Validity of Automated Software Supported Diabetic Retinopathy Screening Compared to Digital Retinal Photograph Evaluation by Retina Subspecialist

Ziaul Haq Yasir, Hassan Al-Dhibi and Rajiv Khandekar
Department of Ophthalmology, King Khaled Eye Specialist Hospital, Riyadh, Saudi Arabia

*Corresponding author: Ziaul Haq Yasir, King Khaled Eye Specialist Hospital, Al oruba street, Um Al-Hamam West, Riyadh, Riyadh 11462, Saudi Arabia, Tel: + 966 547081988; E-mail: yasir0075@yahoo.co.in

Received date: August 12, 2018; Accepted date: August 27, 2018; Published date: August 31, 2018

Abstract

Background: To review the validity of automated screening software (RetinaLyze) for diabetic retinopathy (DR) compared to the evaluation of digital fundus images by a retina subspecialist.

Materials and Methods: This cross-sectional study was conducted at Tertiary eye hospital in 2016 and 2017. The digital fundus images of diabetics at our eye hospital and at a Primary health center (PHC) were obtained using non-mydriatic retcam. The image was linked to the RetinaLyze software (test 1). It was also reviewed by retina subspecialist (test 2). DR was graded into NO DR, Non-proliferative DR (NPDR), proliferative DR (PDR). Agreement rates, sensitivity, specificity and other validity parameters were calculated using SPSS.

Results: Retinal images of 460 eyes and 239 diabetics were included. The prevalence of DR and sight threatening diabetic retinopathy (STDR) were 52.2% and 22.4% respectively. Grading of DR by both tests matched in 281 (61.1%) eyes. RetinaLyze did not detect 47 (10.3%) eyes with STDR. The sensitivity and specificity of software based STDR screening were 35.7% and 83.3% respectively.

Conclusions: RetinaLyze automated screening software is easy to use in the field for DR screening. However, its validity is less than desired for a good DR screening tool.

Keywords: Diabetic retinopathy; Sight threatening diabetic retinopathy; Diabetic macular edema; RetinaLyze

Introduction

Diabetes Mellitus (DM) Type II is a major public health problem worldwide [1]. Diabetic retinopathy (DR) is a leading cause of vision loss in the working-age population [2]. DR affects 126.6 million diabetics globally and 37.3 million diabetics have sight threatening diabetic retinopathy (STDR) [3]. Early detection of DR and timely management of STDR will decrease the public health burden of this disease. The exponential rise in the prevalence of diabetes implies that the current yearly screening efforts for DR will be inadequate [4].

The evaluation of DR using fundus photographs is a reliable and the most commonly used method [5]. Changes in DR are noted and documented using ophthalmoscopy, slit lamp bio-microscopy, fundus cameras and smart phone apps. The images are then sent to the reading centers via tele-ophthalmology or through a secure online website [6].

Automated screening software linked to digital cameras was introduced to decrease the workload of specialists and for a faster diagnosis. The technical enhancements for automated image analysis require accurate algorithms. Automated screening software could be useful in developing countries where the burden of DR is high and resources are limited.

Various automated screening programs are available including ARIA, retmarker (Portugal), EyeArt (USA), IDP (USA), iGrading and RetinaLyze [7]. The diagnostic accuracy of these automated programs is debatable especially in the presence of other ocular comorbidities [8,9]. However, the automated screening programs may assist the primary care physicians in timely referral of the majority of STDR cases [10]. A Dutch study reported high validity of a commercially available automated fundus image analysis software (RetinaLyze, Netherlands) [11]. However, to the best of our knowledge, there is no published study that evaluated this software on a diabetic Arab population. Hence, prior to broad scale application, it is essential to test this software in an Arab population.

Saudi Arabia is facing an epidemic of DM and DR. In 2012, the prevalence of DM in adult Saudis was 29.7%. Among registered diabetics, the prevalence of DR and STDR was 36.8% and 17.5% respectively [12]. In 2016, more than 452,200 diabetic patients attended government health institutions in Saudi Arabia [13]. However, there are only 700 ophthalmologists in these institutions, posing a major challenge for annual DR screening. In developing countries, task shifting has been used as a strategy for effective DR screening [14]. In Saudi Arabia, mid-level eye care professionals are limited yet the geographic spread of patients with DM is vast. Therefore, using the services of mid-level eye care professionals for DR screening remains a major challenge. A software system for automated fundus image analysis could therefore benefit the national health program. This study compares the grading of DR from digital retina images by RetinaLyze and a retina specialist in Saudi Arabia.
Subjects and Methods

The institutional ethics and research board approved this study (P-1309). Diabetics registered at two Primary Health Centers (PHCs) in the Riyadh region of Saudi Arabia and patients who presented at the screening unit of a tertiary eye care hospital were invited to participate in this study. Informed verbal consent was obtained from all patients. This cross-sectional validity study was performed between December 2016 and June 2017. Diabetics with media opacity and hazy digital images as per retina specialist's evaluation were excluded. Those declining to participate were also excluded.

To calculate the sample size for the present study, we assumed that the sensitivity of the software assisted grading of DR was 90% [11]. To achieve a 95% confidence interval (CI), with a 5% acceptable margin of error and a clustering effect of 2, at least 277 eyes of diabetic patients were required for evaluation by the software and by the retina specialist [15].

Medical retina specialist, ophthalmic technician and epidemiologist were the study investigators. Diabetics were defined as individuals who were registered in the diabetes registry of PHCs or referred to the eye hospital for management of DR.

The digital fundus images were obtained using TRC-NW-300 (Topcon Corp., Tokyo, Japan) non-mydriatic retina camera. One central fundus image covering approximately 45° of retina from the fovea was captured [16]. The fundus images were uploaded from the laptop attached to the retina camera to the website of RetinaLyze. For maintaining confidentiality of patient images, a designated login ID and password were used. A specific client reference number was added for each photo. After uploading the image, the option of running analysis for DR and for age related macular degeneration (AMD) was selected. The results were displayed on the monitor within a few seconds. The software detected 'red lesions' (micro-aneurysms and/or hemorrhages) and used this information to grade DR. RetinaLyze can also detect hard exudates and/or cotton-wool spots designated as a 'bright lesion.' The drusen could also be detected by software and only included in the algorithm for age related macular degeneration (AMD) screening. The steps to software interpretation of the image are described in the manual [17]. The software converts each retina image into a gradient representation. Automated lesion detection is based on the advanced mathematical analysis of the gray-level intensity of the images, where the periphery of potential lesions is established from each of a number of seed points. The optic nerve head (ONH) and the arcades are automatically identified and are excluded to define them as a lesion. A measure of visibility was assigned to each potential red and bright lesion, and lesions exceeding user-supplied visibility thresholds were automatically detected and displayed by the system [18]. The RetinaLyze software gave color code results; 'no immediate alteration' (green), 'few alterations' (yellow) or 'severe alterations' (red). This software however, does not indicate diabetic macular edema (DME). DR changes in fundus image can be located by clicking the 'Toggle DR' overlay icon. The software encircles the micro-aneurysms and/or hemorrhages present in that particular photo with a black ring. Based on the automated analysis, the software then recommends if a visit to an ophthalmologist is warranted earlier than scheduled. Software identification of a single 'red lesion' of any type in any image of a diabetic patient labels the patient as having DR and recommends an ophthalmic referral. Image quality is measured by the variation in the gradients in the image and according to a designated cutoff level of the image quality threshold. Images with small or no gradients are rejected and defined as ungradable. The rejection of one image only from a specific patient will classify the patient as having images of insufficient quality and will recommend referral to an ophthalmologist [17]. The analyzed image with overlay can be downloaded and printed.

The digital retina image was physically transferred using high quality external hard disc to the retina specialist to grade DR. An information technology (IT) expert ensured that the image quality was not negatively affected during the image transfer. The retina status of DR and DME was graded separately. Macular edema was defined as the presence of hard exudates or localized retina thickening within 500 μm of the fovea. The severity of DR was defined according to the Early Treatment Diabetic Retinopathy Study (ETDRS) as No DR, Non-proliferative DR (Mild-Moderate-Severe), and Proliferative DR (PDR) [19], STDR was defined as PDR with or without DME [20]. Feedback from field staff was collected on image capture with the retina camera and image upload to the software.

Data were collected on a pretested data collection form and then transferred to an Excel spreadsheet (Microsoft Corp., Redmond, WA, USA). For univariate analysis, a parametric method was used with Statistical Package for Social Sciences (SPSS-24) (IBM Corp., Armonk, NY, USA). The agreement rate was estimated for software-assisted outcomes versus a retina specialist's grading of DR. The retina specialist interpreted the digital images and graded them as: (1) presence of DR, (2) presence of STDR (warranting an ophthalmologist intervention). The percentage proportions were calculated. In cases with more than 2 response options the Kappa value was calculated. The sensitivity was defined as the ability of the software to correctly diagnose the presence of DR as compared to the retina specialist's report. The specificity was defined as the ability of the software to determine an eye without DR correctly compared to the gold standard. A false positive indicated that the software wrongly diagnosed that DR was present when the retina specialist declared the eye did not have DR. A false negative indicated that the software indicated that no DR when the retina specialist had indicated the presence of DR. The 95% confidence intervals (CI) of validity parameters were also calculated. The observations of the retina specialist and RetinaLyze were collected separately to ensure masking of the data and outcome.

Results

We included 476 images of 239 diabetic patients (two patients were monocular). There were 160 (67%) males and 79 (33.0%) females. The mean age was 56.7 ± 11.5 years.

The prevalence of DR in study population was 52.2% (95% CI: 47.7-56.7). The prevalence of STDR was 22.4% (95% CI: 18.6-26.2). The prevalence of DME in the study population was 21.5% (95% CI: 17.7-25.3).

For validation of software based DR grading, 49 images were considered blurred by the software. The retina specialist found 16 images were blurred and the remaining 33 were adequate for grading of DR. Thus 460 images were reviewed by both methods to determine the validity parameters (Table 1). In 281 (61.1%) images there was agreement between the retina specialist and automated software for grading DR.

Total of 47 (10.3%) eyes with STDR were not detected by RetinaLyze as a severe grade warranting referral (Table 2).

Field staff feedback suggested that the RetinaLyze software was easy to use and required very little training on the digital camera and the software.
software negatively affects patients. In the current study, nearly one in ten screened diabetics was missed by the software for timely and accurate DR grading. Therefore, an expert ophthalmologist was needed to review the missed cases. The ability of the software to miss severe DR cases could explain lower level of validity in our study. The presence of these lesions could allow physicians to predict the risk of rapid progression of diabetic complications [23]. The validity of this software can be increased by combining red and bright lesion detection [18].

The validity in our study was lower than that reported by Larsen et al., Hansen et al. and Bouhaimed et al. [11,17,18,24]. Our sample size was much larger than the previous studies [11,17,18,24]. The sensitivity in the previous studies ranges from 82% to 96.7% and the specificity from 71.4% to 100% [11,17,18,24]. The wide difference in the validity parameters among previous studies and ours is difficult to explain. A study regarding glaucoma medication had highlighted outcome differences among researches that were industry sponsored and non-sponsored [25]. It should be noted that our study was not funded however, the software was provided free of cost for independent testing.

We used an ETDRS grading system for diabetic retinopathy [19]. In UK National Health Service (NHS), a more practical grading system is applied [26,27]. Future studies are required to validate this software compared to the NHS grading system.

There are some limitations to this study. The sample size was calculated based on the rate of DR published by Larsen et al. [11] and not for STDR. The prevalence of STDR was lower than DR in our study. Therefore the sample size for validating STDR was not ideal.

### Conclusion

More than half of the diabetics in the study population had DR and nearly one fourth of diabetics had sight-threatening stages of DR. RetinaLyze automated screening software was perceived as a useful tool in the field for DR screening. However, its action oriented recommendation based on the image evaluation matched to the advice of retina specialists in 62% of cases only. Further refinement of software to grade DR is needed before it is applied on a wider scale especially for the diabetic population with low prevalence of DR.

### Conflict of Interest

Authors declare that there is no conflict of interest.

### References


