Diabetic retinopathy screening using digital non-mydriatic fundus photography and automated image analysis

Anja B. Hansen,1 Niels V. Hartvig,1 Maja S. Jensen,1 Knut Borch-Johnsen,2 Henrik Lund-Andersen1,2 and Michael Larsen1

1Department of Ophthalmology, Herlev Hospital, University of Copenhagen, Denmark
2Steno Diabetes Centre, Gentofte, Denmark

ABSTRACT.

Purpose: To investigate the use of automated image analysis for the detection of diabetic retinopathy (DR) in fundus photographs captured with and without pharmacological pupil dilation using a digital non-mydriatic camera.

Methods: A total of 83 patients (165 eyes) with type 1 or type 2 diabetes, representing the full spectrum of DR, were photographed with and without pharmacological pupil dilation using a digital non-mydriatic camera. Two sets of five overlapping, non-stereoscopic, 45-degree field images of each eye were obtained. All images were graded in a masked fashion by two readers according to ETDRS standards and disagreements were settled by an independent adjudicator. Automated detection of red lesions as well as image quality control was made: detection of a single red lesion or insufficient image quality was categorized as possible DR.

Results: At patient level, the automated red lesion detection and image quality control combined demonstrated a sensitivity of 89.9% and specificity of 85.7% in detecting DR when used on images captured without pupil dilation, and a sensitivity of 97.0% and specificity of 75.0% when used on images captured with pupil dilation. For moderate non-proliferative or more severe DR the sensitivity was 100% for images captured both with and without pupil dilation. Conclusion: Our results demonstrate that the described automated image analysis system, which detects the presence or absence of DR, can be used as a first-step screening tool in DR screening with considerable effectiveness.

Key words: diabetic retinopathy – non-mydriatic – automated image analysis – screening


Introduction

Photographic screening of patients with diabetes mellitus entails classification of diabetic retinopathy (DR) based on the type, number and location of any microvascular lesions in the fundus of the eye, combining a count of the number of lesions and a comparison with standard photographs of various stages of retinopathy (Early Treatment Diabetic Retinopathy Study Research Group 1991). This method requires trained personnel, is time-consuming and costly, and also tedious, semi-subjective and prone to error.

The application of automated image analysis to digital fundus images may, however, reduce the workload and costs by minimizing the number of photographs that need to be manually graded, while at the same time providing an objective, repetitive system that reduces intra- and interobserver variability.

In the present study we tested the ability of an automated system, which includes both red lesion detection and image quality control, to correctly identify patients as having or not having DR, based on images captured with and without pharmacological pupil dilation using a digital non-mydriatic camera. The main end-points were: for methodological purposes, a per-eye analysis of the system’s performance in relation to variations in lesion detection sensitivity; and for clinical purposes, a per-patient analysis of the system’s ability to identify patients with any DR, as in clinical practice the decision to refer a patient to an ophthalmologist is based on the patient’s worst eye.
Material and Methods

Patients

The digital images of all 83 patients used in this study were obtained as part of a photographic validation study of the digital non-mydriatic camera carried out in the Department of Ophthalmology, Herlev Hospital.

All patients were recruited at the Steno Diabetes Centre based on a record review of the most recently diagnosed level of retinopathy according to four morphological groups: no retinopathy, background retinopathy (microaneurysms, haemorrhages, hard exudates), preproliferative retinopathy (which can include intraretinal microvascular abnormalities (IRMA) cotton wool spots and venous abnormalities) and proliferative retinopathy. The recruitment was weighted toward less retinopathy in order to pose a greater challenge in the grading process.

Inclusion criteria were: age 18 years or older, diagnosed type 1 or type 2 diabetes mellitus. Exclusion criteria were: pregnancy; previous retinal laser treatment in either eye; history of conditions in either eye that might preclude pupil dilatation, and the use of eye drops (mydriatic or miotic) that might alter pupil size or reactivity.

The study was approved by the local medical ethics committee. All participants gave written informed consent after receiving full information according to the Helsinki Declaration.

Photography

A non-mydriatic digital fundus camera, a TRC-NW6S (TOPCON, Tokyo, Japan) interfaced with a 3CCD colour camera (KY-F70B; JVC, Tokyo, Japan) with a pixel resolution of 1450 × 1026 and IMAGEnet 2000 computer system were used to capture:

1. five overlapping, non-stereoscopic 45-degree photographs of each eye without pharmacological pupil dilatation (posterior pole, nasal, temporal, superior and inferior), covering up to 94 degrees of the fundus including both the macula and the optic disc (Shiba et al. 1999), and
2. five non-stereoscopic images of each eye with pharmacological pupil dilatation using 2.5% phenylephrine and 1% tropicamide, covering the same area.

All digital images were captured in true colour (24 bits), labelled and stored in TIF format on the IMAGEnet computer system. For manual reading, the five images from each eye were arranged as a single mosaic image and stored using the IMAGEnet mosaic software.

All retinal images were viewed on a 24-inch Trinitron colour graphic display (GDM-FW900; Sony, Tokyo, Japan), true colour 1920 × 1200 resolution.

Manual grading

All images were graded according to a modification of the ETDRS extension of the modified Airlie House classification of DR (Diabetic Retinopathy Study Research Group 1981; Early Treatment Diabetic Retinopathy Study Research Group 1991) as described elsewhere (Hansen et al. 2004). Note that macula oedema was not graded in this study, and that ‘Cannot grade’ was used when the image quality made it impossible to determine whether a characteristic was present in a field: if an area of three or more than three disc areas of the retina was visible, and that area was without the characteristic being graded, the field was then graded as ‘no evidence’.

If the characteristic was present in the visible area, the field was then graded as if the characteristic was present with the same degree of severity in the whole field.

Two independent readers graded all digital images in a masked fashion and the results were compared. In cases of disagreement, an independent adjudicator performed a final overall grading. The readers were also asked to evaluate the image quality.

Automated lesion detection

The fundus images were analysed as individual images and not as the mosaic image, using commercial fundus image analysis software (Larsen et al. 2003a, 2003b). The system uses advanced modelling of the grey-level image function of digital images, primarily the green colour channel, and provides automated red microaneurysm and haemorrhage lesion detection as well as image quality measurement.

Each image is converted into a gradient representation. The vessel tree and the optic nerve head are identified and extracted from the image. Seed points of candidate lesion areas (i.e. dark fundus areas) are located and grown with custom-developed algorithms. A visibility parameter describing the densitometric steepness of the edge of each candidate lesion and its contrast relative to the surrounding fundus is calculated; and according to a present cutoff level: the visibility threshold, it is determined whether the candidate lesion is accepted or not according to a preset cut-off level known as the visibility threshold. Note that identification of a single red lesion of any type in any image from a specific patient will classify the patient as having DR and will lead to referral.

Image quality is measured by the variation in the gradients in the image and; and according to a present cutoff level: the image quality threshold, images with small or no gradients are rejected. Note that rejection of one image only from a specific patient will classify the patient as having images of insufficient quality and will lead to referral.

Besides an image-scale parameter, the visibility threshold and the image quality threshold are the only user-supplied parameters in the system. The visibility threshold controls the lesion detection sensitivity and thus the balance between sensitivity and specificity of the patient classification. In a practical screening scenario the threshold should be adjusted to the specific image acquisition and grading protocol, preferably by a calibration study. For the present study, the adjustment was based on the study data, using the same threshold (visibility threshold 2.1 and image quality threshold 0.57) for images captured with and without pharmacological pupil dilatation.

Statistical analysis

Patients were classified as having automatically detected DR if the algorithm identified a single red lesion of any type in any of the images from that patient or if one or more of the images were identified as being of insufficient image quality. The clinical relevance of the detected lesions was characterized by the sensitivity and specificity against the manual grading results (Altman 1999). The receiver operating characteristic (ROC) of the automated red lesion detection was used to characterize the relationship between sensitivity and specificity, with the area under the curve (AUC) serving as a general
measure of performance (van Erkel & Pattynama 1998; Greiner et al. 2000). All quoted AUC values are presented with 95% confidence intervals.

Results

A total of 83 patients (165 eyes) were photographed and their DR status assessed by visual grading (the gold standard) for images captured without and with pharmacological pupil dilation.

Red lesion detection in images captured without pharmacological pupil dilation

One patient was found to have very severe age-related macular degeneration (AMD) in both eyes and was excluded from this part of the analysis based on the following considerations: AMD is known to share many features with DR, including haemorrhages and hard exudates, making it impossible for the automated image analysis algorithms to differentiate between the two diseases. A further 11 eyes were also excluded from this part of the analysis due to images being classified as ungradable by the visual grading (including one patient [two eyes] in whom it was not possible to obtain any images due to small pupils). Thus the total number of eyes was 152 (76 patients).

The per-eye ROC curve (Fig. 1A) of the automated red lesion detection illustrates the balance between sensitivity and specificity for images captured without pupil dilation at eye level. The AUC of the ROC curve was 84.1% (95% CI 77.3–90.9%). When the visibility threshold was set to 2.1 for the automated algorithm, the sensitivity for detecting eyes with DR was 90.1% and the specificity for detecting eyes without DR was 73.8%.

The per-patient ROC curve (Fig. 1D) of the automated red lesion detection illustrates the balance between sensitivity and specificity for images captured with pupil dilation at patient level. The AUC of the ROC curve was 94.0% (95% CI 88.9–99.2%). When the visibility threshold was set to 2.1 for the automated algorithm, a sensitivity of 93.9% and specificity of 75.0% for detecting patients with and without DR, respectively, were obtained.

Red lesion detection in images captured with pharmacological pupil dilation

Images from all 83 patients were classified as gradable. The patient with very severe AMD in both eyes was also excluded from this part of the analysis based on the considerations stated previously. Thus the total number of eyes was 163 (82 patients).

The per-eye ROC curve (Fig. 1C) of the automated red lesion detection demonstrates the range between sensitivity and specificity for images captured with pupil dilation at eye level. The AUC of the ROC curve was 90.8% (95% CI 86.3–95.3%). When the visibility threshold was set at 2.1 for the automated algorithm, the sensitivity for detecting eyes with DR was 90.1% and the specificity for detecting eyes without DR was 73.8%.

The per-patient ROC curve (Fig. 1D) of the automated red lesion detection illustrates the range between sensitivity and specificity for images captured with pupil dilation at patient level. The AUC of the ROC curve was 94.0% (95% CI 88.9–99.2%). When the visibility threshold was set to 2.1 for the automated algorithm, a sensitivity of 93.9% and specificity of 75.0% for detecting patients with and without DR, respectively, were obtained.

Image quality control and red lesion detection in images captured without pharmacological pupil dilation

All 83 patients were included in this part of the analysis. The visibility threshold was set at 2.1 and the image quality threshold at 0.57, i.e. patients with one or more images below this threshold were referred owing to insufficient image quality.

Table 1 summarizes the results of the patient classification with respect to automatically detected red lesions and the image quality measure of the images captured with pupil dilation and compares them with the results of the visual grading; The sensitivity of the automated algorithm with respect to determining whether a patient was in need of referral was 97.0%, and the specificity 75.0% (Table 2). Note that all patients (in total 27) evaluated as having moderate non-proliferative or more severe DR by the visual grading were correctly identified for referral to an ophthalmologist.

False-negative classification (i.e. patients with retinopathy not identified by the automated algorithm) amounted to 8.4% (7/83); false-positive classification (i.e. patients without retinopathy identified as having retinopathy by the automated algorithm) amounted to 2.4% (2/83). Of the seven false-negatives, five had microaneurysms only, one had microaneurysms, haemorrhages and hard exudates, and one was judged as ungradable by the visual grading and thus referred. The two false-positives both had a small hyper pigmentation, which the algorithm mistook for a red lesion.

In total six patients were evaluated as being in need of referral by the visual grading due to ungradable images, including the one patient where it was not possible to obtain any images without pupil dilation due to small pupils (Table 1). The automated algorithms likewise identified six patients for referral due to insufficient image quality, of whom five (including the patient with no images) were among those identified by the visual grading.

Image quality control and red lesion detection in images captured with pharmacological pupil dilation

Images from all 83 patients were included in this part of the analysis. The visibility threshold was set at 2.1 and the image quality threshold at 0.57.

Table 2 summarizes the results of the patient classification with respect to automatically detected red lesions and the image quality measure of the images captured with pupil dilation and compares them with the results of the visual grading; The sensitivity of the automated algorithm with respect to determining whether a patient was in need of referral was 97.0%, and the specificity 75.0% (Table 2). Note that all patients (in total 27) evaluated as having moderate non-proliferative or more severe DR by the visual grading were correctly identified for referral to an ophthalmologist.

The proportion of false-negative classification was 2.4% (2/83) and that of false-positive classification was 4.8% (4/83). The two false-negatives both had microaneurysms only, whereas the four false-positives each had a small hyper pigmentation, which the algorithm mistook for a red lesion.
All images were evaluated as gradable by the visual retinal grading, whereas the automated algorithm identified two patients for referral solely due to insufficient image quality.

Discussion

Our study demonstrates that an automated image analysis system as described in this article can be used in a screening scenario to detect the presence or absence of DR with considerable effectiveness.

At eye level, when analysing images captured without pharmacological pupil dilation using red lesion detection only, we achieved a sensitivity of 81.1% and a specificity of 80.5%. Pharmacological dilation of the pupil resulted in an increased sensitivity of 90.1% but at the cost of a lower specificity of 73.8%. Thus dilation of the pupil increased the number of correctly identified eyes with DR but also the number of false-positives. This is likely due to a difference in image quality, as we have previously found that images captured with pupil dilation tend to be of better image quality (increased brightness) compared to those captured without (Hansen et al. 2004). The better image quality

---

Fig. 1. Receiver operating characteristic (ROC) of the automated diabetic retinopathy red lesion detection algorithm. Sensitivity is depicted as a function of 1 minus specificity. The square denotes the current setting of the algorithm with a visibility threshold of 2.1. With this setting the cited sensitivities and specificities were achieved. (A) Per-eye analysis on images captured without pupil dilation (152 eyes). (B) Per-patient analysis on images captured without pupil dilation (76 patients). (C) Per-eye analysis on images captured with pupil dilation (163 eyes). (D) Per-patient analysis on images captured with pupil dilation (82 patients).
facilitates red lesion detection but also increases the number of false-positives; as previous studies using the automated algorithm (Larsen et al. 2003a, 2003b) have shown that the false-positive automatically detected lesions were mainly found in well defined bright areas of visible nerve fibre layers or in areas with bright posterior hyaloid reflexes. In these areas small, well circumscribed features of normal yellow-red fundus pigmentation, and small vessel segments, not identified as branchings off the main vascular trunks, appeared as isolated high contrast elements that mimicked small haemorrhages and microaneurysms.

When automated image quality control as well as red lesion detection was applied to the images, the sensitivity at patient level increased from 88.7% to 89.9% (without pupil dilation) and from 93.9% to 97.0% (with pupil dilation). The specificity remained unchanged. For images captured without pupil dilation the increased sensitivity was due to the image quality control correctly identifying for referral five of the six patients whose images had been evaluated as ungradable by the visual grading (i.e. excluded from the initial analysis). For images captured with pupil dilation the increase was due to the image quality control correctly identifying for referral two patients who were assessed as having DR by the visual grading, but who were found to have no detectable lesions by the red lesion detection alone. This demonstrates that the addition of an image quality control to an automated system will not only increase the sensitivity of the system, but will also reduce the workload, as previous studies testing automated algorithms have found it necessary to manually grade retinal images for image quality before processing them automatically (Hipwell et al. 2000; Sinthanayothin et al. 2002; Larsen et al. 2003b; Olson et al. 2003).

All patients with false-negative gradings in this study (seven patients based on the images captured without pupil dilation and two patients based on the images captured with pupil dilation) had minimal/mild non-proliferative retinopathy. The potential risk of misclassifying a patient with that level of retinopathy can be estimated based on data published by the Liverpool Diabetic Eye Study. Here, the 1-year incidence for development of sight-threatening DR (defined as ETDRS level ≥ 40) for patients who, at baseline, had no DR or background retinopathy was 0.3% and 3.6%, respectively, for patients with type 1 diabetes (Younis et al. 2003a), and 0.3% and 5.0%, respectively, for patients with type 2 diabetes (Younis et al. 2003b). These numbers indicate that the risk of overlooking a single or few microaneurysms may be inconsequential in a screening situation where the follow-up interval is 1 year. However, if disease in the same patient is missed on subsequent
examinations the risk obviously increases, necessitating future studies to address whether the missing of lesions is a random function or an event that is repeated systematically in a subgroup of eyes or patients; if the latter is shown to be the case, possible markers of poor detectability should be identified so they can be taken into account, if possible. In this study all patients evaluated as having moderate non-proliferative or more severe DR by the visual grading were correctly identified for referral.

In our study we investigated the use of the automated algorithm on images captured with and without pharmacologically pupil dilation. Which of these two formats is preferable will ultimately depend upon the screening setting, the screening population and the preferred standards of sensitivity and specificity. The British Diabetic Association (1997) recommended that, when performing retinal photographic screening for DR, the standard for sensitivity should be 80% and for specificity 95%. It is possible to almost comply with this recommendation by changing the algorithm’s visibility threshold to 4.5 for images captured with pupil dilation, thereby obtaining sensitivity of 80.3% and specificity of 93.4%, and to 3.0 for images captured without pupil dilation, thereby obtaining sensitivity of 80.6% and specificity of 92.9%. However, it should be noted that these are recommendations and are not based on any scientific evidence; whether such accuracy and precision are clinically necessary has yet to be demonstrated. We found that in a screening situation it is important to have a system with high sensitivity. We therefore recommend that patients have their pupils pharmacologically dilated prior to capturing the images, making it possible to obtain a sensitivity of 97%. Noting that in certain situations it may not be practical or desirable to dilate the pupils pharmacologically, in which case data indicate that the experience of the photographer may influence the image quality of images captured without pharmacologically dilated pupils, which may potentially increase the sensitivity of these images (Hansen et al. 2004).

In light of the increasing number of patients with diabetes mellitus (Amos et al. 1997; King et al. 1998; Glumer et al. 2003) and the resultant increase in pressure on the health care system, we propose the following first-line screening system for DR as representative of a cost-effective telemedicine tool: patients would be screened yearly using images obtained with a digital non-mydriatic camera graded by an automated image analysis algorithm. If the automated algorithm detected no retinopathy, the patient would be recalled 1 year later for repeated screening; if retinopathy were detected, the patient’s images would be referred for human inspection. As the majority of patients with diabetes in most screening populations are likely to have no retinopathy the application of such an automated system would substantially reduce the burden of manual grading.

In conclusion, this study showed that automated detection of red lesions combined with image quality control was successful in identifying all diabetic patients with moderate non-proliferative or more severe DR, and 97.0% all of patients with any retinopathy using images captured with pupil dilation. It was also demonstrated that the addition of an image quality control to an automated system will not only increase the sensitivity of the system, in our study by little over 3%, but also reduce the workload, as there will be no need to manually assess image quality before images are processed automatically.

Acknowledgements
This study was funded by the Danish Diabetes Association, the Danish Eye Health Society (Værn om Synet), the Eye Foundation (Œjenfonden), the T. Elmqvist Foundation and Synoptik A/S.

Commercial Interest
None of the authors have any commercial interest in the described automated image analysis system.

References

Received on March 31st, 2004.
Accepted on August 17th, 2004.

Correspondence:
Anja Bech Hansen
Department of Ophthalmology
Herlev Hospital
DK-2730 Herlev
Denmark
Tel: + 45 44 88 46 57
Fax: + 45 44 53 86 69
Email: anbha@herlevhosp.kbhamt.dk