

Published in final edited form as:

Retina. 2008 January ; 28(1): 36–40.

## Agreement Between Clinician and Reading Center Gratings of Diabetic Retinopathy Severity Level at Baseline in a Phase 2 Study of Intravitreal Bevacizumab for Diabetic Macular Edema

Ingrid U. Scott, Neil M. Bressler, Susan B. Bressler, David J. Browning, Clement K. Chan, Ronald P. Danis, Matthew D. Davis, Craig Kollman, Haijing Qin, and the Diabetic Retinopathy Clinical Research Network Study Group

### Abstract

**Summary Statement**—Comparison of clinician and reading center assessments of diabetic retinopathy severity demonstrated moderate agreement on a 4-step scale that pooled a majority of cases in a single broad step of the scale.

**Purpose**—To evaluate agreement in diabetic retinopathy severity classification by retina specialists performing ophthalmoscopy versus Reading Center (RC) grading of 7-field stereoscopic fundus photographs in a phase 2 clinical trial of intravitreal bevacizumab for center-involved diabetic macular edema.

**Methods**—Clinicians' grading scale used 4 levels: microaneurysms only, mild/moderate non-proliferative diabetic retinopathy (NPDR), severe NPDR, and proliferative diabetic retinopathy (PDR) or prior panretinal photocoagulation (PRP) or both. The Reading Center scale used 8 levels: microaneurysms only, mild NPDR, moderate NPDR, moderately severe NPDR, severe NPDR, mild PDR, moderate PDR, and high-risk PDR. Percent agreement and kappa statistic were defined by collapsing RC categories to match those used by clinicians.

**Results**—There was agreement in 89/118 eyes (75%) with kappa=0.55 (95% confidence interval: [0.41, 0.68]). In 6 eyes, disagreements were of potential substantial clinical importance: 5 eyes with subtle retinal neovascularization and 1 with a small preretinal hemorrhage identified only in photographs.

**Conclusions**—Clinician grading of retinopathy severity had moderate agreement with Reading Center grading and might be useful for placing eyes into broad baseline categories.

### Keywords

Diabetic retinopathy level; Diabetic retinopathy severity; fundus photograph grading; non-proliferative diabetic retinopathy; ophthalmoscopic grading; proliferative diabetic retinopathy

---

Corresponding author: Ingrid U. Scott, MD, MPH, c/o Jaeb Center for Health Research, 15310 Amberly Drive, Suite 350, Tampa, FL 33647; Phone: (813) 975-8690, Fax: (800) 816-7601, E-mail: drcrnet@jaeb.org.

DRCRnet investigator financial disclosures are posted at [www.drcr.net](http://www.drcr.net)

There are no conflicts of interest.

An address for reprints will not be provided.

\*The most recently published list of the Diabetic Retinopathy Clinical Research Network investigators and staff participating in this protocol can be found in *Ophthalmology* 2007;114:1860-1867 with a current list available at [www.drcr.net](http://www.drcr.net) Supported through a cooperative agreement from the National Eye Institute EY14231, EY14269, EY14229 and the Juvenile Diabetes Research Foundation, International (New York, NY).

## INTRODUCTION

Grading severity of diabetic retinopathy from 7-field stereoscopic fundus photographs is the accepted standard method to evaluate baseline levels and progression of diabetic retinopathy in clinical trials and epidemiologic studies.<sup>1–6</sup> There have been several publications reporting the sensitivity and specificity of diabetic retinopathy detection and severity classification as determined by clinical examination (ophthalmoscopy) compared with a reference standard of stereoscopic fundus photographs. However, in many of these studies, ophthalmoscopy was performed by a single ophthalmologist<sup>7–9</sup> or other health care professionals,<sup>10</sup> such as optometrists, ophthalmic technicians,<sup>9</sup> diabetologists,<sup>11</sup> physician assistants,<sup>12</sup> and nurses.<sup>11</sup> In studies which compared ophthalmoscopy to grading of 7-field stereoscopic fundus photographs graded at a Reading Center, sensitivity of diabetic retinopathy detection has been reported to range between 32 and 82%, and specificity between 95 and 100%.<sup>7, 9, 12</sup> In the current study with 47 participating retinal specialists in a phase 2 clinical trial of intravitreal bevacizumab (Avastin®, Genentech, South San Francisco, California) for center-involved diabetic macular edema (DME), we evaluated the agreement in diabetic retinopathy severity classification as determined by ophthalmoscopy performed by retina specialists using a clinicians' grading scale as compared with grading of stereoscopic fundus photographs at a Reading Center using a grading scale modified slightly from the ETDRS.<sup>13</sup> Results of this study provide information regarding how ophthalmologists in the Diabetic Retinopathy Clinical Research Network (DRCR.net) categorize levels of retinopathy of subjects compared with detailed grading of fundus photographs of those same subjects by a Reading Center, and may help determine if and when clinician assessment alone may be sufficient for certain study purposes.

## METHODS

The design, methods, and results of the phase 2 intravitreal bevacizumab for DME trial, a multi-center randomized clinical trial conducted by the DRCR.net at 36 clinical sites in the United States, have been published.<sup>14</sup> The study enrolled 121 subjects at 36 different sites by 47 different retina specialists. As the current study relates only to baseline data from this trial, only the pertinent eligibility criteria and aspects of the study design relevant to the current study are summarized here. The entire protocol is available on the public site at <http://drcr.net>. Eligible subjects were at least 18 years old with type 1 or type 2 diabetes. The pertinent eligibility criteria for the study eye included: (1) definite retinal thickening due to DME involving the center of the macula based on clinical exam, (2) OCT central subfield thickness  $\geq 275$  microns, (3) no history of treatment for DME within the prior 3 months and (4) no panretinal photocoagulation (PRP) within the prior 4 months or an anticipated need for PRP in the 6 months following randomization. A subject could contribute only one study eye.

Standard ETDRS 7-field film color stereoscopic fundus photographs were obtained at baseline only of the study eye of each study participant and forwarded to the DRCR.net Reading Center at the University of Wisconsin-Madison for grading by trained readers who were masked to all clinical data, including clinician assessment of diabetic retinopathy level. Occasionally fundus photographs of both eyes were submitted; photographs of non-study eyes were not available to the graders. Photographs were each graded independently by each of 2 of the 6 graders assigned to DRCR.net color photograph grading. Disagreements by one or more retinopathy severity levels were resolved by a senior grader. The graders used standardized definitions for various characteristics of diabetic retinopathy and standard photographs to assess the severity of these characteristics in the 7 photographic fields and derived a composite ETDRS level score of 20, 35, 43, 47, 53, 60–61, 65, 71 or 75.<sup>13, 15</sup> These levels correspond to the following classifications for diabetic retinopathy severity: 1) microaneurysms only, 2) mild non-proliferative diabetic retinopathy (NPDR), 3) moderate NPDR, 4) moderately severe

NPDR, 5) severe NPDR, 6) mild proliferative diabetic retinopathy (PDR), 7) moderate PDR, and 8) high-risk PDR.

Clinicians were asked to grade diabetic retinopathy severity level at baseline in each eye of all study participants as one of the following levels: a) microaneurysms only, b) mild/moderate NPDR, c) severe NPDR (4-2-1 rule), and d) PDR or prior PRP or both. For eyes in category (d) presence/extent of PRP scars and of new vessels/vitreous or preretinal hemorrhage was also recorded. With the exception of required mydriasis, no specific protocol for the examinations was provided, but the principal methods used for ophthalmoscopy were slit-lamp biomicroscopy with a hand-held lens and binocular indirect ophthalmoscopy. Except for the 4-2-1 rule defining severe NPDR (severe intraretinal hemorrhages and microaneurysms in at least 4 quadrants, definite venous beading in at least 2 quadrants, or moderate intraretinal microvascular abnormalities in at least 1 quadrant), no specific definitions were provided to the clinicians for the severity levels with the presumption that their general ophthalmic fund of knowledge would suffice to differentiate these categories.

For comparison with clinician grading, Reading Center categories were collapsed to match those used by the clinicians as determined by the authors; Reading Center categories for mild, moderate and moderately severe NPDR were combined as were categories for mild, moderate and high-risk PDR. The clinician versus Reading Center grading comparison was not planned until follow-up on all patients had been completed. The current study is comprised of a retrospective analysis of data available and collected prospectively within the phase 2 study protocol.

Photographs were not available for 3 subjects, leaving 118 subjects for analysis. All photographs for which there was any disagreement in the gradings by the clinician compared with the Reading Center were re-reviewed by one of us (MDD) in order to ascertain potential reasons for the disagreements. Of the 118 eyes, 4 were initially classified as non-gradable by the Reading Center because of poor photographic quality, but on review for this manuscript could be assigned a retinopathy severity level and were included in the analysis (reviewer was aware of clinician's grading at the time of re-review).

Percent agreement and the unweighted kappa statistic were computed, comparing the clinician and Reading Center gradings.

## RESULTS

Results of clinician and Reading Center diabetic retinopathy severity level gradings are summarized in Table 1. The kappa value was 0.55 (95% confidence interval [0.41, 0.68]). The shaded cells represent consistent gradings (89 of 118 eyes; 75%) and highlighted cells represent agreement within one step (101 eyes; 86%).

Of the 12 one-step disagreements, 9 were between severe NPDR on one scale and the next lower step on the other scale, and one was between severe NPDR and mild PDR (Table 1). The remaining two one-step disagreements occurred in eyes with occasional microaneurysms and small hemorrhages; the hemorrhages were overlooked clinically or categorized as microaneurysms.

Of the 17 disagreements by 2 or more steps, 4 involved small differences in severity of NPDR (2 classified as microaneurysms only clinically and in the mildest part of the moderately severe NPDR category [level 47A] by Reading Center grading and 2 classified as severe NPDR clinically and moderate NPDR [level 43] by Reading Center grading). In 12 disagreements, NPDR was recorded clinically and PDR by Reading Center grading. In 6 of these 12 cases, scars of prior PRP, which were present in the photographs, had apparently been ignored in

answering the retinopathy severity question clinically (new vessels, vitreous or preretinal hemorrhage or PRP were all part of the highest step on the clinical scale); other than PRP scars, these eyes were not noted to have other features that would have put them in the PDR category. The remaining 6 of these 12 disagreements comprised 2 eyes with small patches of neovascularization elsewhere (NVE), 2 with very subtle neovascularization at the disc (NVD), one with a small preretinal hemorrhage and one with definite NVE but only questionable NVD that had been overgraded as definite. Finally, there was one eye classified as PDR clinically and as moderately severe NPDR by Reading Center grading in which neither new vessels, nor vitreous or preretinal hemorrhage or PRP was visible in the photographs; presumably, one or more of these abnormalities was observed clinically outside of the 7 photographic fields.

There was agreement on presence of prior PRP scars in all of the 23 eyes in which the clinical and photographic assessments agreed on presence of PDR. In 5 of these eyes, residual new vessels or vitreous/preretinal hemorrhage were graded present only in the photographs and in 4 noted to be present only clinically. On review of the first 5, subtle NVE was present in 3, subtle NVD in one, and a small preretinal hemorrhage in the fifth. Among the latter 4 eyes, there was one in which a small amount of old white vitreous hemorrhage was not reported by the graders and 3 in which no PDR characteristics were seen in the photographs (and presumably were present only outside of the 7 photographic fields).

## DISCUSSION

We found 86% agreement within one step and a kappa value of 0.55 (95% confidence interval [0.41, 0.68]) between clinical assessments of diabetic retinopathy severity level performed by retina specialists and fundus photograph gradings performed at a Reading Center evaluating one eye of each of 118 individuals participating in a clinical trial of center-involved DME. Clinician assessments were made on a 4-step retinopathy severity scale without standardization of methods and compared with the ETDRS retinopathy severity scale used by trained graders at a Reading Center. These results are generally consistent with results of other studies comparing ophthalmoscopy to grading of 7-field stereoscopic fundus photographs at a Reading Center. Kinyoun et al<sup>7</sup> reported an agreement rate of 86% and 86% (weighted kappa, 0.56 and 0.62) in the grading of diabetic retinopathy in right and left eyes, respectively, with both the retina specialist and the Reading Center using the same scale of diabetic retinopathy level. In 5 of the 8 disagreements, the clinical assessment resulted in a lower diabetic retinopathy severity level classification than the Reading Center grading. Moss et al<sup>9</sup> reported an agreement rate of 86% (kappa, 0.75) for grading retinopathy (using 3 levels: none, nonproliferative, or proliferative) when comparing ophthalmoscopy performed by an ophthalmologist (the principal investigator of the study), an optometrist or an ophthalmic technician (both of the latter had undergone extensive training in ophthalmoscopy and fundus photography) to grading of photographs at a Reading Center. Ophthalmoscopy was more likely to disagree with fundus photography grading in eyes with less severe forms of retinopathy. Among 301 eyes classified by Reading Center grading as microaneurysms only, 50% were classified as having no retinopathy by ophthalmoscopy. Perhaps more importantly, in 35 of 170 eyes (21%) with PDR by Reading Center grading, NPDR (or, in 3 of these eyes, no retinopathy) was found by ophthalmoscopy. Further review of these photographs by the ophthalmologist found PDR present in 34 of the 35 disagreements, but high-risk PDR in only 1 of 34 (representing 1 of the 36 eyes with high-risk PDR by Reading Center grading).

The principal causes of disagreements in our study appeared to be the greater sensitivity of photograph grading to subtle abnormalities and the larger area of retina available to clinical examination. These causes of disagreement have been recognized previously.<sup>9</sup> Although prior PRP was listed in the written definition of PDR provided to clinicians participating in this clinical trial, in 6 of the 12 disagreements in which NPDR was recorded clinically and PDR

by Reading Center grading, scars of prior PRP had apparently been ignored in answering the retinopathy severity question. Other investigators have also found that prior photocoagulation causes confusion in how to classify eyes for diabetic retinopathy.<sup>16</sup> These findings suggest that it may be helpful to emphasize such grading criteria to clinicians in future studies.

In 6 eyes, disagreements were considered to be of potential substantial clinical importance if the patients had not been under regular ophthalmologic follow-up. Among the 6, there were 5 eyes with very subtle new vessels and one with a small preretinal hemorrhage identified only in the photographs.

Limitations of this analysis include the restricted range of diabetic retinopathy included in our sample (center-involved DME was required and anticipated need for PRP in the 6 months following enrollment was an exclusion), the pooling of a broad portion of the NPDR severity range into one category, and the relatively small number of study participants. In addition, no specific manual of definitions or direction towards certain criteria for evaluation was provided to the clinicians; it is unknown whether provision to clinicians of a defined set of criteria for evaluation would have affected the agreement between clinician and Reading Center evaluations.

In summary, comparison of clinical and photographic assessments of retinopathy severity showed moderate agreement on a 4-step scale that pooled a majority of cases in a single broad step of the scale. While this 4-step scale was deemed sufficient for the needs of a trial managing DME—that is, to classify baseline severity in order to check for balance among study groups and any treatment interaction of the diabetic retinopathy severity level on outcomes—it may be insufficient for determining precise retinopathy severity levels and for trials in which the presence of early PDR is an eligibility or outcome variable.

## References

1. Early Treatment Diabetic Retinopathy Study Research Group. Grading diabetic retinopathy from stereoscopic color fundus photographs - an extension of the modified Airlie House classification. ETDRS report number 10. *Ophthalmology* 1991;98:786–806. [PubMed: 2062513]
2. Group DCaCTR. Progression of retinopathy with intensive versus conventional treatment in the Diabetes Control and Complications Trial. *Ophthalmology* 1995;102:647–61. [PubMed: 7724182]
3. Klein R, Klein BE, Moss SE, Cruickshanks KJ. The Wisconsin Epidemiologic Study of Diabetic Retinopathy: XVII. The 14-year incidence and progression of diabetic retinopathy and associated risk factors in type 1 diabetes. *Ophthalmology* 1998;105:1801–15. [PubMed: 9787347]
4. Mitchell P, Smith W, Wang JJ, Attebo K. Prevalence of diabetic retinopathy in an older community. The Blue Mountains Eye Study. *Ophthalmology* 1998;105:406–11. [PubMed: 9499768]
5. Varma R, Paz SH, Azen SP, et al. The Los Angeles Latino Eye Study: design, methods, and baseline data. *Ophthalmology* 2004;111:1121–31. [PubMed: 15177962]
6. The PKC-DRS Study Group. The effect of ruboxistaurin on visual loss in patients with moderately severe to very severe nonproliferative diabetic retinopathy: initial results of the Protein Kinase C beta Inhibitor Diabetic Retinopathy Study (PKC-DRS) multicenter randomized clinical trial. *Diabetes* 2005;54:2188–97. [PubMed: 15983221]
7. Kinyoun JL, Martin DC, Fujimoto WY, Leonetti DL. Ophthalmoscopy versus fundus photographs for detecting and grading diabetic retinopathy. *Invest ophthalmol Vis Sci* 1992;33:1888–93. [PubMed: 1582794]
8. Schachat AP, Hyman L, Leske MC, et al. Comparison of diabetic retinopathy detection by clinical examinations and photograph gradings. Barbados (West Indies) Eye Study Group. *Arch Ophthalmology* 1993;111:1064–70.
9. Moss SE, Klein R, Kessler SD, Richie KA. Comparison between ophthalmoscopy and fundus photography in determining severity of diabetic retinopathy. *Ophthalmology* 1985;92:62–7. [PubMed: 2579361]

10. Hutchinson A, McIntosh A, Peters J, et al. Effectiveness of screening and monitoring tests for diabetic retinopathy--a systematic review. *Diabet Med* 2000;17:495–506. [PubMed: 10972578]
11. Forrest RD, Jackson CA, Yudkin JS. Screening for diabetic retinopathy--comparison of a nurse and a doctor with retinal photography. *Diabetes Res* 1987;5:39–42. [PubMed: 3621804]
12. Pugh JA, Jacobson JM, Van Heuven WA, et al. Screening for diabetic retinopathy. The wide-angle retinal camera. *Diabetes Care* 1993;16:889–95. [PubMed: 8100761]
13. Early Treatment Diabetic Retinopathy Study Research Group. Fundus photographic risk factors for progression of diabetic retinopathy. ETDRS report number 12. *Ophthalmology* 1991;98:823–33. [PubMed: 2062515]
14. Network DRCR. A phase 2 randomized clinical trial of intravitreal bevacizumab for diabetic macular edema. *Ophthalmology*. In Press
15. Group ETDRSR. Fluorescein angiographic risk factors for progression of diabetic retinopathy. ETDRS report number 13. *Ophthalmology* 1991;98:834–40. [PubMed: 2062516]
16. Nathan DM, Fogel HA, Godine JE, et al. Role of diabetologist in evaluating diabetic retinopathy. *Diabetes Care* 1991;14:26–33. [PubMed: 1991432]

Table 1

Reading Center vs Clinician Classification of Retinopathy Severity (N=118)<sup>a</sup>

Reading Center Classification	Clinician Classification				Overall
	Microaneurysms only	Mild/Moderate/ Moderately Severe NPDR	Severe NPDR	PDR and/or prior PRP	
Microaneurysms only					
Mild NPDR	0	0	0	0	0
Moderate NPDR	2	12	0	0	14
Moderately severe NPDR	0	16	2	0	18
Severe NPDR	2	32	5	1	40
Mild PDR	0	4	6	0	10
Moderate PDR	1	6	1	19	27
High Risk PDR	0	4	1	2	7
Overall	5	74	15	24	118

<sup>a</sup>Photos were not available for 3 subjects.

<sup>b</sup>Note different categories used for clinician and Reading Center assessments. Shaded cells deemed to represent agreement between the two assessments. Sum of counts in the shaded cells is 89. Agreement is 75% (89/118).

<sup>c</sup>Agreement within one step was defined according to the Reading Center classification levels (i.e., includes cells immediately above or below a shaded cell). Counts in these cells sum to 12. Agreement within one step is 86% ((89+12)/118).