

Automated Detection of Fundus Photographic Red Lesions in Diabetic Retinopathy

Michael Larsen,¹ Jannik Godt,² Nicolai Larsen,¹ Henrik Lund-Andersen,¹
Anne Katrin Sjølie,³ Elisabet Agardh,⁴ Helle Kalm,⁵
Michael Grunkin,² and David R. Owens⁶

PURPOSE. To compare a fundus image-analysis algorithm for automated detection of hemorrhages and microaneurysms with visual detection of retinopathy in patients with diabetes.

METHODS. Four hundred fundus photographs (35-mm color transparencies) were obtained in 200 eyes of 100 patients with diabetes who were randomly selected from the Welsh Community Diabetic Retinopathy Study. A gold standard reference was defined by classifying each patient as having or not having diabetic retinopathy based on overall visual grading of the digitized transparencies. A single-lesion visual grading was made independently, comprising meticulous outlining of all single lesions in all photographs and used to develop the automated red lesion detection system. A comparison of visual and automated single-lesion detection in replicating the overall visual grading was then performed.

RESULTS. Automated red lesion detection demonstrated a specificity of 71.4% and a resulting sensitivity of 96.7% in detecting diabetic retinopathy when applied at a tentative threshold setting for use in diabetic retinopathy screening. The accuracy of 79% could be raised to 85% by adjustment of a single user-supplied parameter determining the balance between the screening priorities, for which a considerable range of options was demonstrated by the receiver-operating characteristic (area under the curve 90.3%). The agreement of automated lesion detection with overall visual grading (0.659) was comparable to the mean agreement of six ophthalmologists (0.648).

CONCLUSIONS. Detection of diabetic retinopathy by automated detection of single fundus lesions can be achieved with a performance comparable to that of experienced ophthalmologists. The results warrant further investigation of automated fundus image analysis as a tool for diabetic retinopathy screening. (*Invest Ophthalmol Vis Sci.* 2003;44:761-766) DOI: 10.1167/iovs.02-0418

From the ¹Department of Ophthalmology, Herlev Hospital, University of Copenhagen, Copenhagen, Denmark; ²Retinalyze A/S, Hørsholm, Denmark; the ³Department of Ophthalmology, Odense University Hospital, Odense, Denmark; the ⁴Department of Ophthalmology, Malmö University Hospital, Malmö, Sweden; the ⁵Department of Ophthalmology, Sahlgrenska University Hospital, Göteborg, Sweden; and the ⁶Department of Diabetes Research, University of Wales College of Medicine, Cardiff, Wales, United Kingdom.

Submitted for publication April 29, 2002; revised August 7, 2002; accepted August 19, 2002.

Commercial relationships policy: E (JG, MG); C (ML, NL, HL-A, DRO); P (MG); N (EA, AKS, HK).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Corresponding author: Michael Larsen, Department of Ophthalmology, Herlev Hospital, DK-2730 Herlev, Denmark; mla@dadlnet.dk.

Morphologic classifications of diabetic retinopathy are based on the number, location, and type of discrete microvascular lesions in the fundus of the eye.¹ Analysis is generally static, in that the chronological sequence of appearance, maturation, and disappearance of fundus lesions is not accounted for because of the complexity of the task. For low lesion densities, an exact lesion count may be used, but for higher lesion densities no attempts at counting are made. Instead, comparison with standard photographs at various stages of retinopathy is used to achieve a semiquantitative classification.

Digital analysis of fundus photographic images may enable detailed lesion mapping and counting, and dynamic analysis of time series of fundus photographs. The simplest automatable task of practical utility is probably that of distinguishing between subjects without retinopathy and those with any level of retinopathy, of which only the latter may need secondary-level evaluation—in this case, a visual evaluation by an ophthalmologist or a grader. Absence of retinopathy is diagnosed in a large proportion of patients with diabetes in populations undergoing photographic screening.² Diagnosing the absence of retinopathy or presence of a single or very few lesions is often more time consuming than diagnosing mild levels of retinopathy, because it may involve detailed consideration of borderline elements of doubtful identity.

The objective of the present study was to develop a method for automated detection of hemorrhages and microaneurysms in digitized fundus photographic images of patients with diabetic retinopathy and to compare the automated detection with visual identification of diabetic retinopathy.

MATERIALS AND METHODS

The Welsh Community Diabetic Retinopathy Study (WCDRS) was designed to compare ophthalmoscopy and retinal imaging (35-mm color film) in the detection of diabetic retinopathy in a population/community-based screening program.² From this cohort a random sample was extracted of 100 patients with diabetes who had fundus status ranging from no retinopathy through nonproliferative retinopathy to preproliferative or early proliferative retinopathy.

The study includes a baseline examination and a follow-up examination after 18 months. Only photographs from the follow-up examination were included in a substudy of single-lesion detection, whereas both sets of photographs, a total of 800, were included in a substudy of the reproducibility of diabetic retinopathy grading between color transparency viewing and computer monitor display of digitized color versions of the same images. Both the color film and the digitized photographs were classified in random order. No upper limit was set for the time used to grade either set of photographs.

Image quality was graded for each eye according to the following classifications: excellent, good, fair, or poor (EURODIAB criteria³).

The present study is an entirely retrospective analysis of a prospectively planned investigation. The protocol underwent appropriate institutional review and complied with the tenets of the Declaration of

Helsinki.² The present study did not involve patients or biological samples, and as such did not require renewed institutional review under the national laws of the countries of residence of the study participants.

Photography

Fundus photographs were obtained after dilatation of the pupil with one or more drops of phenylephrine hydrochloride 2.5% and/or tropicamide 1%. A fundus camera (CR4 45 NM; Canon Europa, NV, Amstelveen, The Netherlands) set at a 45° angular field of view and 35-mm color transparency film (Ektachrome Elite 100; Eastman Kodak Corp., Rochester, NY) were used.

Four photographs were recorded at each visit in each eye of each of the 100 patients: two with the nasal border of the photographic field transecting the optic disc (macula-temporal field) and two with the temporal border of the photographic field set at 1 disc diameter from the optic disc on its temporal side (disc-nasal field). The two fields in combination span a viewing angle of 45° (vertically) by 80° (horizontally).³ The best photograph of each field was used for grading and the other was retained in the project files.

Digitization

The slide-mounted 35-mm color film transparencies were digitized with a slide film digitizer (Coolscan LS-2000; Nikon Corp., Tokyo, Japan) at 1350 dpi and 12 bits per pixel per color channel (red, green, and blue [RGB]). The images were stretched to 8 bits (values from 0 to 255) per color channel and stored in the uncompressed RGB tagged information file format (TIFF). The full image size was 1448 × 1296 pixels, and the diameter of the circular fundus region defined by the internal mask of the camera was approximately 1170 pixels. The images were displayed on a 21-in. color monitor (Professional Series PT813; ViewSonic Corp., Walnut, CA), unless otherwise specified.

Single-Lesion Classification

For visual inspection, the color and contrast of the digitized fundus images were enhanced according to an automated procedure involving full color stretching separately for each color channel. The lesions were outlined with a manually controlled screen marker. The software annotation tool included circles, ellipses, and freehand drawing to enclose the lesion as narrowly as possible. The lesion's class and outline were displayed in color code as an overlaid image that could be toggled on and off.

Red fundus lesions were defined as any fundus lesion visibly recognizable as being a hemorrhage or a microaneurysm or any of the two. No other types of lesions were accepted. Hemorrhages were defined as comprising round or irregularly shaped, sharply or diffusely outlined, deep red areas of the color of venous blood or darker. Microaneurysms were defined as small round lesions with a well-defined edge and occasionally a brighter rim. The color and saturation of microaneurysms vary according to the oxygenation of the blood, from bright red (nearly pink) to deep red. The definitions of microaneurysms and hemorrhages are overlapping, and therefore the classification did not distinguish between the two.

Overall Visual Grading

Using standard grading conditions, two experienced graders of the Bro Taf Screening Service (Bro Taf Area Health Authority Diabetic Retinopathy Screening Service, Cardiff, Wales, UK), independently and in a randomized order of presentation, graded the color fundus photographs and image quality on transparency film of each eye according to the Bro Taf protocol (a modification of the EURODIAB standard).^{2,3} In a subsequent session, the graders reached consensus on the grades of all eyes.

One month later, the graders independently and in a randomized order of presentation graded the digitized images of each eye accord-

ing to the Bro Taf protocol. The images were presented on two 19-in. color computer monitors (Multiscan G400; Sony, Tokyo, Japan). In a subsequent session, the graders reached consensus on the grades and image quality of all eyes, and the consensus grades were used to classify patients into the gold-standard binary reference classification of no diabetic retinopathy and any diabetic retinopathy, by using the Bro Taf levels of no diabetic retinopathy and questionable diabetic retinopathy for the first classification and levels of manifest diabetic retinopathy for the second classification. The overall visual gradings were reserved for the validation of automated lesion detection and were not used for development of the lesion-detection algorithm.

Visual Single-Lesion Annotation

Each of six ophthalmologists, all of whom subspecialize in retinal disease, identified and outlined as narrowly as possible all red lesions in each of the 400 digitized fundus photographs, independently and in random order of presentation.

First Round of Automated Lesion Detection

The digitized fundus images were analyzed with commercial fundus image-analysis software (RetinaLyze System; Retinalyze A/S, Hørsholm, Denmark). The lesion-detection algorithm of the system applies advanced modeling of the gray-level image function of digital images, primarily the green color channel. Seed points of feature candidates are identified by using local image properties. The vessel tree and optic nerve head are automatically identified and extracted before growing lesion candidates. A visibility parameter describing each candidate lesion is then calculated. This parameter describes the densitometric steepness of the edge of the candidate lesion and its contrast relative to the surrounding fundus, with allowance for other dark areas being permissible in its vicinity, provided that such dark areas are noncontiguous with the vascular branchings of the trunk vessels of the retina. The level of visibility required for acceptance of a candidate lesion is determined by a single user-supplied parameter, the visibility threshold, which controls the sensitivity of the automated lesion detection. The analysis was performed on a personal computer with a standard operating system (Windows NT; Microsoft Corp., Seattle, WA).

Adjudicated Single-Lesion Annotation

Three months after the specialist lesion outline, an adjudicating single-lesion classification was made of each of the 400 digitized fundus images. Three of the six ophthalmologists discussed and reached agreement or decided by majority vote whether to accept or reject a given hemorrhage or microaneurysm they deemed to be at least 50% likely to be a diabetic lesion. To ascertain that the adjudicating panel considered all potential red lesions, the panel members were shown all red lesion outlines set by the six specialists as well as by the automatic lesion-detection system. All outlines were shown as a simultaneous overlay in a masked appearance, using the same marker for automatically detected and specialist-identified lesions. At the discretion of any panel member the manually controlled zoom was linearly increased to approximately 30 × 25 pixels, which allowed the panel to investigate detailed areas of the images and to accept, reject, or redraw existing outlines or to outline unmarked lesions that they detected.

At a training session before the adjudicating session, a protocol amendment was accepted requiring all parts of the fundus images to be investigated simultaneously by the panel members at a specified image zoom factor to approximately 400 × 300 pixels. In addition, the adjudicators were allowed to examine the original slide-mounted transparencies using a 15-D magnifying spectacle lens and a slide projector.

The adjudicated set of single red lesion annotations was used to make the final adjustment of the automated lesion-detection algorithm.

Statistical Analysis

The agreement between overall visual gradings based on inspection of color transparency film and inspection of computer screened versions of the same photographs was calculated using the Cohen weighted κ .⁴

Patients were classified as having automatically detected diabetic retinopathy if a single red lesion of any type was identified by the automated lesion-detection algorithm in any of the four images of the patient's two eyes. The clinical relevance of the detected lesions was characterized by the specificity, sensitivity, and accuracy (the fraction of the entire study population that was correctly classified) against the gold-standard reference classifying each patient as having or not having diabetic retinopathy. The receiver operating characteristic (ROC) curve of the automated lesion detection was used to characterize the relationship between sensitivity and specificity, with area under the curve (AUC) serving as a general measure of performance.⁵

The Cohen κ of agreement between automated classification and the gold standard reference was compared with the mean of agreements between the six ophthalmologists and the gold standard reference. The coefficient of variation was used to characterize the variability between individual ophthalmologists.

Statistical analysis was performed on computer (SAS System, ver. 8.01; SAS Institute Inc., Cary, NC; and StatXact 4; Cytel Software Corp., Cambridge, MA).

TABLE 1. Comparison of Diabetic Retinopathy Grading between Fundus Photographs on Color Transparency Film versus Digitized Image Display

Color Film Grade	Grade Description	Digitized Image Grade						
		0	0+	1	2a	3	4a	4b
0	No DR	288	2	16	0	0	0	0
0+	Questionable DR: no MA; HM >1 DD from fovea; <5 CWS; HE >1 DD from fovea	2	13	0	0	0	0	0
1	Minimal nonproliferative DR: ≤ 5 MA; >3 MA <1 DD from fovea	12	0	23	2	1	0	1
2a	Mild nonproliferative DR: >5 MA, and/or, in the presence of MA, HM >1 DD from fovea; HE outside arcades; ≤ 5 CWS; HE within arcades >1 DD from fovea	2	0	2	9	2	0	0
3	Moderate nonproliferative DR: HM <1 DD from fovea; HE <1 DD from fovea; possible CSME (unexplained v.a. <0.5)	1	0	0	0	12	2	0
4a	Preproliferative DR: >5 CWS; extensive HM	0	0	0	0	1	1	0
4b	Preproliferative DR: venous beading/loops; IRMA	0	0	1	1	3	0	2

The gradings represent the consensus of two graders using the Bro Taf modification of the EURODIAB grading scale. No eyes had diabetic retinopathy of more advanced grades than those shown. Both eyes in the 100 patients were graded by visual examination of color film transparencies and digitized image display in two separate grading sessions. A grading was made for each of two photographic examinations at 18 months' interval. Photographs of one eye from one examination were erroneously omitted from the grading session. DR, diabetic retinopathy; MA, microaneurysm; HM, hemorrhage; DD, disc diameter; CWS, cotton wool spot; HE, hard exudate; IRMA, intraretinal macular abnormalities; CSME, clinically significant macular edema; v.a., visual acuity. Figures in bold indicate agreement between the two gradings. Semicolons in table should read as "and/or."

TABLE 2. Agreement within and between Two Observers Grading Diabetic Retinopathy by Inspection of Fundus Photographic Color Transparency Film and Digitized Color Display of the Same Photographs

	Cohen Weighted κ	95% Confidence Interval
Intragrader agreement between image formats	0.78	0.72-0.83
Intergrader agreement, assessment of fundus photographs on color film	0.71	0.62-0.80
Intergrader agreement, assessment of digitized fundus images	0.68	0.57-0.78

The grading system was the Bro Taf modification of the EURODIAB grading scale (Table 1).³

RESULTS

The diabetic retinopathy status of the 400 eyes, as seen on digitized fundus photographic images, was found to range from no retinopathy to preproliferative diabetic retinopathy, with a considerable preponderance of eyes without retinopathy (Table 1). Comparison of color transparencies viewed through a magnifier with digitized fundus photographs viewed on a computer monitor display demonstrated an intragrader agreement (95% confidence intervals) between image formats on a seven-stage ordinal retinopathy scale of 0.78 (0.72-0.83; Table 2). The intergrader agreement within image formats was 0.71 (0.62-0.80) for color transparencies and 0.68 (0.57-0.78) for digitized images. Consequently, the grading was observed to be more reproducible between image formats than between graders, although no statistically significant differences appeared between agreements.

All digitized fundus images were classified as gradable, with the distribution of image quality among the photographs being 10.4% excellent, 82.5% good, 6.9% fair, and 0.26% poor. No patient in the study population had bright lesions (exudates or cotton-wool spots) without also having dark lesions (hemorrhage and microaneurysms).

The ROC curve (Fig. 1) of the automatic red lesion detection demonstrates the range of options in setting the balance between sensitivity and specificity. The AUC of the ROC curve was 90.3%. With the key parameter of the automated-detection algorithm, the visibility threshold (see the Methods section), set to its default value of 1.2, the specificity of detecting patients without retinopathy was 71.4% and the sensitivity of detecting patients with retinopathy was 96.7%, resulting in an overall accuracy (the fraction of the entire study population that was correctly classified) of 79.0% (Table 3; Fig. 2).

With this parameter setting, 50 of 70 patients without diabetic retinopathy were correctly classified as having no retinopathy by the automated red lesion-detection algorithm. One patient was incorrectly classified as having no diabetic retinopathy. This patient had minimal nonproliferative diabetic retinopathy in one eye with vaguely defined microaneurysms in an underilluminated peripheral crescent of one of two photographs.

Twenty patients without diabetic retinopathy or with questionable diabetic retinopathy were incorrectly classified by automated red lesion detection as having retinopathy. The false-positive lesions arose in eyes with well-defined bright

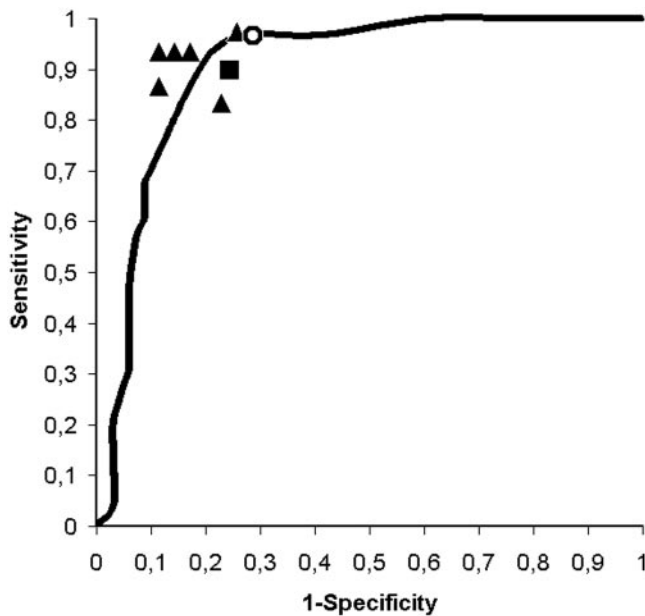


FIGURE 1. Receiver operating characteristic (ROC) of an automated diabetic retinopathy red-lesion-detection algorithm. Sensitivity is depicted as a function of 1 minus specificity. Also shown are the sensitivity and specificity coordinates of six independent graders of single lesions (\blacktriangle) and of an adjudicating panel (\blacksquare) consisting of three of the aforementioned graders who evaluated all candidate lesions defined by the graders or the automated lesion-detecting algorithm. The circle denotes the default setting of the algorithm with visibility threshold 1.2. This ROC coordinate was chosen to represent a setting of potentially practical usefulness in screening for diabetic retinopathy. With this setting, a sensitivity of 96.7% and a specificity of 71.4% in the identification of subjects with diabetic retinopathy was achieved.

areas of visible retinal nerve fibers or with bright posterior hyaloid reflexes. On the background of these features, small, interspersed, well-circumscribed areas of normal yellow-red fundus pigmentation appeared as isolated high-contrast elements (intensity minima) that mimicked small hemorrhages and microaneurysms.

When we increased the visibility threshold to 1.5, the specificity was changed to 85.7%, and the sensitivity to 83.3%, thus resulting in an overall accuracy 85.0%. This demonstrates the possibility of adjusting the algorithm for specific objectives by variation of the visibility threshold.

The mean specificity of the six individual specialists was 82.9% (range, 74.3%–88.6%). This was higher than the specificity of 71.4% achieved by the automated lesion detection (Table 4). The mean sensitivity of the individual specialists was 91.1% (range, 83.3%–96.7%). This was lower than the sensitivity of 96.7% achieved by the automated lesion detection at the default threshold value. The specificity of 75.5% of the adjudicating panel was higher than that of the automated lesion detection and the sensitivity of 90.0% was lower than that of the automated lesion detection. At a visibility threshold of 1.2 the sensitivity of the automated lesion detection matched the sensitivity of the best-performing specialist, at a cost of 2.9 percentage points of specificity.

At a visibility threshold of 1.5 the agreement of the automated lesion detection with the overall visual grading was 0.659. The corresponding mean value of the six individual ophthalmologists was 0.648, with a coefficient of variation of 14.5%. Consequently, the automated grading could be adapted to match the average performance of experienced ophthalmologists.

Discussion

We have shown that a red-lesion-detection algorithm operating on digitized fundus images can achieve considerable effectiveness in sorting fundus photographs that show no disease from photographs with nonproliferative diabetic retinopathy. Thus, 50 of 70 patients without retinopathy were correctly identified.

The present study was intended to determine whether a sound ROC curve function can be achieved in automated detection of diabetic retinopathy. It was not intended to measure the sensitivity and specificity that is ultimately obtainable by automated fundus image analysis, because it involved a partial circular argument, in that the algorithm was tested on the same fundus photographs that were used to develop it. Consequently, a study of an independent set of photographs has been undertaken.⁶

We specifically chose not to use fluorescein angiography to classify red lesions into microaneurysms and hemorrhages because angiography does not permit distinction between small spot hemorrhages and clotted microaneurysms. Indeed, there is currently no practicable method whereby an anterior ruling can be made on candidate lesions, and color fundus photographic grading is the established method used in clinical trials for morphologic assessment of diabetic retinopathy.

Ultimately, there is no means of objectively determining an ulterior truth, not even by fluorescein angiography. Although it may reveal microaneurysms that are not visible on fundus images, a candidate lesion that does not show up on the angiogram cannot be classified, because it may be a clotted microaneurysm, a hemorrhage, or another type of dark area that does not warrant the diagnosis of any of these lesions. Consequently, it would be difficult or impossible to determine whether the adjudicating panel or the overall graders were most correct in determining the absence or presence of diabetic retinopathy.

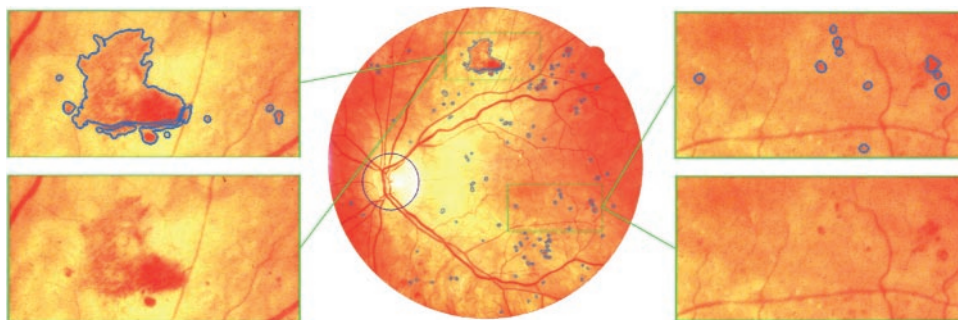
The frequency of diabetic retinopathy in the general population is low, and consequently the identification of one definite lesion in a fundus photograph may induce the observer to decrease his or her threshold for acceptance of candidate

TABLE 3. Classification of Patients with Diabetes by Absence or Presence of Diabetic Retinopathy on Digitized Fundus Photographs with Automated Versus Visual Assessment

Overall Visual Classification	Automated Classification Based on Digital Detection		
	No DR n (%)	DR n (%)	Patients Examined (n)
No DR	50 (71.4)	20 (28.6)	70
DR	1 (3.3)	29 (96.7)	30
Patients examined	51	49	100

The classification is based on two 45° color fundus photographs from each of a patient's two eyes. The key parameter in controlling the sensitivity of the automated lesion detection, the visibility threshold, was set to its default value of 1.2, where one or more red lesions were automatically detected in 96.7% of the patients visually classified as having diabetic retinopathy (sensitivity) and no lesions were found in 71.4% of the patients visually classified as having no diabetic retinopathy (specificity). The fraction of patients correctly classified with diabetic retinopathy by the automated lesion detection was 59.2% (positive predictive value). The fraction of patients correctly classified without diabetic retinopathy by the automated lesion detection was 98.0% (negative predictive value). Percentage of total is shown in parentheses.

FIGURE 2. Digitized color fundus photograph (*middle*) of a left eye with nonproliferative diabetic retinopathy with annotation markers outlining each of the red type of microangiopathic lesion (hemorrhages and microaneurysms) identified by an automated image-analysis algorithm. Magnified sections of the photograph are shown in raw format (*bottom left and right*) as well as with annotation markers (*top left and right*).



elements in the same photograph as being true diabetic lesions. Such a tendency to perceive an increase in the likelihood that morphologically similar elements are indeed specific pathologic lesions and not just random fluctuations in normal fundus morphology could not be confirmed by statistical analysis in the present study, and such reasoning was not implemented in the algorithm.

The present study demonstrates the possibility of achieving algorithmic identification of fundi with diabetic retinopathy with a near-perfect 96.7% sensitivity while maintaining a specificity of 71.4%. The suggested practical usefulness of the automated system is to reduce the workload of visual graders by approximately 50%, depending on the distribution of reti-

nopathy in the screening population. In addition, the high workload and monotonous procedures involved in photographic screening for diabetic retinopathy appear to warrant practical assessment of computer-assisted diabetic retinopathy screening, at a time when digital fundus photography is becoming gradually more common.

In the present study automated lesion detection was applied in a conservative strategy optimized to achieve a significant potential for workload reduction in a screening setting without overlooking even very mild grades of retinopathy. Further studies should investigate whether the option of tuning the lesion detection to values other than its default setting can be safely used.

The effect of potentially confounding pathologic fundus changes has not been thoroughly investigated in the present study because of the low prevalence of such changes in the study population. Preliminary investigations indicate that the algorithm for red lesion detection is virtually insensitive to bright lesions such as drusen, epiretinal fibrosis, and hypopigmentation of the retinal pigment epithelium, whereas it is sensitive to microaneurysms and hemorrhages, regardless of whether these are caused by diabetic retinopathy or other types of retinal disease that produce such elements of microangiopathy.

In both the UK Prospective Diabetes Study (UKPDS) and the Early-Treatment Diabetic Retinopathy Study (ETDRS) the presence of microaneurysms only, in the absence of other lesions, is associated with an increased risk of progression of diabetic retinopathy.^{7,8} Patients in the UKPDS with microaneurysms only had a 3.4% risk of progression to moderate nonproliferative diabetic retinopathy (level 43) or higher levels of retinopathy or photocoagulation treatment within 6 years, whereas the risk in patients with no retinopathy at baseline was 2.7%.⁹ These relatively small numbers indicate that the risk of overlooking a single or a few microaneurysms is inconsequential in screening settings in which the follow-up interval never exceeds 2 years. The risk obviously increases if disease in the same patient is missed more than once in successive examinations. Future studies should address whether the missing of lesions is a random function or a systematically repeating event in a subgroup of eyes or patients, so that markers of poor detectability can be identified and taken into account, if possible.

The present study demonstrates the feasibility of using automated digital detection of red lesions in diabetic retinopathy. Further studies are in progress to determine the reproducibility of the procedures involved. The results of the present study firmly document the potential for incorporating automated image analysis in the diabetic retinopathy screening services. The exact role of such instruments remains to be determined in practice.

TABLE 4. Methods for Identification of Patients without or with Diabetic Retinopathy

Method	Specificity (range)	Sensitivity (range)
Automated computer detection	71.4	96.7
Individual ophthalmologists	82.9 (74.3–88.6)	91.1 (83.3–96.7)
Adjudicating panel of three ophthalmologists	75.7	90.0

Statistical characteristics of methods for identification of patients without or with diabetic retinopathy on digitized versions of color transparency fundus photographs using automated computer algorithmic lesion detection, grading by individual ophthalmologists, or visual inspection by an adjudicating panel of ophthalmologists. All three methods were based on single-lesion detection. These were compared with an overall visual grading by two independent graders. Two fundus photographs were examined from each of the two eyes of 100 patients with diabetes (Table 1). Lesion detection by retina specialists was based on visual examination of digitized color fundus photographs displayed on computer monitor. The independent reference graders examined the same digitized fundus photographs to determine the presence of any diabetic retinopathy on either photograph of a patient's two eyes. After an initial ophthalmologist session, an adjudicating session was arranged, with three of the six ophthalmologists working in conjunction on the basis of a display of all photographs showing lesions marked by any of the specialists or by the automated lesion detection algorithm masked as expert lesion annotations. The adjudication procedure mandated active acceptance or rejection of all doubtful candidate lesions, doubtful being defined as candidate lesions marked by only one ophthalmologist during the initial single lesions annotation, by the algorithm alone, or by the algorithm and a single ophthalmologist. The criterion for acceptance was a perceived likelihood of a candidate lesion as being a true red lesion of at least a 50%. Automatic annotations were displayed in the same color code as the visually identified and manually outlined lesions. Candidate lesions were accepted without active annotation from the consensus panel if they were marked by two or more ophthalmologists, but the consensus panel were free to override this automated function. Data are expressed as percentages.

References

1. 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 number 10. *Ophthalmology*. 1991;98:786-806.
2. Gibbins RL, Owens DR, Allen JC, Eastman L. Practical application of the European Field Guide in screening for diabetic retinopathy by using ophthalmoscopy and 35 mm retinal slides. *Diabetologia*. 1998;41:59-64.
3. Aldington SJ, Kohner EM, Meuer S, Klein R, Sjolie AK. Methodology for retinal photography and assessment of diabetic retinopathy: the EURODIAB IDDM complications study. *Diabetologia*. 1995;38:437-444.
4. Agresti A. *Categorical Data Analysis*. New York: John Wiley & Sons; 1990.
5. Metz CE. Basic principles of ROC analysis. *Semin Nucl Med*. 1978;8:283-298.
6. Larsen N, Godt J, Grunkin M, Lund-Andersen H, Larsen M: Automated detection of diabetic retinopathy in a fundus photographic screening population. *Invest Ophthalmol Vis Sci*. 2003;44:767-771.
7. Kohner EM, Stratton IM, Aldington SJ, Turner RC, Matthews DR. Microaneurysms in the development of diabetic retinopathy (UKPDS 42). UK Prospective Diabetes Study Group. *Diabetologia*. 1999;42:1107-1112.
8. Early Treatment Diabetic Retinopathy Study Research Group. Fundus photographic risk factors for progression of diabetic retinopathy. ETDRS report number 12. *Ophthalmology*. 1991;98:823-833.
9. Stratton IM, Kohner EM, Aldington SJ, et al. UKPDS 50: risk factors for incidence and progression of retinopathy in type II diabetes over 6 years from diagnosis. *Diabetologia*. 2001;44:156-163.