninety-two eyes of 198 patients that were graded either as disease-negative or positive by the eight primary graders were regraded by a single expert grader blinded to both the primary grading results and to the identity of patient and primary grader. The 392 photos were also audited for picture quality; pictures must be focused, each eye must have two pictures (the first image must be fovea-centered with retinal vessels clearly visible within one disc diameter of center of fovea, and the second image must be disc-centered with fine retinal vessels clearly visible on surface of disc, and vessels visible across >90% of image). The expert grader was an experienced ophthalmologist with a special interest in diabetic eye disease and the clinical lead for the DRS-OPT program.

In accordance with standard objectives 4 and 5 of the DRS-OPT program, screeners/graders are required to ensure photographs are of adequate quality (percentage of ungradable patients in at least one eye <10%), and that grading is accurate. The interobserver agreement was defined as the level of agreement between the primary graders and the expert grader in grading both eyes of the patient (all four images within an image set). The interobserver agreement was determined for overall grading outcome as well as nonreferable, referable, and ungradable subcategories. Nonreferable disease included all normal images (R0M0) plus those graded R1M0 (mild nonproliferative retinopathy without maculopathy). Referable disease or positive recalls were graded as R2, R3, M1, or U; R2 (moderate and severe nonproliferative retinopathy), R3 (proliferative retinopathy and preretinal fibrosis ± tractional retinal detachment), M1 (sight-threatening maculopathy: lesion in the center of the macula (i.e., exudates, hemorrhages, and microaneurysms), and U (ungradable). An ungradable image was one that failed

to meet the definition of adequate quality. Positive recalls and ungradable images (technical recalls) were referred for a dilated fundus examination by an ophthalmologist for further assessment and management.

The Kappa coefficient for overall grading referral outcome and each subcategory was calculated as a measure of reliability of the interobserver agreement values. The interobserver agreement and the Kappa coefficient were calculated using the website vassarstats.net.

Published data for intergrader/interobserver agreement of grading referral outcomes within the diabetic retinopathy program are very scarce; therefore, the audit standard for interobserver agreement was set arbitrarily at 80%. A Kappa coefficient of 0.7 was decided according to the results from a previous paper.[7]

Results

Of the 392 eyes screened for picture quality, 82% of pictures showed two positions, 75% of pictures were of good and adequate quality, and 24% were of inadequate quality. Of the 198 image sets screened in the audit, 73.2% (145/198) and 72.2% (143/198) were found to be normal or disease-negative by the primary graders and the expert grader, respectively.

The primary graders and the expert grader agreed on the DRS-OPT retinopathy grade in 314 of 392 eyes, with an interobserver agreement of 80.1%. The primary graders and the expert grader agreed on grading referral outcomes in 174 of 198 image sets of 198 patients. The interobserver agreement for overall grading outcome was 87.9% [Tables 1 and 2].

The interobserver agreement was 87.9% for the overall grading referral outcome with a Kappa coefficient of 0.70.

The interobserver agreement for nonreferable retinopathy was 98.6%, with a Kappa of 0.86. The expert found nonreferable disease in 143 patients, and the grader found nonreferable disease in 145 patients. Both expert and grader agreed on 133 image sets as nonreferable.

The interobserver agreement for referable retinopathy was 85.7%, with a Kappa of 0.72. The expert found referable disease in 42 patients, and the grader found referable disease in 49 patients. Both expert and grader agreed on 38 image sets as referable disease.

The interobserver agreement for ungradable images was 30.8% with a Kappa coefficient of 0.21 [Table 3]. The expert graded 13 image sets as ungradable compared with only four image sets by the primary graders, agreeing on only three image sets Table 2. There were eight image sets that were graded by the expert grader as normal, and two image sets were graded as referable disease.

Discussion

This audit was initiated by UNRWA and undertaken as part of the internal quality assurance procedure of the DRS-OPT program. The main aim was to establish the accuracy with which fully trained primary graders in the program were able to detect referable, nonreferable, and ungradable disease when compared with a gold standard (i.e., expert grader).

The national diabetic retinal screening program[5] and the UK national screening committee[8] recommend that, as part of the internal quality assurance system, 10% of disease-negative cases and 100% of disease-positive cases should be regraded independently. In this audit, we decided to include 100% of all screen-positive and screen-negative cases.

The interobserver agreement was better than the audit standard for the overall grading referral outcome, and also better than the audit standard for referable and nonreferable disease. The Kappa coefficient achieved the audit standard of 0.7 for the overall outcome and was better than the audit standard for the nonreferable and referable disease. The interobserver

Table 1: Interobserver agreement between the primary graders and the expert grader for 392 eyes

		Expert grader								
		Normal	Early NPDR no maculopathy	Positive recall maculopathy		Positive recall M/S NPDR-PDR maculopathy		Positive recall	Technical recall	Total
		R0M0	R1M0	R0M1	R1M1	R2M0	R2M1	R3M1	U	
Primary grader	R0M0	**254**	2	1	2				13	272
	R1M0	5	**17**	3	4		1	1	2	33
	R0M1	6		**1**						7
	R1M1	3	1	3	**15**		1	2	1	26
	R2M0	1	2		1	**3**	5	1		13
	R2M1	1		1	7		**18**		2	29
	R3M1						2	**3**		5
	U	3						1	**3**	7
	Total	273	22	9	29	3	27	8	21	**392**

The bold numbers indicate the numbers of eyes/image sets with agreement between primary graders and expert grader. R0: Normal, R1: Mild nonproliferative retinopathy, M0: No maculopathy, R2: Preproliferative retinopathy, R3: Proliferative retinopathy, M1: Sight-threatening maculopathy, U: Ungradable, PDR: Proliferative diabetic retinopathy, NPDR: Nonproliferative diabetic retinopathy

Table 2: Interobserver agreement between the primary graders and the expert grader for 198 image sets

		Expert grader			
		Normal	Positive recall	Technical recall	Total
Primary grader	Normal	**133**	4	8	145
	Positive recall	9	**38**	2	49
	Technical recall	1	0	**3**	4
	Total	143	42	13	198

The bold numbers indicate the numbers of eyes/image sets with agreement between primary graders and expert grader. Positive recall indicates referable disease (R2, R3, and M1); Technical recall indicates ungradable images. R2: Preproliferative retinopathy, R3: Proliferative retinopathy, M1: Sight-threatening maculopathy

Table 3: Audit results for overall agreement of grading referral outcomes and the three subcategories

	Agreement (%)	κ	Reliability
Overall (grading referral outcome)	87.9	0.7	Substantial
Referable disease (positive disease)	85.7	0.72	Substantial
Nonreferable disease (normal and R1)	98.6	0.86	Almost perfect
Ungradable disease (technical recall)	30.8	0.21	Fair

agreement for ungradable disease was 30.8% with a Kappa coefficient of 0.21.

It must be noted that the Kappa coefficient, while being a measure of the reliability of interobserver agreement, is dependent on the prevalence of the disease of interest. Lower Kappa values for relatively unusual findings may not necessarily reflect poor agreement.

Data on the prevalence of diabetes mellitus and diabetic retinopathy are scarce in the Palestine. One published study claims that the prevalence of diabetes in men and women aged 30–65 years were 11.3% and 13.9% in rural and urban populations, respectively,[9] while another claim that the prevalence of self-reported diabetes mellitus in ≥50-year-old patients was 26.4%.[10] There is no data on the prevalence of diabetic retinopathy in Palestine. Data collected from the diabetic retinopathy screening program should provide light onto this challenge.

From the clinical perspective, any discrepancy in grading between the first and second graders is subjected to arbitration grading, which is usually undertaken by an ophthalmologist with experience in diabetic retinopathy. The main concern would relate to disease-negative image sets, of which only 10% are regraded as recommended by the DRS-OPT screening program. In this audit, therefore, we focused on the eight image sets which were graded normally by the primary graders and therefore would not normally be regraded. Those eight patients were recalled and were examined by an ophthalmologist.

The audit recommendations are (1) to provide refresher training to all primary graders in the program, with an emphasis on identifying ungradable images, and (2) to repeat the audit after 12 months with a larger number of image sets and using the results of this audit as a standard for comparison.

Methodology features of the audit include (1) all disease-negative and disease-positive image sets were regraded, (2) the interobserver agreement was a comparison of the outcome between primary graders and expert grader, and (3) the expert grader was blinded to both the results and the identity of the primary grader. The main drawback of the audit was the small sample size.

Conclusions

This audit has demonstrated an acceptable level of quality and accuracy of primary grading in the DRS-OPT program. It also provides UNRWA with baseline standards against which future interobserver agreement can be measured for quality assurance within the DRS-OPT program.

The main recommendation of the audit is to conduct regular auditing and coaching for primary graders that include discussion of samples of audited pictures.

Authors' contributions

The author is responsible for the conception of the study, data collection and analysis and for drafting, revising, and approving the final version of the manuscript.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

Compliance with ethical principles

Ethical approval was granted by the St John's Hospital Ethical Committee. All data were de-identified before analysis.

References

1. Photocoagulation treatment of proliferative diabetic retinopathy: The second report of diabetic retinopathy study findings. Ophthalmology 1978;85:82-106.
2. Early photocoagulation for diabetic retinopathy. ETDRS report number 9. Early treatment diabetic retinopathy study research group. Ophthalmology 1991;98:766-85.
3. Harding SP, Broadbent DM, Neoh C, White MC, Vora J. Sensitivity and specificity of photography and direct ophthalmoscopy in screening for sight threatening eye disease: The Liverpool diabetic eye study. BMJ 1995;311:1131-5.
4. American Academy of Ophthalmology Retina/Vitreous Panel. Preferred Practice Pattern® Guidelines. Diabetic Retinopathy. San Francisco, CA: American Academy of Ophthalmology; 2016. Available from: http://www.aao.org/ppp. [Last accessed on 2018 Sep 18].
5. Diabetic Retina Screen - The National Diabetic Retinal Screening Programme. Standards for Quality Assurance in Diabetic Retinopathy Screening. First Edition Revision 4.0. Published 2017. Ireland.
6. Patra S, Gomm EM, Macipe M, Bailey C. Interobserver agreement between primary graders and an expert grader in the Bristol and Weston diabetic retinopathy screening programme: A quality assurance audit. Diabet Med 2009;26:820-3.
7. Landis JR, Koch GG. The measurement of observer agreement for

categorical data. Biometrics 1977;33:159-74.
8. UK National Screening Committee. Essential Elements in Developing a Diabetic Retinopathy Screening Programme 2009. Workbook 4.3. Available from: http://www.rcophth-website.www.premierithosting.com/docs/publications/published-guidelines/ENSPDR_Workbook_2009.pdf. [Last accessed on 2016 Nov 06].
9. Abdul-Rahim HF, Husseini A, Bjertness E, Giacaman R, Gordon NH, Jervell J, *et al.* The metabolic syndrome in the west bank population: An urban-rural comparison. Diabetes Care 2001;24:275-9.
10. Chiang F, Kuper H, Lindfield R, Keenan T, Seyam N, Magauran D, *et al.* Rapid assessment of avoidable blindness in the occupied palestinian territories. PLoS One 2010;5:e11854.