Computer Classification of Nonproliferative Diabetic Retinopathy

Samuel C. Lee, PhD; Elisa T. Lee, PhD; Yiming Wang, MS; Ronald Klein, MD; Ronald M. Kingsley, MD; Ann Warn, MD

Objective: To propose methods for computer grading of the severity of 3 early lesions, namely, hemorrhages and microaneurysms, hard exudates, and cotton-wool spots, and classification of nonproliferative diabetic retinopathy (NPDR) based on these 3 types of lesions.

Methods: Using a computer diagnostic system developed earlier, the number of each of the 3 early lesions and the size of each lesion in the standard photographs were determined. Computer classification criteria were developed for the levels of individual lesions and for NPDR. Evaluation of the criteria was performed using 430 fundus images with normal retinas or any degree of retinopathy and 361 fundus images with no retinopathy or the 3 early lesions only. The results were compared with those of the graders at the University of Wisconsin Ocular Epidemiology Reading Center and an ophthalmologist.

Main Outcome Measures: Agreement rates in the classification of NPDR between the computer system and human experts.

Results: In determining the severity levels of individual lesions, the agreement rates between the computer system and the grading reader were 82.6%, 82.6%, and 88.3% using the 430 images and 85.3%, 87.5%, and 93.1% using the 361 images, respectively, for hemorrhages and microaneurysms, hard exudates, and cotton-wool spots. When the “questionable” category was excluded, the agreement rates were 86.5%, 92.3%, and 91.0% using the 430 images and 89.7%, 96.3%, and 97.4% using the 361 images. In classifying NPDR, the agreement rates between the computer system and the ophthalmologist were 81.7% using the 430 images and 83.5% using the 361 images.

Conclusions: The proposed criteria for computer classification produced results that are comparable with those provided by human experts. With additional research, this computer system could become a useful clinical aid to physicians and a tool for screening, diagnosing, and classifying NPDR.


Nonproliferative diabetic retinopathy (NPDR) is characterized by structural damage to small retinal blood vessels, causing them to dilate, leak, or rupture. Visible retinal lesions include microaneurysms, at first alone but later accompanied by 1 or more of the following: dot or blot hemorrhages, hard exudates (HE), soft exudates or cotton-wool spots (CWS), intraretinal microvascular abnormalities (IRMA), and venous beading. Nonproliferative diabetic retinopathy can be classified into mild (or early), moderate, and severe stages depending on the severity of the lesions.

Early diagnosis and accurate staging are essential prerequisites for effective treatment of diabetic retinopathy and the reduction of visual disability risk. Advancing technology in the computer analysis of retinal images now has major efficiency and cost-effectiveness potential to improve the detection and selection of at-risk patients. Digital fundus imaging makes computer diagnosis and classification of diabetic retinopathy possible. A few studies have investigated whether computer systems can be as reliable as humans in the classification and grading of diabetic retinopathy or whether human diagnosis and classification can be improved or aided by the computer.1-3

Computer diagnostic systems probably will not replace the current screening modalities, as suggested by the American Diabetes Association.4 They will, however, provide a modality that greatly reduces subjective human variation and the cost of the current screening process and therefore facilitate a more objective and economical method of diagnosis. In addition, the technology offered by the computer system will likely be an aid to human experts to produce diagnostic results of higher quality and accuracy.

We have developed a computer system using image-processing and pattern recognition techniques to grade the severity of early diabetic retinopathy lesions. The validation of this system and its potential as a clinical aid is the subject of this report.

Author Affiliations: School of Electrical and Computer Engineering, University of Oklahoma, Norman (Dr S. C. Lee); Center for American Indian Health Research, University of Oklahoma Health Sciences Center, Oklahoma City (Dr E. T. Lee and Mr Wang); Department of Ophthalmology, University of Wisconsin Medical School, Madison (Dr Klein); and Dean A. McGee Eye Institute, Oklahoma City (Dr Kingsley), and Lawton, Okla (Dr Warn).

Financial Disclosure: None.
recognition techniques to detect 3 early lesions of NPDR, namely, retinal hemorrhages and microaneurysms (HMA), HE, and CWS, from color fundus slides or real-time images directly transmitted from the fundus camera to the computer. The computer system was developed for fundus photographs taken at 45° and between standard fields 1 and 2 (referred to as “1-2 field,” with both the optic disc and macula present). It can also be applied to real-time images. The diagnostic criteria of a local retinal specialist (R.M.K.) and fundus images from a large number of American Indians of southwestern Oklahoma were used to develop the system. The performance of the system was found to be closely comparable with that of a general ophthalmologist (A.W.), a retinal specialist, and certified graders at the University of Wisconsin Ocular Epidemiology Reading Center (UWOERC).  

Of the various available methods for the classification of diabetic retinopathy, grading of fundus photographs using Standard Photographs (SPs) has become an accepted method in clinical studies. The 7 standard fields used in the Early Treatment Diabetic Retinopathy Study (ETDRS) and the modified Airlie House classification system are commonly used to classify the level of severity of retinopathy. In 1986, an alternative method of categorizing diabetic retinopathy into 8 levels was proposed. However, in many epidemiologic studies, only 1 fundus photograph is taken of a field of the retina, and therefore classification criteria involving 7 fields are not applicable and must be modified.

We propose a method for computer classification of severity levels of HMA, HE, and CWS and for classification of NPDR. This method is based on a modification of the ETDRS criteria. The performance of this computer classification system is evaluated by comparing the computer results with those of the UWOERC and a general ophthalmologist.

**METHODS**

Using image-processing and pattern recognition techniques including noise removal and image normalization, a computer vision system for the diagnosis of 3 early diabetic retinopathy lesions, HMA, HE, and CWS, was developed. The system detects these lesions on the basis of the color contrast between lesions and the retinal background. The diagnosis also depends on the shape of the lesion. The development of this computer system was part of an epidemiologic study of eye disease called Vision Keepers. Participants in Vision Keepers were American Indians who resided in southwestern Oklahoma. Informed consent was obtained from all participants to allow their fundus photographs to be used in the development of a computer vision system for the diagnosis of retinopathy. The study was approved by the University of Oklahoma Health Sciences Center Institutional Review Board.

Fundus photographs were taken by a 45° camera (CR5-45 NM; Canon, Tokyo, Japan) and included SP fields 1 and 2, with both the disc and macula present. These photographs were used as training images in the development of the computer system. To use it, a fundus photograph (35-mm slide) was first scanned and digitized using the Nikon LS1000 scanner (Melville, NY). The dimension of the digitized images was 512 × 512 pixels (screen resolution). The digitization resolution was 10 000 pixels per inch. The digitized image was then processed using the computer vision system. It has been shown that the diagnostic results obtained from the computer system are comparable with those obtained by human experts' in agreement rates, sensitivity, and specificity. We have expanded the capability of the computer system to classify the individual lesions and NPDR into levels of severity.

The grading criteria used by the UWOERC for the 3 early retinal lesions are based on the ETDRS protocol and SPs 2A and 3, as follows:

- **HMA**: the levels are 0, none; 1, questionable; 2, <SP 2A; and 3, ≥SP 2A.
- **HE**: the levels are 0, none; 1, questionable; 2, <SP 3; and 3, ≥SP 3.
- **CWS**: the levels are 0, none; 1, questionable; and 2, present.

The “questionable” category was used only when the grader thought that the likelihood of the lesion being present was 50% or greater but less than 90%. When an abnormality was definitely present but its nature was uncertain, the grader assigned the grade of “questionable” for the lesion considered to be most likely and “none” for the lesions considered less likely.

The grading criteria used in this study were based on a single 45° photographic field, whereas the ETDRS classification requires 7 standard 30° fields. The ETDRS classification criteria were modified. The major difference was that the ETDRS mid-ninproliferative levels were collapsed into 2 levels owing to the inability to count the number of fields involved with specific lesions. The ETDRS classification gives weight to multiple fields having a lesion present, but the criteria used in this study categorize retinopathy severity based on the lesion amount present in 1 broader field (R.K. and S. Meuer, written communication, 2004). Although it is difficult to compare the 2 methods directly, the scale in this study has been used in many large epidemiologic studies; for example, the Third National Health And Nutrition Examination Survey.

To develop a set of computer classification criteria that were similar to these criteria, we digitized the SPs, identified the 3 lesions using the developed computer system, and measured the size of each HMA and HE in pixels. The total lesion area of all HMAs in SP 2A was found to be 1717 pixels, and the total area of all HEs in SP field 3 was 86 pixels. The SPs were taken using a Zeiss camera (Carl Zeiss, Oberkochen, Germany) at 30°. The mean optic disc diameter of the 30° photographs was 4.7 mm; for photographs taken by the Canon 45° camera, the mean optic disc diameter was 3.4 mm, as measured on the film by the UWOERC. Therefore, the 45° photographs were approximately 72.3% magnification compared with the 30° pictures. Using this fraction, the 30°-to-45° area conversion ratio was 0.723=0.527. We also measured the diameter of the optic disc of the digitized 30° SPs and the digitized 45° images from our study population, the American Indians. The mean diameter was found to be 111 pixels for the 30° SPs and 80 pixels for the 45° photographs. The resultant resolution was approximately 19 µm per pixel. The diameter conversion ratio was 0.7207 and the area conversion ratio was approximately 0.52, which was in agreement with the conversion ratio estimated by the UWOERC. We used 0.52 as the scaling factor in converting lesion areas in the 30° SPs to a 45° equivalent.

Using this scaling factor, the total area of HMA in SP 2A and the total area of HE in SP 3 were found to be equivalent to 893 pixels and 45 pixels, respectively, in a 45° photograph. On the basis of these converted values, we propose the following criteria for computer grading:

- **HMA**: 0, none detected; 1, questionable; 2, HMA detected and the total area is <893 pixels; and 3, HMA detected and the total area is ≥893 pixels.
- **HE**: 0, none detected; 1, questionable; 2, HE detected and...
the total area is <45 pixels; and 3, HE detected and the total area is ≥45 pixels.

- CW5: 0, none detected; 1, questionable; and 2, present.

The "questionable" category was assigned when the lesions were very small and therefore suggestive of NPDR. The classification also depended on whether HMA was present in the cases of HE and CWs. For HMA, very small lesions (area ≤4 pixels) were classified as questionable HMA. If the image contained only questionable HMA, it was classified as showing questionable HMA. If the image had no HMA or only questionable HMA, the image was considered as showing questionable HE even if HE was detected by the computer system. If the image was classified as showing definite HMA and none of the detected HE in the image was greater than 4 pixels, the image was classified as showing questionable HE. For CWs, if 1 or more CWs was detected by the computer system and the image had no HMA or only questionable HMA, the image was classified as showing questionable CWs.

This set of criteria was used to grade the level of severity of the 3 early retinal lesions in 430 images. None of these selected slides were used in developing the computer system. These images were selected from more than 1000 American Indians in southwestern Oklahoma who participated in an epidemiologic study of eye disease called Vision Keepers. In this study, participants were given a funduscopic examination by a general ophthalmologist using indirect ophthalmoscopy. In addition, for each participant, 2 fundus photographs were taken of each eye using a Canon 45° CR5 camera and focusing on the area between fields 1 and 2, including the optic disc and macula. The better image of each eye was sent to the UWOERC for grading. The overall rate of diabetic retinopathy in this cohort was 23%. The quality of each eye was sent to the UWOERC for grading. The overall agreement rate and κ statistic was 0.55. Twenty-two (5.1%) and 38 images (8.8%), respectively, were classified by the UWOERC grader and the computer system as showing questionable (level 1) HMA. The false-positive rate of the computer system (excluding level 1) for HMA was 7.7% (22/284), and the false-negative rate (excluding level 1) was 18.9% (25/132). When the "questionable" category was ignored, the agreement rate was 86.5% and the κ statistic was 0.70, indicating very good agreement between the 2 diagnostic methods. Major causes of misclassification were that (1) the HMA was too small, (2) the color intensity of the HMA was too close to the background or was low contrast, (3) the HMA was adjacent to a vessel, and (4) the shape of the HMA was long and thin, which could be mistaken as a vessel by the computer system.

We first report the results for individual lesions. Table 1 gives the comparisons of computer classification results from the 430 images with those of the UWOERC readers. For HMA, the overall agreement rate was 82.6% with a κ statistic of 0.63, indicating good agreement. Fourteen (3.3%) and 10 images (2.3%), respectively, were classified by the UWOERC grader and the computer system as showing questionable (level 1) HMA. The false-positive rate of the computer system (excluding level 1) for HMA was 7.7% (22/284), and the false-negative rate (excluding level 1) was 18.9% (25/132). When the "questionable" category was ignored, the agreement rate was 86.5% and the κ statistic was 0.70, indicating very good agreement between the 2 diagnostic methods. Major causes of misclassification were that (1) the HMA was too small, (2) the color intensity of the HMA was too close to the background or was low contrast, (3) the HMA was adjacent to a vessel, and (4) the shape of the HMA was long and thin, which could be mistaken as a vessel by the computer system.

For HE, the overall agreement rate was 82.6% and the κ statistic was 0.55. Twenty-two (5.1%) and 38 images (8.8%), respectively, were classified by the UWOERC grader and the computer system as showing questionable HE. The computer system categorized a total of 47 images as showing questionable, level 2, or level 3 HE, whereas the UWOERC found them free of HE; among these, 30 were classified as questionable by the computer. Most of the other misclassifications were due to the computer's calling drusen, light reflection, and artifacts HE. On the other hand, UWOERC graders classi-
UWOERC Grader

<table>
<thead>
<tr>
<th>Computer</th>
<th>Level 0</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMA†</td>
<td>255</td>
<td>10</td>
<td>25</td>
<td>0</td>
<td>290</td>
</tr>
<tr>
<td>Level 0</td>
<td>255</td>
<td>10</td>
<td>25</td>
<td>0</td>
<td>290</td>
</tr>
<tr>
<td>Level 1</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Level 2</td>
<td>22</td>
<td>2</td>
<td>94</td>
<td>5</td>
<td>123</td>
</tr>
<tr>
<td>Level 3</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>284</td>
<td>14</td>
<td>123</td>
<td>9</td>
<td>430</td>
</tr>
</tbody>
</table>

HE†

| Level 0 | 303 | 9 | 1 | 314 |
| Level 1 | 30   | 7 | 1 | 38  |
| Level 2 | 15   | 5 | 0 | 43  |
| Level 3 | 2    | 1 | 2 | 35  |
| Total   | 350  | 22 | 19 | 390 |

CWS‡

| Level 0 | 319 | 11 | 15 | NA  |
| Level 1 | 11   | 4  | 0  | NA  |
| Level 2 | 10   | 3  | 56 | NA  |
| Total   | 340  | 18 | 71 | 429 |

Abbreviations: CWS, cotton-wool spots; HE, hard exudates; HMA, hemorrhages and microaneurysms; NA, not applicable; UWOERC, University of Wisconsin Ocular Epidemiology Reading Center.

*For HMA, the agreement rate was 82.6% or 86.5% (with or without level 1: questionable); \(\kappa = 0.63\) or 0.70 (with or without level 1: questionable).

†For HE, the agreement rate was 82.6% or 92.3% (with or without level 1: questionable); \(\kappa = 0.55\) or 0.74 (with or without level 1: questionable).

‡For CWS, the agreement rate was 88.3% or 91% (with or without level 1: questionable); \(\kappa = 0.65\) or 0.78 (with or without level 1: questionable).

The UWOERC had 1 case in which CWS was ungradable.

fied a total of 58 images as showing level 2 or level 3 HE, and the computer system misclassified only 3 (5.2%) as showing level 0 or level 1 HE. When the “questionable” category was excluded, the overall agreement rate and the \(\kappa\) statistic increased to 92.3% and 0.74, respectively, and the false-positive rate was only 5.3% (17/320), indicating excellent agreement.

The agreement rate was 88.3% between the classifications of CWS severity level by the computer and those of the UWOERC graders. The \(\kappa\) statistic was 0.65. The UWOERC graders gave 18 “questionable” diagnoses (4.2%), and the computer system gave 15 (3.5%). Of the 340 images classified by the UWOERC as showing no CWS, 21 were categorized as either questionable (11 images) or CWS-positive (10 images), for a false-positive rate of 6.2%. Of the 71 images with a classification of CWS by the UWOERC, the computer system missed 15 (21.1%). When the “questionable” category was omitted, the agreement between UWOERC graders and the computer system was excellent, with an agreement rate of 91.2% and a \(\kappa\) statistic of 0.78. The primary reason for computer misclassification was the vagueness of the CWS or too little contrast between CWS and the background.

For the 361 images that had the 3 early lesions only or no lesions at all, the agreement rates (Table 2) were about 3% to 5% higher (85.3%, 87.5%, and 93.1%, respectively, for HMA, HE, and CWS) than those in Table 1. However, the \(\kappa\) statistic did not improve except for CWS (0.60, 0.38, and 0.90, respectively, for HMA, HE, and CWS). When the “questionable” category was excluded, both the agreement and the \(\kappa\) statistic improved considerably. The agreement rates were 89.7%, 96.3%, and 97.4%, and the \(\kappa\) statistics were 0.69, 0.63, and 0.71, respectively, for HMA, HE, and CWS. Table 3 gives the results for the classification of NPDR in the 430 images from the study ophthalmologist and the computer system. In this table, we have excluded 16 images that were classified as showing questionable NPDR by the computer system because the ophthalmologist did not use the
“questionable” category in her diagnosis. The overall agreement rate was 81.2%, and the $\kappa$ statistic was 0.67. Of the 284 images that were categorized by the ophthalmologist as showing no NPDR, 27 were classified by the computer as showing early NPDR and 6 as moderate NPDR, for a 11.6% overall false-positive rate. Of the 79 images categorized as showing early NPDR, 26 (32.9%) were classified as showing no NPDR by the computer. However, in the moderate and severe groups, the computer made very few mistakes. Of the 361 images that had no lesions or the 3 early lesions only, 11 were classified as showing questionable NPDR by the computer system. Omitting these 11 images, Table 4 shows a slightly higher agreement rate of 83.5%.

Using pattern recognition and image-processing techniques in medical diagnosis has been successful in many areas. One example is the machine interpretation of electrocardiograms. This technique has been a useful aid to physicians to provide an expert interpretation, particularly when cardiologists are not readily available. Diabetic retinopathy is another disease that might benefit by machine- or computer-aided diagnosis, particularly because of the possible interobserver and intraobserver variation. As reported earlier, the agreement rates among a retinal specialist, a general ophthalmologist, and graders at a fundus reading center in diagnosing HMA, HE, and CWS ranged between 85% and 89% when classification consisted of yes/questionable/no categories and between 91% and 96% when classification consisted of only yes/no categories. Similar variation may exist in the classification of severity of individual lesions and NPDR.

Since the computer system to recognize the 3 early retinal lesions, HMA, HE, and CWS, was developed, we have attempted to expand its use to classify these 3 lesions into levels of severity. The human classification has been based on a comparison of the image under study with the SPs. Because the images were digitized into pixels before they were processed by the computer system, we also digitized the SPs and then identified the areas of various lesions. To determine whether the lesion in the image under study was more severe than in the SPs, we proposed to compare the total areas of lesions. In converting the lesion area in a 30° image to a 45° equivalent, we used a scaling factor of 0.52. The factor is an estimate. However, it was interesting to find that the magnification fraction using digital images (0.7207) was very close to that obtained by measuring the film (0.7230).

In addition to the 4 levels of NPDR defined by the ETDRS (none, early, moderate, and severe), we propose an additional level, “questionable NPDR,” which includes cases where all lesions detected by the computer are not definite or in which HE and CWS are detected without the presence of HMA. If the computer system is used alone, patients with a computer diagnosis of questionable or more severe NPDR should be referred to an ophthalmologist for further examination. Patients with a computer diagnosis of moderate or severe NPDR must be referred more urgently for confirmation and treatment.

In this study, we compared the diagnostic results of the computer with those of human graders and a general ophthalmologist using images that had some retinal lesions or had no lesions or the 3 early lesions only in the absence of other, more severe lesions characterizing NPDR. Of the 430 images that had no or some retinal lesions, the overall agreement rates for the levels of severity and the classification of NPDR were between 82.6% and 92.3% and the $\kappa$ statistics were between 0.55 and 0.78. The agreement rates and $\kappa$ statistics were higher when the “questionable” category was excluded. These agreement rates were good given that the graders at the UWOERC and the ophthalmologist often used the presence of other lesions or lesions outside of the SP fields 1 and 2 to grade or diagnose a specific lesion and the computer considered only the 3 early lesions inside the SP field 1 and 2. We also compared the results using images that had the 3 early lesions only or no lesions. Of the 361 such images, the agreement rates were between 85.3% and 97.4%.

The reproducibility of both the computer diagnosis and that of human graders is high. The computerized method is deterministically reproducible if the same digitized images are processed or if the images are transmitted to the computer directly without going through a scanner. Agreement between pairs of graders for lesions typical of diabetic retinopathy was 95% to 97%, and $\kappa$ statistics ranged from 0.72 to 0.86, indicating excellent agreement.

Our data indicate that the classification differences between the computer system and the UWOERC graders were due to the “questionable” category more than whether the image had only the 3 early lesions or other, more severe lesions. There are no uniform definitions for questionable lesions that can be used by both human experts and the computer system. The criteria used by graders were qualitative and involved subjective judgment. The criteria used by the computer system considered the small size ($\leq 4$ pixels) of the lesion and whether definite HMA was present (for HE and CWS only). The selection of 4 pixels as the cutoff point was based on the suggestion that the size of the smallest clinically visible microaneurysm is 30 µm$^3$ or a 4-pixel square in a circular image of 576 pixels in diameter. Our images were in a
circular field of 512 pixels in diameter, and therefore we considered that HMA of 4 pixels or less in area was questionable. We have also considered using lesion color or its contrast with the retinal background to determine if a lesion was questionable. However, there was no uniform standard for the color or contrast, so we chose to use the size without considering these factors. The difference in criteria used by the graders and computer made it difficult to compare the results from these 2 diagnostic modalities. To improve the diagnostic capability of the computer system and to facilitate comparisons, a set of universally accepted quantitative and/or objective criteria is needed to diagnose these questionable lesions.

Another problem is how to improve the accuracy of classification by computer to reduce the rates of false-positive and false-negative diagnoses. This question appears to be more difficult to answer than in the case of electrocardiograms, in which the processing and tracking of shape of the electrocardiographic signals need to be studied. With fundus images, it involves not only the shape but also the color of the lesions and the color contrast between lesions, vessels, and background. Considerable variations exist in the color and levels of contrast of the lesions in fundus images from one patient to another. However, the human eye can distinguish lesions from normal retinal vessels and the retinal background despite these variations. To improve the capability of a computer system to diagnose and classify diabetic retinopathy, more research is needed on how human experts make diagnostic decisions and how human intelligence can be translated into computer algorithms. In addition, data on other, more severe lesions such as IRMA, venous beading, and new vessels need to be incorporated into the computer vision system and used in the classification criteria. More research is being conducted to further improve the algorithms in computer diagnosis and classification of diabetic retinopathy, including these clinically important lesions.

In conclusion, this article describes a method for a computer system using new software to grade the severity of 3 types of early lesions (HMA, HE, and CWS) of diabetic retinopathy and a method to classify NPDR based on these 3 lesions only. Classification results from the computer system were compared with those of graders from an established fundus reading center and a general ophthalmologist. Agreements rates were good to excellent. Our results show that the computer system could become a useful aid for physicians and fundus image graders and a tool for large studies, such as screening and epidemiologic studies, in diagnosing early lesions and classifying early diabetic retinopathy. For physicians and fundus image graders, the computerized diagnosis can serve as a second opinion. The computer system could be particularly helpful when ophthalmologists are not readily available and eye examinations are performed or fundus photographs are read by physicians in other specialties. Thus, computer-assisted diagnosis of the presence and severity of diabetic retinopathy may be beneficial to those with a high risk of this disease, especially in areas where ophthalmologic expertise is lacking.

Submitted for Publication: October 28, 2003; final revision received March 29, 2004; accepted April 29, 2004.

Correspondence: Samuel C. Lee, PhD, School of Electrical and Computer Engineering, University of Oklahoma, 202 W Boyd St, Norman, OK 73019 (samlee@ou.edu).

Funding/Support: This study was supported by grant EY-09898 from the National Eye Institute, Bethesda, Md.

Additional Information: This work was performed at the Center for American Indian Health Research, University of Oklahoma Health Sciences Center.

Acknowledgment: We thank the Vision Keepers participants for allowing us to use their fundus photographs in this study. We also thank Dana Russell, Stacy Meuer, Michelle Roberts, and Cathy Morales for their assistance.