# Monoscopic versus Stereoscopic Retinal Photography for Grading Diabetic Retinopathy Severity

*Helen K. Li,*[1,2] *Larry D. Hubbard,*[3] *Ronald P. Danis,*[3] *Adol Esquivel,*[4] *Jose F. Florez-Arango,*[2,5] *and Elizabeth A. Krupinski*[6]

**Purpose.** To assess agreement between monoscopic and stereoscopic photography for research classification of the severity of diabetic retinopathy (DR).

**Methods.** Monoscopic digital (MD) images were compared with stereo digital (SD) and film (SF) photographs from a 152-eye cohort with full-spectrum Early Treatment Diabetic Retinopathy Study (ETDRS) severity levels for agreement on severity level, DR presence with ascending severity threshold, presence of DR index lesions, and repeatability of grading.

**Results.** There was substantial agreement classifying ETDRS DR severity levels between MD and SF ($\kappa = 0.65$, $\kappa_w$ [linear weighted] $= 0.87$), MD and SD ($\kappa = 0.66$, $\kappa_w = 0.87$), and SD and SF ($\kappa = 0.62$, $\kappa_w = 0.86$) images. Marginal homogeneity analyses found no significant difference between MD and SF images ($P = 0.53$, Bhapkar test). The $\kappa$ agreement between MD and SF ranged from 0.80 to 0.94 for the presence or absence of eight ascending DR severity thresholds. Repeatability between the readers of the MD images was equal to or better than that of the readers of SD or SF images. Severity threshold grading repeatability between readers was similar with the MD and SF images. The $\kappa$ agreement between MD and SF for identifying diabetic retinopathy lesions ranged from moderate to almost perfect. The $\kappa$ comparisons showed that performance of grading new vessels on the disc in MD images was slightly lower than that with the SF images.

**Conclusions.** Monoscopic photography can equal the reliability of stereo photography for full ETDRS DR severity scale grading. (*Invest Ophthalmol Vis Sci.* 2010;51:3184–3192) DOI: 10.1167/iovs.09-4886

From the [1]Department of Ophthalmology and Visual Sciences, University of Texas Medical Branch, Galveston, Texas; the [2]School of Health Information Sciences, University of Texas Health Science Center, Houston, Texas; the [3]Department of Ophthalmology and Visual Sciences, University of Wisconsin-Madison, Madison, Wisconsin; the [4]Houston Veterans Affairs Medical Center and Baylor College of Medicine, Houston, Texas; [5]Universidad De Antioquia, Medellin, Columbia; and the [6]Department of Radiology, University of Arizona, Tucson, Arizona.

Supported by a grant from Juvenile Diabetes Foundation Research International, New York, NY (HKL), and by unrestricted grants from Research to Prevent Blindness (Department of Ophthalmology and Visual Sciences, The University of Texas Medical Branch, and the Department of Ophthalmology and Visual Sciences, University of Wisconsin School of Medicine and Public Health).

Submitted for publication November 9, 2009; revised December 11, 2009; accepted December 30, 2009.

Disclosure: **H.K. Li**, None; **L.D. Hubbard**, None; **R.P. Danis**, None; **A. Esquivel**, None; **J.F. Florez-Arango**, None; **E.A. Krupinski**, None

Corresponding author: Helen K. Li, Department of Ophthalmology and Visual Sciences, The University of Texas Medical Branch, 301 University Boulevard, Galveston, TX 77555-1106; hlimed@mac.com.

Stereopsis is the ability to perceive three dimensions by merging two slightly different views of the same scene. Stereo fundus photography has been a cornerstone of diabetic retinopathy assessment since the 1968 Airlie House Symposium established the first diabetic retinopathy classification system.[1,2] Their stereo photography protocol and severity classifications were modified during the Diabetic Retinopathy Study[3–5] and were later expanded in the Early Treatment Diabetic Retinopathy Study (ETDRS).[6] Today, stereo, 30°, seven-field, 35-mm color slides remain the gold standard for clinically evaluating diabetic retinopathy. Stereo fundus photography has been used in many ongoing clinical studies including the Diabetic Retinopathy Clinical Research Network studies,[7] the Action to Control Cardiovascular Risk in Diabetes Eye Study,[8] Epidemiology of Diabetes Interventions and Complications,[9] and the Diabetes Control and Complications Trial.[10] Some telemedicine programs interested in managing diabetic retinopathy also include stereo photography.[10–12]

A stereo image is produced by taking photographs of the same scene from two slightly different positions. The distance between these positions is called the stereo base. Many fundus cameras have the ability to stereoscopically image the central and peripheral retina by sequentially taking two images. Parallax, the apparent displacement of objects in the stereo scene, is achieved by horizontally shifting the fundus camera between photographs with a joystick or other positioning device. A 2-mm stereo base is considered necessary for an adequate stereo effect.[2,11]

It is generally assumed that depth perception helps distinguish subtle extraretinal neovascularization elevated above the plane of the retina from intraretinal microvascular abnormalities (IRMAs). This discrimination is important because eyes could otherwise be misclassified on the ETDRS severity scale. Stereopsis may also aid in detecting new vessels elsewhere (NVE), new vessels on the disc (NVD), fibrous proliferations elsewhere (FPE), fibrous proliferations on the disc (FPD), preretinal hemorrhages (PRH), and vitreous hemorrhages (VH)—the diabetic vascular abnormalities found on or above the retina. Confusing these advanced abnormalities with other lesions could result in missed opportunities for timely intervention to prevent vision loss. Correct classification of the diabetic retinopathy severity level is also essential in clinical and epidemiology studies in which diabetic retinopathy progression is observed.[12] It is also believed that stereo photography's illusion of depth is useful for assessing the severity of diabetic macular edema. Detailed classification of macular edema is dependent on identifying and measuring retinal thickening.

Whether grading diabetic retinopathy or macula edema severity, stereoscopic evaluation adds burdens to the photographic protocol. For example, sufficient illumination of both images in a stereo pair is necessary to view details that provide depth cues. Left and right eye images must be equally sharp and illuminated. Maintaining focus in a sequential stereo pair

requires clear media and well-dilated pupils, preferably 6 mm or more.[2] Since dilated pupil size in diabetic subjects is often limited by autonomic neuropathy, taking quality stereo photographs can be challenging in patients with long-duration diabetes.[5,13] Photographers must have good patient skills because the patient's cooperation is needed to minimize movement between sequential photographs. Sequential stereo fundus photography doubles the number of light flashes a patient must endure (a minimum of 16 flashes to create one pair of anterior segment and seven pairs of retinal photos, per the ETDRS photography protocol).[5] After stereo photos are taken, special equipment such as optical viewers or LCD goggles is needed to review them. Correct mounting of photographs is essential for stereo viewing. Ultimately, the quality of the stereo effect is also dependent on the observer's personal ability to fuse stereoscopically. Observers with eyes of unequal visual acuity may have difficulty appreciating stereo depth.

The value of assessing diabetic retinopathy by stereo fundus photography is the subject of debate. The utility of single images versus stereo pairs in assessing ETDRS levels of diabetic retinopathy severity has not been well studied. Some clinical trials and epidemiology studies have forgone stereo photography, such as the Liverpool Diabetes Eye Study,[14] the UK Prospective Diabetes Study,[15] and the EURODIAB IDDM Complications Study.[15] There is no consensus on how stereo photography affects diagnostic accuracy in diabetic retinopathy[13] or whether its added burden and cost are justified. Determining whether stereo photography is a critical factor in the optimal assessment of diabetic retinopathy could resolve important questions of need and practicality. Similar to grading of the severity of diabetic retinopathy, grading of diabetic macular edema severity is a topic of considerable complexity and subtlety, well deserving of its own analysis. Our study design included comparisons of monoscopic versus stereoscopic grading of diabetic macular edema, the results of which will be reported in another paper. In this study, we focused on grading diabetic retinopathy severity levels from digital monoscopic and stereo photographs compared with ETDRS stereo 35-mm slides (film).

## Material and Methods

### Study Design

We conducted a three-way comparison of ETDRS diabetic retinopathy severity levels graded from monoscopic and stereo digital images and from film. This study is one arm of a comprehensive evaluation of multiple digital photography formats compared with ETDRS photography: uncompressed color stereoscopic, compressed color stereoscopic, monoscopic, and monoscopic wide-angle mosaic. Additional papers will describe the other digital formats' effectiveness in evaluation of diabetic retinopathy compared with film.

### Fundus Photographs

Patients from The University of Texas Medical Branch (UTMB) Department of Ophthalmology and Visual Sciences Eye Clinic gave written consent for eye photography. Institutional review board approval was obtained and the protocol complied with the Declaration of Helsinki. Patients were informally screened in the clinic, and those with severity levels needed for the study were invited for photography. Study patients had a broad range of diabetic retinopathy but no other retinal vascular disease. Patients with media opacities preventing an adequate retina image were excluded from the study. Patients with previous retinal laser photocoagulation were also excluded, to avoid having the readers overgrade macular edema or proliferative retinopathy. Each study eye had 16 slides and 16 digital photographs taken per the ETDRS protocol: seven nonsimultaneous color stereo field pairs of the fundus and one pair of the anterior segment. Film and digital photographs were taken with the same mydriatic camera (model TRC-50EX/IX; Topcon Medical Systems, Paramus, NJ) at its 35° setting. Fundus camera optics were coupled with a 35-mm camera for one set of images and a digital camera (MegaVision, Santa Barbara, CA) for the other set. The mydriatic camera's digital sensor resolution was 2400 × 2000 pixels. Patients also underwent dilated photography with a nonmydriatic camera for the mosaic format. Patients rested a minimum of 30 minutes between photography format sessions. The same photographer, certified by the University of Wisconsin (UW) Fundus Photograph Reading Center for ETDRS protocol photography, took each patient's digital and film photographs in the same format sequence. Slide film was processed at a commercial facility (Kodak-certified Q-Lab; Eastman Kodak, Rochester, NY), as recommended by the UW Fundus Photograph Reading Center.[7]

A total of 152 eyes from an 85-patient cohort were selected to represent a stratified sample across the full range of diabetic retinopathy severity levels. Patients included 32 (37.6%) men, 53 (62.4%) women, 37 (43.5%) Caucasians, 24 (28.2%) Hispanics, and 24 (28.2%) African-Americans. Seventy-five right eye images and 77 left eye images were used. Photographs included both eyes of 67 (78.8%) patients, only right eyes of eight (9.4%) patients, and only left eyes of 10 (11.8%) patients. Patients ranged from 33 to 83 years of age, with a median of 60.5 years and a mean of 59.4 years. Photographs were assigned coded ID numbers, to remove identifying patient information.

Color and contrast of digital images were adjusted to conform to a formal color model based on standard ETDRS slides. Custom software generated red/green/blue (RGB) luminance histograms and adjusted each color channel curve to fit the model parameters. This method maximized the contrast of diabetic retinopathy abnormalities against retinal pigment epithelial (RPE) backgrounds without creating artifacts. The color-balancing algorithm was modified from one included in the UW Fundus Photograph Reading Center algorithm for reading Age-Related Eye Disease Study 2 (AREDS2) digital images.[16] Digital photography from the 152 stereo pair image sets were copied and processed to create the other digital formats. Monochromatic green channel images are thought to provide better contrast of small lesions against their background than does full color photography.[17] Stereo monochromatic green channel photographs were created by extracting the green channel from digital RGB color images. Green channel images were referred to as necessary when grading digital photographs to confirm suspected subtle diabetic retinopathy lesions. The left or right eye image from each stereo pair exhibiting the best focus and contrast was selected to create the color and green-channel monoscopic format.

### Grading Photographs

Three readers from the University of Wisconsin-Madison independently evaluated all format images from every eye. The images were reviewed in batches of the same format. The batch presentation order was counterbalanced by format to minimize bias. To minimize recall, evaluation was regulated by custom scheduling software that separated grading the same eye by at least 2 weeks. Readers graded severity based on abnormalities present in relevant photographic fields of the ETDRS classification system.[6,18] Findings were entered into a computer algorithm that calculated retinopathy severity on a nine-step ETDRS level scale: 10, no retinopathy; 15/20, microaneurysm (Ma) or retinal hemorrhage (RH) only; 35, mild nonproliferative diabetic retinopathy (NPDR); 43, moderate NPDR; 47, moderately severe NPDR; 53, severe NPDR; 61, mild proliferative retinopathy (PDR); 65, moderate PDR; 71/75, severe PDR; and 90, cannot grade.[18] Confidence levels were assigned to each study eye image format based on the reader's confidence in assessing retinopathy severity: *high*, high confidence, due to good image quality and typical lesions; *moderate*, adequate confidence due to less satisfactory image quality and/or atypical lesions; or *low*, inadequate confidence leading to the selection of "cannot grade." Readers also determined the presence and severity of the index lesions and diabetic macular edema (DME) according to ETDRS definitions.[6]

Index lesions were defined as abnormalities the presence and severity of which qualify an eye for a specific ETDRS level of severity.

Stereo sets of film were graded as described in ETDRS report 10.[6] Stereo pairs were viewed on daylight fluorescent light boxes with a Donaldson 5× stereo viewer (George Davco, Holcombe, MA). Overall magnification was approximately 12.5×, accounting for the combined magnification of the fundus camera and the viewer.

Customized software facilitated viewing digital images monoscopically at full screen with 13× magnification, approximating film review magnification. Digital images were displayed on 21-in. CRT displays at 1600 × 1200-pixel resolution at a viewing distance of approximately 26 in. Monitors were set at a color temperature of 6500°K and 2.2 $\gamma$, checked monthly with calibration hardware and software (Greytag Macbeth; X-Rite Inc.; Grand Rapids, MI). Digital images were displayed at 6.5× magnification when shown as stereo pairs and viewed with a hand-held stereo viewer (Screen-Vu; Eye Supply USA., Tampa, FL). During grading sessions, readers first reviewed a proof sheet of digital thumbnail images before examining each 35° field image in detail. Each field was reviewed monoscopically (13×) and stereoscopically (6.5×) in full color (and, when needed, as green channel images).

### Statistical Analysis

The severity of diabetic retinopathy in each eye was calculated as the central tendency (median grade) among the three independent graders, allowing pair-wise comparisons of all readers within each imaging format.

Stereo film grading results were considered the reference standard. The presence or absence of index lesions, severity level, and retinopathy severity at different thresholds were compared. We defined threshold as the cutoff for the presence of retinopathy at a particular ETDRS severity level or worse. Thresholds ranged from any retinopathy (level 15/20) to high-risk PDR (level 71/75).

Severity level agreement was cross-tabulated, and $\kappa$ statistics (unweighted and weighted [linear scheme]) were calculated. Eyes with photographs classified as ungradable (level 90) were excluded from the analysis. Guidelines for interpretation were based on Landis and Koch, as used in ETDRS report 10: 0.0 to 0.2, slight agreement; 0.21 to 0.40, fair agreement; 0.41 to 0.60, moderate agreement; 0.61 to 0.80, substantial agreement; and 0.81 to 1.00, almost perfect agreement.[19] Differences in frequency distributions (cross-tab marginal homogeneity) were checked for significance by using the McNemar test for dichotomized scales and the Bhapkar test for multistep scales. McNemar's test of overall bias was used to assess marginal distributions.

Severity threshold and index lesion grading were compared by using the $\kappa$ and receiver operator characteristic statistics, including sensitivity and specificity percentages, positive and negative predictive values, and disease prevalence percentages. Agreements between monoscopic and stereoscopic photographic grades, and within each format between readers, were assessed by percentage of agreement and $\kappa$.

Ungradable eyes were excluded from pair-wise interreader performance analyses within each format. Statistics comparing one imaging format to another were calculated by using the common subset of eyes gradable in both formats.

Statistical analyses were performed with several programs (SAS ver. 9.2, SAS, Inc. Cary, NC; Analyze-It, Analyze-It, Ltd., Leeds, UK; MedCalc, MedCalc Software Bvba, Mariakerke, Belgium; the MH [marginal homogeneity] Program ver. 1.2[20]).

## Results

### Monoscopic and Stereoscopic Agreement of Severity Levels

The distribution of diabetic retinopathy severity levels determined by film grading was level 10, 26 (17.1%) eyes; level 15/20 8 (5.3%) eyes; level 35, 34 (22.4%) eyes; level 43, 25 (16.4%) eyes; level 47, 24 (15.8%) eyes; level 53, 5 (3.3%) eyes; level 61, 10 (6.6%) eyes; level 65, 10 (6.6%) eyes; and level 71/75, 10 (6.6%) eyes (Table 1). The monoscopic digital photographs of three eyes were ungradable (Table 1).

There was substantial agreement between assessments of monoscopic digital and stereo film images ($\kappa$ = 0.65 [95% confidence interval (CI), 0.56–0.73]; linear weighted $\kappa$ = 0.87 [95% CI, 0.83–0.91]). Agreement was exact in 69.8% of the eyes: within one step in 96.0%, and within two steps in 98.7% (Table 2). ETDRS severity scale distribution from grading with monoscopic digital images was not significantly different from that with stereo film. According to the Bhapkar test of marginal homogeneity, the difference between the two was significant at $P$ = 0.53.

Grading agreement between monoscopic digital images versus stereo film was similar to stereo digital images versus stereo film (Li HK, et al., manuscript submitted 2010; Table 2).

Interreader agreement was better on the monoscopic digital images than that on either stereo format. Pair-wise interreader agreement on monoscopic photographs was similar to or better than that on stereo digital images and similar to or better than that on stereo film images (Table 3). Our intrareader $\kappa$ for monoscopic digital images is also similar to that for film (data not shown).

### Severity Threshold Agreement

There was good comparability of monoscopic digital photographs to stereo film in the assessment of the severity threshold (Table 4, Fig. 1). Sensitivity and specificity were at or above 90% in all divisions. The $\kappa$ agreement was 0.80 or above for all thresholds. Repeatability of grading severity threshold between readers was similar for monoscopic digital or stereo film images (Fig. 2). In both formats, agreement appeared somewhat better for any retinopathy, mild NPDR, moderate NPDR, and PDR thresholds than for moderately severe and severe NPDR thresholds.

### Index Lesion Agreement

The $\kappa$ statistic for identifying diabetic retinopathy lesions in monoscopic digital images compared to stereo film ranged from moderate to perfect (Table 5, Fig. 3). Grading venous beading (VB) had lower sensitivity and $\kappa$ than other intraretinal lesions (sensitivity = 81%, $\kappa$ = 0.71). The $\kappa$ and sensitivity were equally high when evaluating monoscopic photographs for IRMA and NVE (sensitivity of IRMA = 87%, $\kappa$ = 0.76; sensitivity of NVE = 81%, $\kappa$ = 0.77). NVD's and FPD's sensitivity and $\kappa$ statistics were lower than those of other lesions anterior to the retina (sensitivity of NVD = 33%, $\kappa$ = 0.43; sensitivity of FPD = 50%, $\kappa$ = 0.66). Grading with monoscopic images achieved high specificity and narrow specificity confidence intervals for all index lesions anterior to the retina.

A three-way comparison of monoscopic versus stereo imaging formats showed the performance of monoscopic grading of NVD to be lower than grading from stereo digital or stereo film photographs (Fig. 4). The comparison also showed almost equal agreement in grading NVE from monoscopic images. The prevalence of NVD and FPD in our study sample was low, as evidenced by wide $\kappa$ confidence intervals. There were only six eyes with NVD and two eyes with FPD.

### Monoscopic Digital and Stereoscopic Film Disagreement

There were six eyes with differences of two or more severity levels between monoscopic digital image and stereo film grading. Three eyes were assigned to a higher severity level when evaluated with monoscopic digital images than with stereo film. Three eyes were assigned to a higher severity level when

**TABLE 1.** ETDRS Diabetic Retinopathy Severity Level Assigned from Grading Monoscopic versus Stereoscopic Images

**A. Monoscopic Digital vs. Stereo Film**

| | Monoscopic Digital | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Stereo Film** | **10** | **15/20** | **35** | **43** | **47** | **53** | **61** | **65** | **71/75** | **90** | **Total** |
| 10 | 24 | 2 | | | | | | | | | 26 |
| 15/20 | 2 | 3 | 3 | | | | | | | | 8 |
| 35 | 1 | 2 | 25 | 5 | | | | | | 1 | 34 |
| 43 | | | 5 | 11 | 8 | | 1 | | | | 25 |
| 47 | | | | 4 | 14 | 2 | 2 | | | 2 | 24 |
| 53 | | | | | | 5 | | | | | 5 |
| 61 | | | | 1 | 1 | 1 | 5 | 2 | | | 10 |
| 65 | | | | | | | | 8 | 2 | | 10 |
| 71/75 | | | | | | | | 1 | 9 | | 10 |
| 90 | | | | | | | | | | | 0 |
| Total | 27 | 7 | 33 | 21 | 23 | 8 | 8 | 11 | 11 | 3 | 152 |

**B. Monoscopic Digital vs. Stereo Digital**

| | Monoscopic Digital | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Stereo Digital** | **10** | **15/20** | **35** | **43** | **47** | **53** | **61** | **65** | **71/75** | **90** | **Total** |
| 10 | 26 | 1 | | | | | | | | | 27 |
| 15/20 | 1 | 3 | 3 | | | | | | | | 7 |
| 35 | | 3 | 26 | 1 | | | | | | 1 | 31 |
| 43 | | | 2 | 11 | 8 | | | | | | 21 |
| 47 | | | 2 | 8 | 8 | 2 | 2 | | | 2 | 24 |
| 53 | | | | | 6 | 5 | | | | | 11 |
| 61 | | | | 1 | 1 | 1 | 6 | | | | 9 |
| 65 | | | | | | | | 10 | 1 | | 11 |
| 71/75 | | | | | | | | 1 | 10 | | 11 |
| 90 | | | | | | | | | | | 0 |
| Total | 27 | 7 | 33 | 21 | 23 | 8 | 8 | 11 | 11 | 3 | 152 |

Dark gray shading indicates perfect agreement; light gray shading indicates agreement within one step.

assessed with stereo film than with monoscopic digital images. An experienced reader not involved in this study reviewed our image sets side by side, to characterize the nature of the disagreements. Disagreement in five of the six eyes was due to uneven quality between the digital and film images. Views of abnormalities of interest were better in film in three eyes (one RH and one Ma in one eye; NVE in two eyes) but better with monoscopic digital photography in another two eyes (NVE). Grading differences in the remaining eye were due to reader misclassification of RPE abnormalities as secondary to laser photocoagulation.

NVD was missed in monoscopic digital images in four of the six eyes with NVD. A side-by-side review of the corresponding monoscopic digital and stereo film found that, in two eyes, the digital image view of the optic disc was oversaturated in the red channel. These images were too brightly illuminated to grade the definite presence of NVD vessels, even when using the green channel. In the other two eyes, the NVD missed in the one side of the stereo pair selected as the monoscopic view (based on overall quality) was detected in the other member of the stereo pair.

## DISCUSSION

Study results show substantial agreement between monoscopic digital photographs and stereo 35-mm film in assessing ETDRS severity level. Grading agreement using monoscopic digital photographs also compared well with stereo digital images. The results showed that grading agreement of stereo digital images was comparable to that with stereo film (Table 2). Given the declining use of film photography in ophthalmol-

**TABLE 2.** Diabetic Retinopathy Severity Level: Monoscopic Compared to Stereoscopic Images

| | **Monoscopic Digital vs. Stereo Film ($n$ = 149)*** | **Stereo Digital vs. Stereo Film† ($n$ = 152)*** | **Monoscopic Digital vs. Stereo Digital ($n$ = 149)*** |
|---|---|---|---|
| Complete Agreement | 69.8% | 67.8% | 70.7% |
| Agreement within 1 step | 96.0% | 96.1% | 96.0% |
| Agreement within 2 steps | 98.7% | 99.30% | 99.30% |
| $\kappa$ | 0.65 (95% CI, 0.56–0.73) | 0.62 (95% CI, 0.54–0.71) | 0.66 (95% CI, 0.57–0.74) |
| Weighted $\kappa$ | 0.87 (95% CI, 0.83–0.91) | 0.86 (95% CI, 0.82–0.90) | 0.87 (95% CI, 0.84–0.91) |

* Number represents the common subset of eyes gradable in both formats.
† Li HK, et al., manuscript submitted.

**TABLE 3.** Interreader Diabetic Retinopathy Severity Level Agreement by Format

| | Monoscopic Digital ($n$ = 149)* | Stereo Digital ($n$ = 152) | Stereo Film ($n$ = 152) | ETDRS Report 12 |
|---|---|---|---|---|
| Complete agreement | | | | 53% |
| Reader A vs. B | 63.1% | 56.6% | 50.0% | |
| Reader A vs. C | 61.7% | 53.3% | 53.3% | |
| Reader B vs. C | 74.5% | 63.2% | 63.2% | |
| Agreement within 1 step | | | | 88% |
| Reader A vs. B | 89.9% | 82.9% | 76.3% | |
| Reader A vs. C | 88.6% | 86.2% | 83.6% | |
| Reader B vs. C | 95.3% | 94.7% | 93.4% | |
| Agreement within 2 steps | | | | — |
| Reader A vs. B | 97.3% | 90.8% | 92.1% | |
| Reader A vs. C | 97.3% | 94.1% | 91.4% | |
| Reader B vs. C | 98.0% | 98.7% | 98.0% | |
| $\kappa$ | | | | 0.42 |
| Reader A vs. B | 0.58 (95% CI, 0.49–0.66) | 0.51 (95% CI, 0.42–0.59) | 0.43 (95% CI, 0.34–0.52) | |
| Reader A vs. C | 0.56 (95% CI, 0.47–0.65) | 0.47 (95% CI, 0.38–0.56) | 0.47 (95% CI, 0.38–0.56) | |
| Reader B vs. C | 0.70 (95% CI, 0.62–0.78) | 0.57 (95% CI, 0.48–0.66) | 0.57 (95% CI, 0.48–0.66) | |
| Weighted $\kappa$ | | | | (0.65)† |
| Reader A vs. B | 0.81 (95% CI, 0.76–0.87) | 0.73 (95% CI, 0.66–0.80) | 0.67 (95% CI, 0.59–0.75) | |
| Reader A vs. C | 0.80 (95% CI, 0.75–0.86) | 0.74 (95% CI, 0.67–0.81) | 0.71 (95% CI, 0.64–0.78) | |
| Reader B vs. C | 0.88 (95% CI, 0.84–0.92) | 0.82 (95% CI, 0.76–0.88) | 0.83 (95% CI, 0.78–0.87) | |

* Ungradable eyes are excluded.

† Different weighting scheme from linear scheme used in this study.

ogy, we did not compare monoscopic film with stereo film. However, our analyses showed high agreement between monoscopic digital images and both the film and digital stereo formats, implying that the grading agreement between monoscopic and stereo slide photography would be similar to our monoscopic versus stereo digital image results. Disagreements in results appear to reflect typical variability in the grading process and fundus photography illumination and focus, rather than intrinsic differences between stereo film, stereo digital, or monoscopic digital photography.

We also found reproducibility of classification of diabetic retinopathy severity in monoscopic or stereo digital images comparable to that with stereo film. Reproducibility compared favorably to that described in ETDRS report 12 (Table 3).

To the best of our knowledge, this is the only comparative study to apply the ETDRS diabetic retinopathy grading photography protocol and severity classification system to both monoscopic and stereo images. Except for replacing a 35-mm camera's film sensor with a digital chip, we minimized changes to the ETDRS protocol for monoscopic grading. Photographs were taken with the same camera optics, field magnification, area, and size. Digital photographs were balanced for color and contrast to closely match 35-mm fundus slides.[16] Our 2400 × 2000-pixel digital photography resolving power was approximately 13 $\mu$m on the retina, sufficient for detecting the smallest IRMA, NVE, or Ma seen on the film. As much as possible, we minimized differences between viewing stereo and monoscopic images. Viewing magnification was similar between stereo film (12.5×) and digital images (13×). For grading monoscopic digital images, each field was displayed full screen at 13×. When grading stereo digital images, each field was first reviewed monoscopically at full screen. Each stereo pair was reviewed at 6.5× to maximally fit both the left and right images to the monitor. Stereo viewing is not only influenced by how images are acquired, but also by how they are displayed.[21] We used stereo viewers that maintain color fidelity and image contrast, avoiding anaglyph (red/blue-green) glasses, polarized viewers, or interlaced liquid crystal shuttering systems. Display monitors were calibrated to ensure adequate and consistent brightness. Readers with a minimum of 10 years of ETDRS protocol grading experience reviewed all images.

Maintaining consistently high quality in stereo fundus photographs is difficult. Dark fundus pigmentation and patient flash tolerance are barriers to maximizing depth of field. Poor focus also limits stereopsis and identification of subtle neovascularization, IRMA, and fibrous proliferation. An experienced reader not involved in grading reviewed the stereo effect in 10% of digital and film image sets. Good stereo effect was found in both media. The stereo effect in sampled photographs was deemed typical for multicenter clinical trials or epidemio-

**TABLE 4.** Diabetic Retinopathy Severity Thresholds: Monoscopic Digital Images Compared with Stereo Film ($n$ = 149 eyes)

| Retinopathy Threshold | Sensitivity | Sensitivity 95% CI | Specificity | Specificity 95% CI | Eyes at or above Threshold ($n$) | Positive Predictive Value | Negative Predictive Value | Rate of Agreement | $\kappa$ | $\kappa$ 95% CI |
|---|---|---|---|---|---|---|---|---|---|---|
| ≥15/20 | 0.98 | 0.93–0.99 | 0.92 | 0.75–0.99 | 123 | 0.98 | 0.89 | 0.97 | 0.89 | 0.79–0.98 |
| ≥35 | 0.97 | 0.93–0.99 | 0.91 | 0.76–0.98 | 115 | 0.97 | 0.91 | 0.96 | 0.89 | 0.80–0.98 |
| ≥43 | 0.94 | 0.86–1.00 | 0.93 | 0.83–0.98 | 82 | 0.94 | 0.93 | 0.93 | 0.86 | 0.78–0.95 |
| ≥47 | 0.91 | 0.81–0.97 | 0.90 | 0.82–0.95 | 57 | 0.85 | 0.94 | 0.91 | 0.80 | 0.71–0.90 |
| ≥53 | 0.94 | 0.81–0.99 | 0.96 | 0.90–0.99 | 35 | 0.87 | 0.98 | 0.95 | 0.87 | 0.78–0.96 |
| ≥61 | 0.90 | 0.73–0.98 | 0.98 | 0.93–0.99 | 30 | 0.90 | 0.98 | 0.96 | 0.87 | 0.78–0.97 |
| ≥65 | 1.00 | 0.83–1.00 | 0.99 | 0.95–1.00 | 20 | 0.91 | 1.00 | 0.99 | 0.94 | 0.87–1.00 |
| 71/75 | 0.90 | 0.55–0.98 | 0.99 | 0.95–1.00 | 10 | 0.82 | 0.99 | 0.98 | 0.85 | 0.68–1.00 |

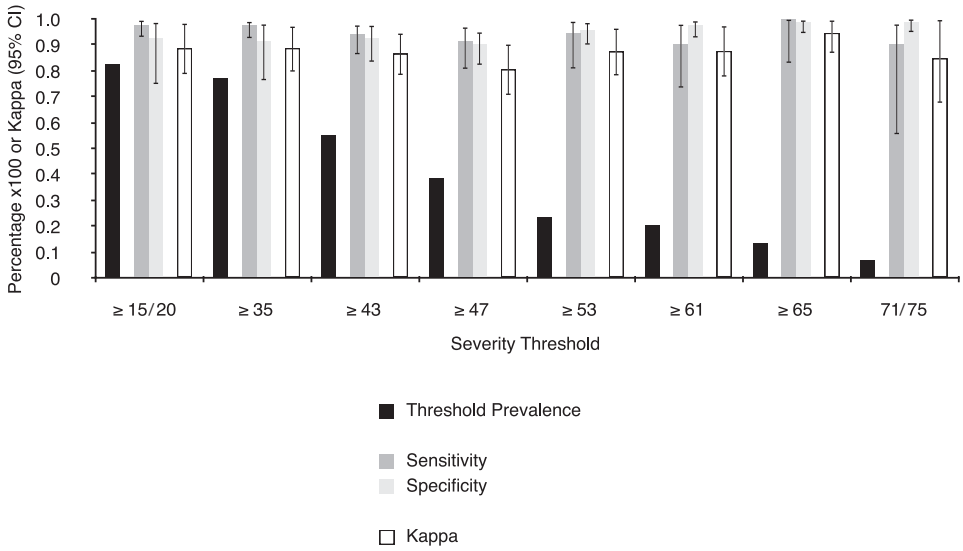


**FIGURE 1.** Diabetic retinopathy severity thresholds: monoscopic digital images compared with stereo film ($n$ = 149 eyes).

logic studies using certified photographers. Some eyes had better stereo effect than others, regardless of medium. Greater or lesser stereo effect seemed to be associated with individual eyes rather than medium.

Readers appeared to be more comfortable in assessing film than digital images, perhaps due to their many years of experience in grading from 35-mm slides. The graders rated their confidence as *high* in 41 (26.9%) of 152 eyes imaged on film, 16 (10.5%) of 152 eyes with stereo digital images, and 18 (12.1%) of 149 eyes with monoscopic digital images. However, lower levels of confidence did not lead to lower reproducibility among graders using digital stereo or monoscopic images. All graders agreed on the same severity level in 59 (38.8%) of 152 eyes using film, in 63 (41.4%) of 152 eyes using stereo digital images, and in 65 (43.6%) of 149 using monoscopic digital images.

Previous studies have compared monoscopic 35-mm slides[22,23] or monoscopic digital images[24–26] to the ETDRS stereo photography protocol. Moller et al.[22] compared 60°, monoscopic, single-field, 35-mm color slides to the ETDRS 30°, stereo, seven-field, 35-mm color slide protocol. Neovascularization falling within the field area of both protocols was missed in two eyes in grading the 60° images. Low magnification of 60° photography was believed to be the reason.[22] The authors suggested that a higher magnification would be better for achieving depth perception than are wider fields. EURODIAB compared a 45°, monoscopic, two-field photography protocol to the ETDRS reference standard.[27] They cited lack of stereopsis and/or low image magnification as reasons for some grading disagreement. In both studies, the reliability of monoscopic fundus photography was confounded by magnification differences between studied systems and the ETDRS protocol. Most studies have reported pooled ETDRS severity levels rather than the full scale. In some studies, the digital image resolutions were insufficient for detecting small diabetic retinopathy lesions. The lack of comparability across these protocols precludes a systematic literature review of monoscopic digital photography versus the ETDRS stereo film grading protocol.

The interpretation of the results of this study is not without caveats. We selected patients with relatively clear media, a variable that cannot be readily controlled in the course of typical research or clinical care. Some degree of lens opacity is

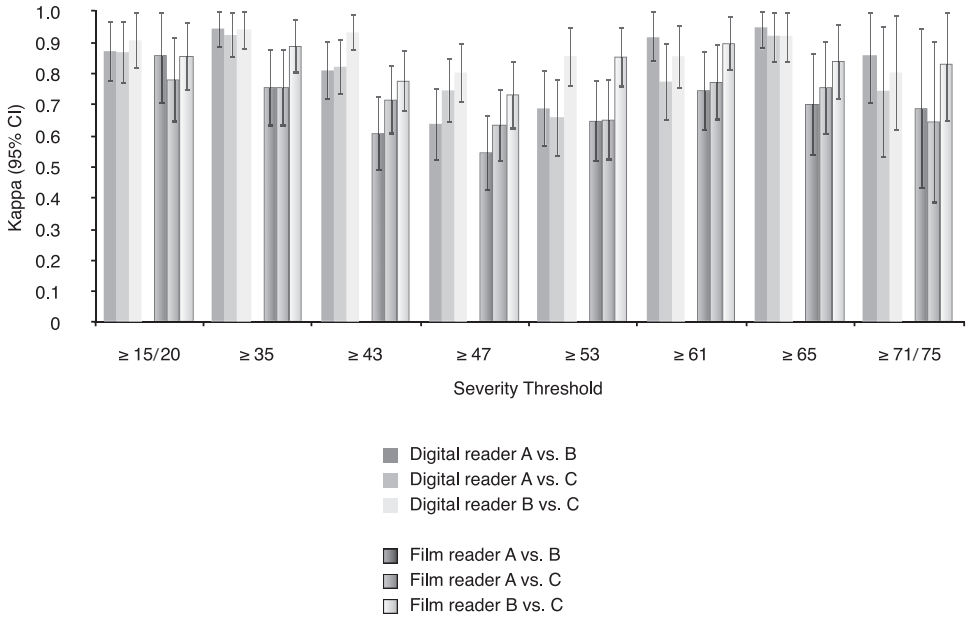


**FIGURE 2.** Interreader diabetic retinopathy severity threshold agreement: monoscopic digital images compared with stereo film (monoscopic digital, $n$ = 149 eyes; stereo film, $n$ = 152 eyes).

**TABLE 5.** Diabetic Retinopathy Severity Index Lesions: Monoscopic Digital Images Compared with Stereo Film ($n$ = 149 Eyes)

| Abnormality | Sensitivity | Sensitivity 95% CI | Specificity | Specificity 95% CI | Eyes with Abnormality ($n$) | Positive Predictive Value | Negative Predictive Value | Rate of Agreement | $\kappa$ | $\kappa$ 95% CI |
|---|---|---|---|---|---|---|---|---|---|---|
| Ma | 0.94 | 0.89–0.98 | 0.93 | 0.76–0.99 | 122 | 0.98 | 0.78 | 0.94 | 0.81 | 0.69–0.93 |
| RH | 0.98 | 0.94–1.00 | 0.93 | 0.76–0.99 | 121 | 0.98 | 0.93 | 0.97 | 0.91 | 0.83–1.00 |
| HE | 0.94 | 0.87–0.98 | 0.95 | 0.86–0.99 | 87 | 0.96 | 0.92 | 0.95 | 0.89 | 0.82–0.96 |
| SE | 0.87 | 0.77–0.94 | 0.89 | 0.79–0.95 | 77 | 0.89 | 0.86 | 0.88 | 0.76 | 0.65–0.86 |
| IRMA | 0.87 | 0.76–0.94 | 0.90 | 0.81–0.95 | 61 | 0.85 | 0.91 | 0.89 | 0.76 | 0.66–0.87 |
| VB | 0.81 | 0.54–0.96 | 0.96 | 0.91–0.98 | 16 | 0.68 | 0.98 | 0.94 | 0.71 | 0.53–0.89 |
| NVE | 0.81 | 0.61–0.93 | 0.96 | 0.91–0.99 | 26 | 0.81 | 0.96 | 0.93 | 0.77 | 0.63–0.90 |
| FPE | 0.84 | 0.60–0.96 | 0.99 | 0.96–0.99 | 19 | 0.94 | 0.98 | 0.97 | 0.87 | 0.75–1.00 |
| NVD | 0.33 | 0.05–0.77 | 0.99 | 0.96–1.00 | 6 | 0.67 | 0.97 | 0.97 | 0.43 | 0.02–0.84 |
| FPD | 0.50 | 0.08–0.92 | 1.00 | 0.98–1.00 | 2 | 1.00 | 0.99 | 0.99 | 0.66 | 0.04–1.00 |
| PRH | 1.00 | 0.63–1.00 | 1.00 | 0.97–1.00 | 8 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00–1.00 |
| VH | 1.00 | 0.40–1.00 | 1.00 | 0.97–1.00 | 4 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00–1.00 |

HE, hard exudates; SE, soft exudates;

common, as cataracts develop earlier in diabetic patients than the general population. Drug treatments may also have cataractogenic side effects. Our results, therefore, may not be applicable to general diabetic retinopathy patient populations.

It is widely believed that viewing stereo images is critical for distinguishing IRMA abnormalities from extraretinal NVE. Both abnormalities are important index lesions. The absence or presence of IRMA separates level 35 (moderate NPDR) from more severe levels, and the absence or presence of NVE distinguishes NPDR (<level 61) from PDR (≥level 61). Our results show reasonably good agreement between monoscopic digital images versus stereo film in detecting IRMA and NVE (Table 5). This agreement may be related to the readers' experience in recognizing diabetic retinopathy features and characteristic anatomic sites that are not dependent on stereopsis, such as IRMA and neovascularization, which are common in areas of vascular dropout.[1] Spoke-and-wheel networks,[28] saccular or fusiform dilated tips are also retinal neovascularization features.[6] Neovascularization in areas other than the disc is commonly located at arteriovenous crossings.[29] Large nets of neovascularization can cross over retinal vessels, a telltale sign not dependent on stereopsis[1] and is more frequently found on temporal than nasal veins.[29] The superotemporal quadrant is the most frequent initial site.[30] Most initial neovascularization is within 6 disc diameters of the optic disc. Three-fourths are in fields 1 (optic nerve), 4 (superotemporal), and 7 (inferonasal).[30] Opaque strands or sheets of fibrous proliferation are distinguishable from white, feathery, soft exudates.[6] Preretinal hemorrhages are commonly associated with a fluid level in addition to obscuring retinal vessels.[6]

We had a small sample of eyes with NVD and FPD. The eyes with NVD that was missed in monoscopic digital images were assigned appropriate retinopathy severity levels because there were co-existing proliferative index lesions.

Agreement between monoscopic digital images and stereo film grading proliferative eyes was high. We learned from additional review of NVD monoscopic data that most would be detected if the optic disc single image was in good focus, color balance, and proper exposure. The depth perception revealed in a stereo pair may have been helpful in only one NVD eye.

Monoscopic images in this study were not acquired as single images. Instead, we used the left or right image of a stereo pair that exhibited the best focus and contrast. The monoscopic view thus could never be better than its stereo view. This limitation accounted for three ungradable eyes among the monoscopic photographs. Suboptimal image quality of the monoscopic images also accounted for missed NVD. Some stereo views were probably better than their corresponding single image because human vision ignores artifacts in a single photograph when fusing a stereo pair. To achieve adequate stereo base, we took the first photograph of a stereo pair

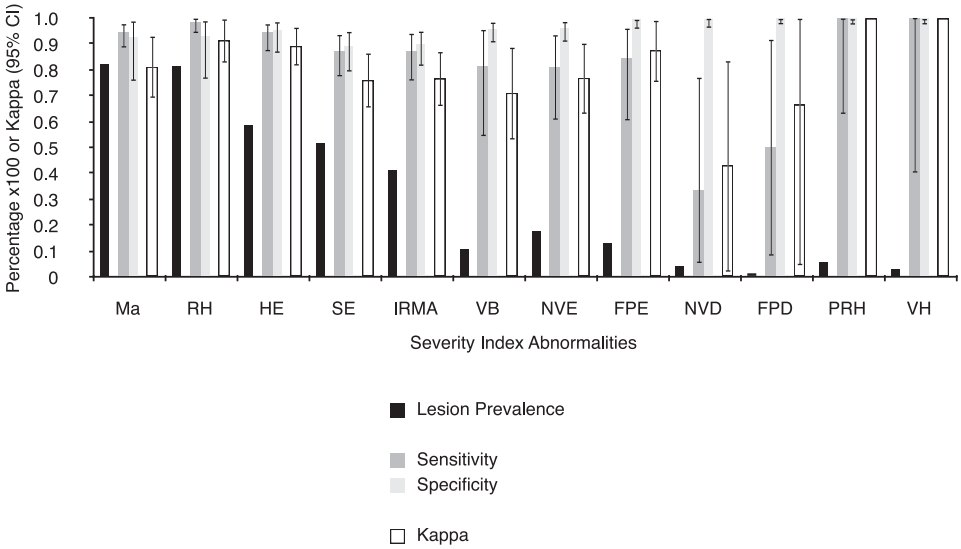


**FIGURE 3.** Diabetic retinopathy severity index lesions: monoscopic digital images compared with stereo film ($n$ = 149 eyes). PRH $\kappa$ 95% CI, 1.00–1.00; VH $\kappa$ 95% CI, 1.00–1.00. HE, hard exudates; SE, soft exudates.

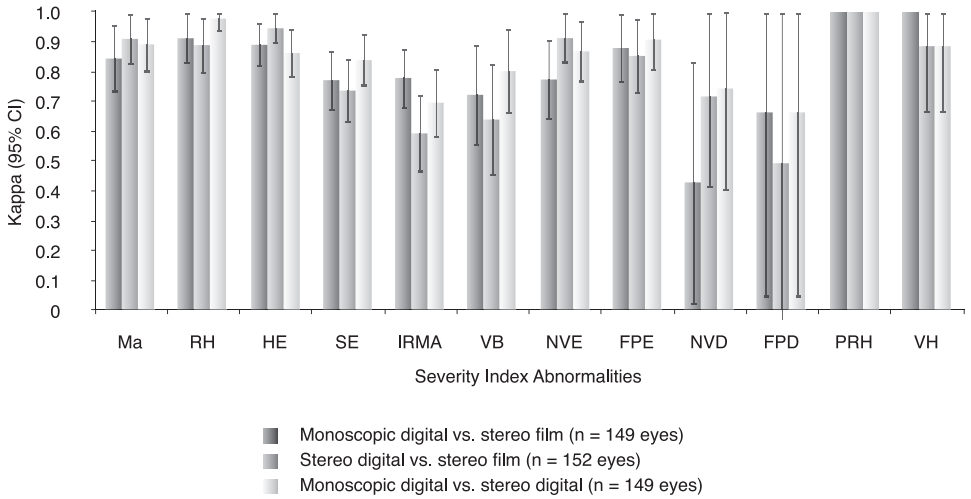


**FIGURE 4.** Diabetic retinopathy severity index lesions: three-way $\kappa$ comparison. PRH $\kappa$ 95% CI, 1.00–1.00 for all format comparisons; VH $\kappa$ 95% CI, 1.00–1.00 for monoscopic digital versus stereo film. HE, hard exudates; SE, soft exudates.

as far to one side of a pupil as possible without introducing shading or edge artifacts. It is likely that our monoscopic image quality would have been better without this requirement. The combined information from both images of a stereo pair compensated for the suboptimal quality of the constituent single image due to the readers' cognitively combining good and poor images.

There are trade-offs in either monoscopic or stereo photography. Sequential stereo imaging creates depth perception but constrains photography to oblique views near the left and right sides of a dilated pupil. This limitation can compromise image clarity in both members of a stereo pair in eyes with suboptimal dilation or substantial peripheral media opacities. Monoscopic imaging allows a straight shot of the fundus, particularly useful in eyes with restricted pupillary dilation or media opacities. However, there is no stereo effect and the reader does not have the insurance of two images to choose from. In their study of digital imaging, Rudnisky et al.[31] chose stereo photography only for views of the disc and macula (ETDRS fields 1 and 2). They used monoscopic photography of the other five (peripheral) fields. This method allowed the investigators some reduction in time, effort, and patient discomfort due to the need for fewer images, while retaining the benefit of the stereo effect for the most critical retinal sites.

Airlie House believed stereo fundus photography would facilitate qualitative and quantitative comparisons of diabetic vascular abnormalities in collaborative studies. They advocated its use whenever possible.[1] The ETDRS stereo protocol was designed to be rigorous, so that even small progressions of diabetic retinopathy could be distinguished in the large number of patients seen in multicenter clinical trials.[27] However, the benefit of grading from stereo versus single photographs was never fully tested. We found that grading a broad range of diabetic retinopathy severity levels by using monoscopic digital images produced results equivalent to using stereo digital images or stereo 35-mm slides. These results suggest that a stereo effect may not be critical for accurate classification of ETDRS diabetic retinopathy severity when using current technology and an optimized framework for fundus photography acquisition and reviewing. If so, the added cost and burden of stereo photography may not be justified.

## *Acknowledgments*

The authors thank the staff from the Department of Ophthalmology and Visual Sciences, University of Wisconsin-Madison, including the Ocular Epidemiology Group of Barbara E. K. Klein, MD, and Ronald Klein, MD: Andrew F. Ewen, Anne E. Mosher, and Maria K. Swift for grading diabetic retinopathy images; Stacy M. Meuer for grading supervision and Daniel P. Murach for computer support; the Fundus Photograph Reading Center directed by Ronald P. Danis, MD: Trina M. Harding for grading orientation, Qian Peng for statistical advice, Jeff T. Klaves for statistical analyses, and Matthew D. Davis, MD, for suggestions regarding the manuscript.

## *References*

1. Davis MD, Norton WD, Myers FL. Airlie classification of diabetic retinopathy: In: Goldberg MF, Fine SL, eds. *Symposium on the treatment of diabetic retinopathy.* Publication no. 1890. Arlington, VA: US Department of Health, Education and Welfare; 1968: 7–22.
2. Allen L. Ocular fundus photography: suggestions for achieving consistently good pictures and instructions for stereoscopic photography. *Am J Ophthalmol.* 1964;57:13–28.
3. Diabetic Retinopathy Research Study Group. A modification of the Airlie House classification of diabetic retinopathy. DRS report number 6. *Invest Ophthalmol Vis Sci.* 1981;21:210–216.
4. *Diabetic Retinopathy Study: Manual of Operations*. Baltimore: Diabetic Retinopathy Study Coordinating Center; 1972.
5. Diabetic Retinopathy Research Study Group: Design, methods, and baseline results. DRS report number 6. *Invest Ophthalmol Vis Sci.* 1981;21:149–209.
6. 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 10. *Ophthalmology*. 1991;98:786–806.
7. Diabetic Retinopathy Clinical Research Network. *DRCRnet Manual of Operations.* Ver 3.0. Tampa, FL: Jaeb Center for Health Research Foundation; April 8, 2005.
8. Chew EY, Ambrosius WT, Howard LT, et al. Rationale, design, and methods of the Action to Control Cardiovascular Risk in Diabetes Eye Study (ACCORD-EYE). *Am J Cardiol.* 2007;99:103i–111i.
9. Epidemiology of Diabetes Interventions and Complications (EDIC). Design, implementation, and preliminary results of a long-term follow-up of the Diabetes Control and Complications Trial cohort. *Diabetes Care.* 1999;22:99–111.
10. The Diabetes Control and Complications Trial (DCCT). Design and methodologic considerations for the feasibility phase. The DCCT Research Group. *Diabetes.* 1986;35:530–545.
11. Modified 7-standard field color fundus photography (7M-F) and film fluorescein angiography (FA-F). Madison, WI: University of Wisconsin-Madison Fundus Photograph Reading Center; 2009. Available at: http://eyephoto.ophth.wisc.edu/Photography/Protocols/Mod7&FA-ver1.pdf. Accessed: June 21, 2009.
12. Davis MD, Hubbard LD, Trautman J, Klein R. Conference on insulin pump therapy in diabetes: multicenter study effect on microvascular disease—studies of retinopathy—methodology for

assessment and classification with fundus photographs. *Diabetes.* 1985;34(suppl)3:42–49.
13. Rudnisky CJ, Tennant MT, de Leon AR, et al. Benefits of stereopsis when identifying clinically significant macular edema via teleophthalmology. *Can J Ophthalmol.* 2006;41:727–732.
14. Younis N, Broadbent DM, Vora JP, Harding SP. Incidence of sight-threatening retinopathy in patients with type 2 diabetes in the Liverpool Diabetic Eye Study: a cohort study. *Lancet.* 2003;361:195–200.
15. Kohner EM, Aldington SJ, Stratton IM, et al. United Kingdom Prospective Diabetes Study, 30: diabetic retinopathy at diagnosis of non-insulin-dependent diabetes mellitus and associated risk factors. *Arch Ophthalmol.* 1998;116:297–303.
16. Hubbard LD, Danis RP, Neider MW, et al. Brightness, contrast, and color balance of digital versus film retinal images in the age-related eye disease study 2. *Invest Ophthalmol Vis Sci.* 2008;49:3269–3282.
17. Yassur Y, Siegel R, Karp M, Laron Z, Topper E. The detection of early microangiopathy in juvenile diabetics by monochromatic light photography. *Pediatr Adolesc Endocrinol.* 1980;9:181–184.
18. Fundus photographic risk factors for progression of diabetic retinopathy. ETDRS report number 12. Early Treatment Diabetic Retinopathy Study Research Group. *Ophthalmology.* 1991;98:823–833.
19. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–174.
20. Uebersax JS. User guide for the MH program (vers. 1.2). Statistical Methods for Rater Agreement; 2006. Available at: http://ourworld.compuserve.com/homepages/jsuebersax/mh.htm. Accessed: May 1, 2009.
21. Tyleer ME. Stereo fundus photography: principles and techniques. In: Saine PJ, Tyler ME. eds. *Ophthalmic Photography: Retinal Photography, Angiography, and Electronic Imaging.* 2nd ed. Boston: Butterworth-Heinemann; 2002:118–135.
22. Moller F, Hansen M, Sjolie AK. Is one 60 degree fundus photograph sufficient for screening of proliferative diabetic retinopathy? *Diabetes Care.* 2001;24:2083–2085.
23. Atan D, Foy C, Scanlon PH. Reply to: Evaluation of the effect of JPEG and JPEG2000 image compression on the detection of diabetic retinopathy (author reply 3 to a letter). *Eye.* 2008;22:471;.
24. Scanlon PH, Malhotra R, Greenwood RH, et al. Comparison of two reference standards in validating two field mydriatic digital photography as a method of screening for diabetic retinopathy. *Br J Ophthalmol.* 2003;87:1258–1263.
25. Lawrence MG. The accuracy of digital-video retinal imaging to screen for diabetic retinopathy: an analysis of two digital-video retinal imaging systems using standard stereoscopic seven-field photography and dilated clinical examination as reference standards. *Trans Am Ophthalmol Soc.* 2004;102:321–340.
26. Massin P, Erginay A, Ben Mehidi A, et al. Evaluation of a new non-mydriatic digital camera for detection of diabetic retinopathy. *Diabet Med.* 2003;20:635–641.
27. Aldington SJ, Kohner EM, Meuer S, et al. Methodology for retinal photography and assessment of diabetic retinopathy: the EURODIAB IDDM complications study. *Diabetologia.* 1995;38:437–444.
28. Shah KB, Han DP. Proliferative diabetic retinopathy. *Int Ophthalmol Clin.* 2004;44:69–84.
29. Taylor E, Dobree JH. Proliferative diabetic retinopathy: site and size of initial lesions. *Br J Ophthalmol.* 1970;54:11–18.
30. Feman SS, Leonard-Martin TC, Semchyshyn TM. The topographic distribution of the first sites of diabetic retinal neovascularization. *Am J Ophthalmol.* 1998;125:704–706.
31. Rudnisky CJ, Tennant MT, Weis E, et al. Web-based grading of compressed stereoscopic digital photography versus standard slide film photography for the diagnosis of diabetic retinopathy. *Ophthalmology.* 2007;114:1748–1754.