

# Automated Detection of Diabetic Retinopathy in a Fundus Photographic Screening Population

Nicolai Larsen,<sup>1,2</sup> Jannik Godt,<sup>3</sup> Michael Grunkin,<sup>3</sup> Henrik Lund-Andersen,<sup>1,2</sup> and Michael Larsen<sup>1</sup>

**PURPOSE.** To evaluate the performance of an automated fundus photographic image-analysis algorithm in high-sensitivity and/or high-specificity segregation of patients with diabetes with untreated diabetic retinopathy from those without retinopathy.

**METHODS.** This was a retrospective cross-sectional study of 260 consecutive nonphotocoagulated eyes in 137 diabetic patients attending routine photographic retinopathy screening. Mydriatic 60° fundus photography on 35-mm color transparency film was used, with a single fovea-centered field. Routine grading was based on visual examination of slide-mounted transparencies. Reference grading was performed with specific emphasis on achieving high sensitivity. Computer-assisted automated red lesion detection was performed on digitized transparencies.

**RESULTS.** When applied in a screening population comprising patients with diabetes with untreated diabetic retinopathy in any eye and patients with diabetes without retinopathy, the automated lesion detection correctly identified 90.1% of patients with retinopathy and 81.3% of patients without retinopathy. A per-eye analysis for methodological purposes demonstrated that the automated lesion detection could be adapted to simulate various visual evaluation strategies. When adapted at high sensitivity, the automated system demonstrated sensitivity at 93.1% and specificity at 71.6%. When adapted at high specificity the automated system demonstrated sensitivity at 76.4% and specificity at 96.6%, closely matching routine visual grading at sensitivity 76.4% and specificity 98.3%.

**CONCLUSIONS.** Automated detection of untreated diabetic retinopathy in fundus photographs from a screening population of patients with diabetes can be made with adjustable priority settings, emphasizing high-sensitivity identification of diabetic retinopathy or high-specificity identification of absence of retinopathy, covering opposing extremes of visual evaluation strategies demonstrated by human observers. (*Invest Ophthalmol Vis Sci.* 2003;44:767-771) DOI:10.1167/iovs.02-0417

The application of digital analysis to fundus photography requires careful consideration of the virtues and limitations of this new technology. Practical application has been proposed in photographic screening for diabetic retinopathy, to

improve overall performance and productivity by introducing utilities that assist or replace current procedures that depend solely on manual labor and visual inspection. This is particularly relevant when fundus photography services are used to remedy logistic problems in providing qualified biomicroscopic fundus examination to all patients with diabetes.

In the present study we examine the ability of an automated image-analysis system for correct classification of patients as having or not having diabetic retinopathy in a series of untreated eyes with either no retinopathy, mild to moderate nonproliferative diabetic retinopathy, or untreated proliferative diabetic retinopathy. The assumption is that this task is practically meaningful in screening services, because patients who have undergone any treatment by photocoagulation or vitrectomy in either of their eyes should continue to be observed by ophthalmologists, and that these patients will not be considered eligible for automated photographic retinopathy screening. The study involved a consecutive series of fundus photographs from the retinopathy screening service of a large diabetology clinic.

In clinical practice, diabetic retinopathy screening concerns individual patients (pairs of eyes) rather than single eyes, because the decision to refer the patient for ophthalmic examination or treatment is made if one eye has reached or exceeded a threshold level of retinopathy. Consequently, as our primary end point, we evaluated the system's performance in relation to patients, not eyes. The specific objective was to test the ability of automated lesion analysis to identify patients with any diabetic retinopathy in either of their eyes in the subpopulation of diabetic patients who had not previously undergone retinal photocoagulation treatment—the rationale being that patients who have undergone photocoagulation in either eye should be designated a priori to visual retinopathy grading.

For methodological purposes, a per-eye analysis is relevant when seeking to characterize system performance in relation to specific diseases. Consequently, we used a per-eye analysis to evaluate variation in performance in relation to variation in the lesion-detection sensitivity as controlled by the user-supplied visibility threshold parameter.

The development of the automated red lesion detection applied in this study was based exclusively on fundus photographs selected from other diabetic populations. Images from the screening clinic involved in the present study were not used in the development of the automated lesion-detection system; thus, the results of this study serve as validation of the results presented in a previous study.<sup>1</sup>

## MATERIALS AND METHODS

The study photographs were retrieved retrospectively from the photographic archives of all 184 patients who attended the Steno Diabetes Center Retinopathy Service during two randomly selected consecutive weeks of November 1999. Inclusion criteria per patient were an established history of diabetes mellitus, type 1 or type 2, and availability of at least one macula-centered (center of fovea within 10° of the center of the photographic frame) color fundus photograph from at least one eye. The slides were routinely mounted in legal-sized plastic

From the <sup>1</sup>Department of Ophthalmology, Herlev Hospital, University of Copenhagen, Herlev, Denmark; <sup>2</sup>Steno Diabetes Center, Gentofte, Denmark; and <sup>3</sup>Retinalyze A/S, Hørsholm, Denmark.

Supported by the Danish Eye Research Foundation (Øjenfonden) and the Danish Diabetes Association (Diabetesforeningen).

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); P (MG).

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: Nicolai Larsen, Department of Ophthalmology, Herlev Hospital, DK-2730 Herlev, Denmark; nicolai.larsen@dadnet.dk.

pockets, from which the slide in the upper left pocket was systematically chosen. This slide represented the left slide of a stereo pair. Exclusion criteria per eye were previous photocoagulation or surgical treatment of diabetic retinopathy or other vitreoretinal disease and media opacities precluding adequate visualization of the fundus by fundus photography, as evaluated at the routine screening sessions.

From the photographic archives, 365 macula-centered photographs obtained in eyes of the 184 patients were retrieved. Thus, three patients were represented by photographs from only one eye each because the fellow eye was blind. Among the remaining 365 eyes, four eyes in four patients had previously been exempted from screening because media opacities precluded adequate visualization of the fundus (two eyes with cataract, one eye with vitreous hemorrhage, and one eye with asteroid hyalosis). Photocoagulation treatment had been performed in both eyes in 43 patients and in one eye in 12 patients. Three eyes in three patients were not included because of decentration of the photographs. One fundus photograph from each of the remaining 260 eyes in 137 patients was included in the study. It follows that 14 of the 137 patients were represented by one eye, because photocoagulation had been applied in the fellow eye (eight patients), the patient had one blind eye (one patient), the photographs of the fellow eye were decentered (two patients), or the eye had previously been exempted from photographic screening because of media opacities (three patients).

The Steno Diabetes Center serves more than 4000 patients with type 1 and more than 1000 patients with type 2 diabetes from the greater Copenhagen area. More than 80% of the patients are systematically screened for diabetic retinopathy by fundus photography under a predefined clinical protocol that was established 12 years before the initiation of the study.

The present study was entirely retrospective and complied with the tenets of the Declaration of Helsinki. The study did not involve patients or biological samples and as such did not require institutional review under Danish law.

### Photography

Fundus photography was made after pupil dilation with one or more drops of phenylephrine hydrochloride 2.5% and/or tropicamide 1%, a fundus camera (CF-60UV; Canon Europa NV, Amstelveen, The Netherlands) set at 60° angular field of view, and 35-mm color transparency film (Ektachrome Elite 100; Eastman Kodak, Rochester, NY). A stereoscopic set of photographs was recorded, if possible, from of the macular region of each eye.

### Digitization

The fundus photographic 35-mm diapositive slides were digitized (Coolscan LS-2000 film-slide digitizer; Nikon Corp., Tokyo, Japan) at 1350 dpi and 12 bits per color channel. 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 full image size was  $1947 \times 1296$  pixels, and the diameter of the circular fundus region defined by the internal mask of the camera was approximately 1530 pixels. The images were displayed on a 21-in. color monitor (Professional Series PT813; ViewSonic Corp., Walnut, CA).

### Routine Retinopathy Grading

The slides were originally graded in November and December of 1999 by visual examination of slide-mounted transparency pairs. Ideally, all eyes were represented by stereoscopic pairs of macula-centered fundus photographs. Stereoscopic examination was not possible when no stereo separation was visible or when one of the photographs from the stereo pair was of inferior quality. The grading was made by a retina specialist who assigned each macula-centered photographic field a routine screening grade of diabetic retinopathy, using a multilevel grading system, of which the lowest retinopathy grade was one to five microaneurysms per eye in the absence of other elements of nonpro-

liferative diabetic retinopathy. The eyes were classified as to whether diabetic retinopathy was present or absent in the macula-centered photographic field. The routine screening was conducted before initiation of the study and is representative of the high-volume grading of the institution. The average net time available for review of a set of fundus photographs from both of a patient's eyes has previously been found to be 10 minutes (Larsen N, unpublished data, 2001). The routine screening grades, the indication for photocoagulation, and other clinical information were retrospectively obtained from the clinical database.

### Reference Retinopathy Grading

A reference grading of all 260 macula-centered slides was made by an experienced ophthalmologist given the task of classifying each eye as having or not having diabetic retinopathy. The photographs were shown in random order of presentation. The reference grader was allowed as much time as needed to review the slides, and the task was repeated after an initial training session including the entire series. The reference grading was made with the aid of magnifiers and a slide projector.

### Automated Lesion Detection

A commercial computer software system for automated fundus image analysis (RetinaLye System; Retinalyze A/S, Hørsholm, Denmark) was used to provide automated red microaneurysm and hemorrhage lesion detection. The system uses advanced modeling of the gray-level image function of digital images, primarily the green color channel, with custom-developed algorithms that identify dark fundus areas in the color fundus image.<sup>1</sup> Only areas that are not contiguous with the branches of the central retinal artery or vein and are not within  $\frac{1}{4}$  optic nerve head diameter distance of the optic nerve head were considered.

A single, user-supplied parameter, the visibility threshold, controls the lesion-detection sensitivity of the system. The visibility threshold is the cutoff level for identification of a lesion in a three-dimensional matrix representation of the fundus image after complex numerical filtering, including adjustment for uneven illumination, and removal of the optic nerve head and the trunk vessels of the retina. Practical considerations and improvements have motivated the adjustment of the default visibility threshold from 1.2 to 1.0. The former value was used in a previous study.<sup>1</sup> The default visibility threshold is the value used by the system when the user has not explicitly specified the threshold. The default visibility threshold represents a preliminary optimization for use in screening of patients with diabetes for diabetic retinopathy, at a high sensitivity priority for detection of patients with retinopathy. No official approval for such use has been obtained at this writing. A scale parameter has also been introduced in the automated lesion-detection algorithm, to specify the resolution of the digital fundus image. This is given as the expected size of the diameter of the optic nerve head measured in pixels (vertical optic nerve head diameter). The default value of the optic nerve head scale parameter of 200 pixels was used throughout the study.

In the first round of analysis, the visibility threshold of the system was used at its default setting. After completion of the first round of analysis, the capacity of the system to adapt to variations in screening priorities was tested by adjusting the visibility threshold to fit the two different visual grading strategies represented in the study. First, the visibility threshold was adjusted to fit the high sensitivity of the reference grading. The objective was to measure the loss of specificity when the automated lesion-detection system was set to achieve per-eye sensitivity above 90%. Second, the visibility threshold was adjusted to fit the sensitivity of the automated lesions detection to that of the routine grading. In other words, the sensitivity of the automated lesion detection was set equal to that of the routine grader, thus enabling a one-dimensional comparison of performance by comparing the respective specificities of the computer and the routine grader.

## Computer-Assisted Retinopathy Grading

Until the completion of the study procedures, the routine grader was involved only in routine grading. Finally, the routine grader was given the task of reviewing all 260 fundus photographs included in the study on the computer display. The review was assisted by the automated lesion-detection system, which displayed the fundus image and the outline of each automatically identified lesion as identified at the high-sensitivity setting, as described earlier. This outline was shown as an overlaid image that could be toggled on and off at the discretion of the grader. This part of the study was appended after completion of the primary studies to explore the potential for methodological development of the automated lesion-detection algorithm.

## Statistical Analysis

Patients without prior photocoagulation treatment in either eye 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 either of the patient's eyes. The per-patient performance of the automated lesion-detection algorithm was characterized by the specificity, sensitivity, and accuracy—that is, the fraction of the entire population that was correctly classified—as measured against the high-sensitivity reference grading categorizing patients as having no diabetic retinopathy or any diabetic retinopathy. The receiver operating characteristic (ROC) described the relationship between sensitivity and specificity, with the area under the curve serving as a general measure of performance.<sup>2</sup>

Eyes without prior photocoagulation treatment were classified as having automatically detected diabetic retinopathy if a single red lesion of any type was identified by the automated lesion-detection algorithm. The per-eye performance of the algorithm was characterized by the specificity, sensitivity, and accuracy measured against the high-sensitivity reference grading categorizing eyes as having no diabetic retinopathy or any diabetic retinopathy.

The routine screening grading was collapsed into a binary classification of no diabetic retinopathy and any diabetic retinopathy. The specificity, sensitivity, and accuracy of the routine grading were measured against the high-sensitivity reference grading.

## RESULTS

### Automated Per-Patient Screening

In the proposed screening scenario in which only patients who had not undergone photocoagulation treatment in either eye attended automated photographic screening, the automated lesion detection correctly identified 90.1% of the 129 patients with untreated diabetic retinopathy in either eye and 81.3% of the patients with no retinopathy (Table 1). The overall accuracy was 86.8%. The proportion of false-negative grading (i.e., patients with retinopathy who were not identified by the algorithm) was 9.9%, whereas that of false-positive grading was 18.8%. All patients with false-negative grading had been reference-graded as having questionable retinopathy or minimal nonproliferative retinopathy with red lesions only.

The ROC demonstrated the range of options available in balancing the sensitivity and specificity of the algorithm (Fig. 1). The area under the per-patient ROC curve was 93.6%.

### Automated High-Sensitivity Per-Eye Grading

When compared with the expert reference grading, the automated analysis of the 260 color fundus photographs of nonphotocoagulated eyes correctly identified 93.1% of eyes with retinopathy and 71.6% of eyes without retinopathy (Table 2). The overall accuracy was 83.5%. This result was achieved with the use of a lower visibility threshold (i.e., a higher lesion-detection sensitivity) than the default setting of the automated analysis system. The 10 eyes with false-negative identification

**TABLE 1.** Automated Fundus Lesion Detection Versus Reference Grading: Per-Patient Analysis of Patients without Prior Photocoagulation in Either Eye

Expert Scrutiny	Automatically Detected Red Lesions (At Recommended Screening Settings)		Total
	0	>0	
No DR	39 (81.3)	9 (18.8)	48
DR	8 (9.9)*	73 (90.1)	81
Total	47	82	129

Data are number of patients, with the percentage of the entire group in parentheses. Performance characteristics of automated red lesion detection on color fundus photographs from nonphotocoagulated diabetic eyes. The automated segregation between patients was performed at the default settings of the system at visibility threshold 1.00 and image-scale parameter of the expected diameter of the optic nerve head at 200 pixels (default setting). Comparison was made with a reference grading made for the specific study purpose of achieving high sensitivity in the detection of diabetic retinopathy. Automated grading of color fundus photographs by single-lesion detection and classification per patient into individuals with or without diabetic retinopathy in either eye demonstrated a sensitivity of 90.1% and a specificity of 81.3%. The overall accuracy was 86.8%.

\* Questionable or minimal nonproliferative diabetic retinopathy (eight patients).

had questionable or minimal nonproliferative retinopathy (eight eyes), hard exudates without the presence of red lesions (one eye), and proliferative diabetic retinopathy (one eye). The area under the per-eye ROC curve was 94.1%.

### Routine Visual Per-Eye Grading

When comparing the routine visual grading of the 260 color fundus photographs of nonphotocoagulated eyes with the expert reference grading, the routine grading correctly identified 76.4% of eyes with retinopathy and 98.3% of eyes without retinopathy (Table 3). The overall accuracy was 86.2%. The proportion of false-negative classifications (i.e., eyes with retinopathy not identified by the routine grader) was 23.6%, whereas that of false-positives was 1.7%.

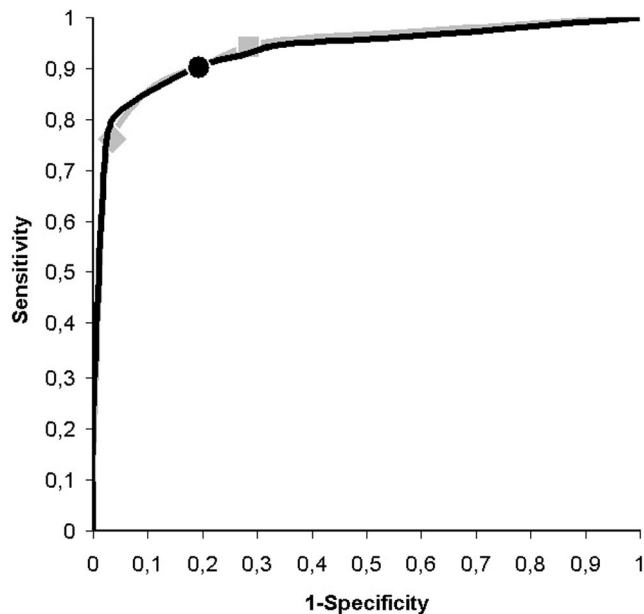
### Automated High-Specificity Per-Eye Grading

When adjusting the visibility threshold to the value that most closely reproduced the sensitivity of the routine grader, the algorithm led to the correct identification of 76.4% of the eyes with diabetic retinopathy and 96.6% of the eyes without retinopathy (Table 4). The overall accuracy was 85.4%. These results, as well as the proportions of false-positive (3.4%) and false-negative (23.6%) classifications, were not significantly different from those of the routine grader. The 34 false-negative eyes had questionable or minimal nonproliferative retinopathy (31 eyes), hard exudates without the presence of red lesions (one eye), and proliferative diabetic retinopathy (two eyes).

### Computer-Assisted Visual Per-Eye Grading

The automated high-sensitivity single-lesion detection was used in a visual regrading session in which the same routine grader examined the 260 fundus photographs on a computer display with mandatory consideration of the annotated image—that is, the digitized fundus image with overlaid lesion outlines in contrasting color. The grader was allowed to use image zooming and stretching facilities and to toggle lesion outlines off and on as needed. This resulted in correct identification of 91.0% of eyes with retinopathy and 95.7% of eyes without retinopathy





**FIGURE 1.** ROC (i.e., sensitivity as a function of 1 - specificity) of an automated red lesion-detection algorithm. For any given sensitivity, which follows from the choice of visibility threshold, the specificity is determined by the curve function, and vice versa. The ROC per patient (*black line*) relates to the automated diagnosis of the presence of diabetic retinopathy in either of a patient's eyes. The curve demonstrates the continuum of possible combinations of sensitivity and specificity as well as the combination resulting from the default visibility threshold setting (*filled circle*). The area under the per-patient ROC curve was 93.6%. The ROC per eye (*gray line*) demonstrates the characteristic in diagnosing the presence of diabetic retinopathy per eye. Alternate settings of the algorithm were applied to demonstrate the capacity of the system to adapt to variations in screening priorities. First, the automated lesion-detection system was set to achieve per-eye sensitivity above 90% (*square*). Then, the visibility threshold was adjusted to fit the sensitivity of the automated lesion detection to that of the routine grading (*diamond*). The area under the per-eye ROC curve was 94.1%.

(Table 5). The overall accuracy was 93.1%, a clear improvement over the performance of both routine visual grading and the automated system alone, with a proportion of false negatives of 9.0% and of false positive, 4.3%.

As in our previous study,<sup>1</sup> false-positive automatically detected lesions were found mainly in fundus regions with bright areas of nerve fibers striation interspersed with small vessel segments that were not identified as branchings off the main vascular trunks of the retina or in relation to areas with intense posterior hyaloid reflexes.

## DISCUSSION

The present study documented that computerized image analysis can achieve considerable effectiveness in distinguishing patients with diabetic retinopathy from patients without retinopathy. The sensitivity of the algorithm is adjustable, allowing for adaptation to specific screening protocols. At the default sensitivity setting, 73 of 81 patients (90.1%) with retinopathy were correctly identified by the algorithm in a population of patients with diabetes who had never undergone photocoagulation treatment. All patients with false-negative grading in this study had questionable retinopathy or minimal nonproliferative retinopathy with red lesions only. The potential risk of missing a diagnosis of such levels of retinopathy can be estimated on the basis of previous studies of the rate of progres-

**TABLE 2.** Automated Fundus Lesion Detection Versus Reference Grading, Per-Eye Analysis

Expert Scrutiny	Automatically Detected Red Lesions (High-Sensitivity Setting)		Total
	0	>0	
No DR	83 (71.6)	33 (28.4)	116
DR	10 (6.9)*	134 (93.1)	144
Total	93	167	260

Data are number of eyes, with the percentage of the entire group in parentheses. Performance characteristics of automated red lesion detection on color fundus photographs from nonphotocoagulated eyes in diabetic patients. The automated lesion detection was performed at increased single-lesion detection sensitivity by lowering the visibility threshold from default 1.00 to 0.80, the latter value being used. Comparison was made with a reference grading made for the specific study purpose of achieving high sensitivity in the detection of diabetic retinopathy. Automated grading by single-lesion detection demonstrated a sensitivity of 93.1% and a specificity of 71.6%. The overall accuracy was 83.5%.

\* Questionable and minimal nonproliferative diabetic retinopathy with only red lesions (eight eyes), minimal nonproliferative diabetic retinopathy with only hard exudates (one eye), and proliferative diabetic retinopathy (one eye).

sion of retinopathy.<sup>3-5</sup> In the Early-Treatment Diabetic Retinopathy Study (ETDRS) the risk of progression to treatment requiring retinopathy over a 2-year period was below 1%.

The United Kingdom Prospective Diabetes Study (UKPDS) identified systemic parameters permitting stratification of patients with identical levels of retinopathy into subgroups with variable progression rates.<sup>5</sup> Such information, if accessible in screening clinics, may enable a refined estimation of appropriate follow-up intervals for individual patients in retinopathy screening. This should minimize the concern that lesions in the far periphery, visible only by binocular contact lens biomicroscopic examination, are missed when only photographic methods are used for retinopathy screening. Ultimately, the choice of screening method must be based on an overall assessment of the options available in a given health care environment.

In the present study, we demonstrated the practicability of using automated digital detection of red lesions in photographic diabetic retinopathy screening. The objective was to sort patients who have not undergone photocoagulation treatment for diabetic retinopathy in two groups: those without retinopathy and those with possible or definite retinopathy who must have visual grading of their fundus photographs.

**TABLE 3.** Routine Visual Screening Versus Reference Grading: Per-Eye Analysis

Expert Scrutiny	Routine Visual Grading		Total
	No DR	DR	
No DR	114 (98.3)	2 (1.7)	116
DR	34 (23.6)	110 (76.4)	144
Total	148	112	260

Data are number of eyes, with percentage of the entire group in parentheses. Performance characteristics of routine visual grading of diabetic retinopathy on color fundus photographs from nonphotocoagulated eyes in diabetic patients. Comparison was made with a reference grading made for the specific study purpose of achieving high sensitivity in the detection of diabetic retinopathy. The routine visual grading made for screening purposes 1 year before initiation of the study, demonstrated a sensitivity of 76.4% and a specificity of 98.3%. The overall accuracy was 86.2%.

An automated prescreening system that automatically categorizes patients with diabetes into those who should have their photographs visually graded and those whose photographs can reliably be handled automatically should be able to accomplish a series of functions. These include identification of images that are unsuitable for automated lesion detection (i.e., low-quality images caused by refractive media opacity or technical failures in the photographic process). Photographs with such characteristics should be automatically designated to visual inspection. Indeed, a preliminary evaluation of an automated quality-assessment algorithm developed subsequent to the completion of this study has demonstrated a significant potential for increasing the sensitivity of the automated classification by elimination of several of the false-negative gradings in this study, one of which had proliferative diabetic retinopathy (Table 2). Additional enhancement appears to be possible by incorporating automated detection of bright fundus lesions. Thus, one eye with a false-negative grade (Table 2) had bright lesions only.

An ancillary finding was that the quality of high-volume visual routine grading was enhanced considerably when the attention of the routine grader was directed to potential fundus lesions by the automated lesion-detection system. This part of the study was appended after completion of the primary studies. The results should be considered indicative of a potential for methodological development, but the tentative conclusion should be verified in a larger prospective study.

Clinical ophthalmology emphasizes the detailed assessment of retinopathy at or above treatment thresholds. The task of making a fine distinction between presence and absence of any retinopathy is rarely addressed outside epidemiologic studies. Clinical ophthalmology practice is not traditionally subjected to validation or quality control, and it is therefore difficult to find standards against which new methods can be compared.

TABLE 4. Automated Fundus Lesion Detection Versus Reference Grading: Per-Eye Analysis

Expert Scrutiny	Automatically Detected Red Lesions (High-Specificity Setting)		Total
	0	>0	
No DR	112 (96.6)	4 (3.4)	116
DR	34 (23.6)	110 (76.4)	144
Total	146	114	260

Data are the number of eyes, with the percentage of the entire group in parentheses. Performance characteristics of automated red lesion detection on color fundus photographs from nonphotocoagulated eyes in diabetic patients. The automated lesion detection was performed at increased single-lesion detection specificity by increasing the visibility threshold from default 1.00 to 1.55. Comparison was made with a reference grading made for the specific study purpose of achieving high sensitivity in the detection of diabetic retinopathy. Automated red lesion detection resulted in a sensitivity of 76.4% and a specificity of 96.6%. The overall accuracy was 85.4%.

TABLE 5. Computer-Assisted Visual Grading Versus Reference Grading: Per-Eye Analysis

Expert Scrutiny	Computer-Assisted Visual Grading		Total
	No DR	DR	
No DR	111 (95.7)	5 (4.3)	116
DR	13 (9.0)	131 (91.0)	144
Total	124	136	260

Data are the number of eyes, with the percentage of the entire group in parentheses. Performance characteristics of visual grading of diabetic retinopathy in color fundus photographs from nonphotocoagulated eyes in diabetic patients when assisted by the automated red lesion detection (same grader as in Table 3, 2 years after the original grading session). Comparison was made with a reference grading made for the specific study purpose of achieving high sensitivity in the detection of diabetic retinopathy. When prompted by computer-identified lesion markings, the routine grader achieved a sensitivity of 91.0% and a specificity of 95.7%. The overall accuracy was 93.1%.

Even when validation is attempted, there is usually no quality-assurance system in place to ensure the widespread application of certified standards of clinical ophthalmology. It appears that technological progress may enable application of such systems, even in complex clinical fields such as ophthalmology.

Populations that undergo systematic fundus photographic screening have been found to have a favorable and improving visual prognosis.<sup>6</sup> Consequently, the task of automatically assisting the separation of patients with retinopathy from patients without retinopathy is a clinically meaningful one, and our results warrant further study of the method's practical applicability in clinical screening practice.

## References

- Larsen M, Godt J, Larsen N, et al. Automated detection of fundus photographic red lesions in diabetic retinopathy. *Invest Ophthalmol Vis Sci.* 2003;44:761-766.
- Metz CE. Basic principles of ROC analysis. *Semin Nucl Med.* 1978; 8:283-298.
- 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.
- 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.
- 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.
- Rossing K, Jacobsen P, Rossing P, Lauritzen E, Lund-Andersen H, Parving HH. Improved visual function in IDDM patients with unchanged cumulative incidence of sight-threatening diabetic retinopathy. *Diabetes Care.* 1998;21:2007-2015.