

Methods and Reproducibility of Grading Optimized Digital Color Fundus Photographs in the Age-Related Eye Disease Study 2 (AREDS2 Report Number 2)

Ronald P. Danis,¹ Amitha Domalpally,¹ Emily Y. Chew,² Traci E. Clemons,³ Jane Armstrong,¹ John Paul SanGiovanni,² and Frederick L. Ferris III² for the AREDS2 Study Group⁴

¹Department of Ophthalmology and Visual Sciences, University of Wisconsin–Madison, Madison, Wisconsin

²National Institutes of Health, National Eye Institute, Bethesda, Maryland

³EMMES Corporation, Rockville, Maryland

Correspondence: Ronald P. Danis, Department of Ophthalmology and Visual Sciences, University of Wisconsin–Madison, 8010 Excelsior Drive, Madison, WI 53717; rpdanis@wisc.edu.

⁴See the Appendix in the Supplementary Material for the members of the AREDS2 Study Group.

Submitted: February 1, 2013

Accepted: April 1, 2013

Citation: Danis RP, Domalpally A, Chew EY, et al. Methods and reproducibility of grading optimized digital color fundus photographs in the Age-Related Eye Disease Study 2 (AREDS2 Report Number 2). *Invest Ophthalmol Vis Sci*. 2013;54:4548–4554. DOI:10.1167/iovs.13-11804

PURPOSE. To establish continuity with the grading procedures and outcomes from the historical data of the Age-Related Eye Disease Study (AREDS), color photographic imaging and evaluation procedures for the assessment of age-related macular degeneration (AMD) were modified for digital imaging in the AREDS2. The reproducibility of the grading of index AMD lesion components and for the AREDS severity scale was tested at the AREDS2 reading center.

METHODS. Digital color stereoscopic fundus photographs from 4203 AREDS2 subjects collected at baseline and annual follow-up visits were optimized for tonal balance and graded according to a standard protocol slightly modified from AREDS. The reproducibility of digital grading of AREDS2 images was assessed by reproducibility exercises, temporal drift (regrading a subset of baseline annually, $n = 88$), and contemporaneous masked regrading (ongoing, monthly regrade on 5% of submissions, $n = 1335$ eyes).

RESULTS. In AREDS2, 91% and 96% of images received replicate grades within two steps of the baseline value on the AREDS severity scale for temporal drift and contemporaneous assessment, respectively (weighted Kappa of 0.73 and 0.76). Historical data for temporal drift in replicate gradings on the AREDS film-based images were 88% within two steps (weighted Kappa = 0.88). There was no difference in AREDS2–AREDS concordance for temporal drift (exact $P = 0.57$).

CONCLUSIONS. Digital color grading has nearly the same reproducibility as historical film grading. There is substantial agreement for testing the predictive utility of the AREDS severity scale in AREDS2 as a clinical trial outcome. (ClinicalTrials.gov number, NCT00345176.)

Keywords: color photographic imaging, AMD, reproducibility

The Age-Related Eye Disease Study 2 (AREDS2, www.areds2.org, in the public domain) is a 5-year, multicenter, randomized, controlled clinical trial designed to evaluate the efficacy and safety of daily dietary supplementation with lutein/zeaxanthin (10 mg/2 mg) and/or omega-3 long-chain polyunsaturated fatty acids (LCPUFAs; docosahexaenoic acid [DHA] 350 mg + eicosapentaenoic acid [EPA] 650 mg) on the rate of progression to advanced age-related macular degeneration (AAMD) among people with at least intermediate age-related macular degeneration (AMD) at baseline. AAMD is defined as neovascular AMD determined by a standardized grading procedure applied at the AREDS2 reading center (University of Wisconsin–Madison, Fundus Photograph Reading Center) or treatment for choroidal neovascularization (CNV) or geographic atrophy (GA) involving the center of the macula determined by the AREDS2 reading center.¹ Findings from the Age-Related Eye Disease Study (AREDS) indicate that nutritional supplementation with a zinc and antioxidant combination formulation could decrease the rate of progression to AAMD by 25%.² In AREDS2, a second-tier randomization was applied to investigate the impact of four variations of the original AREDS formulation on the study endpoints.

The color photograph grading procedures developed for AREDS were based upon the Wisconsin Age-Related Maculopathy Grading System,³ which assessed presence, location, and severity of the abnormalities characteristic of AMD. AREDS color fundus image gradings were used to develop severity scales for AMD.^{4,5} These include a detailed nine-step scale for individual eyes and a simplified five-step person-based scale. In the development of the eye-based scale, AMD lesion components such as drusen and pigment abnormalities were examined as risk factors for 5-year progression to AAMD in the right eyes of 3212 participants without AAMD in either eye at baseline. Three closely associated drusen characteristics (area, size, and type) were evaluated. Area of drusen within the AREDS macular grid was the strongest and most consistent predictor. Areas of increased and decreased pigmentation and presence of noncentral GA were assessed and combined in a pigment abnormality assessment. These drusen and pigment subscales were then combined and tested on AREDS AMD endpoints. The 5-year risk of neovascular AMD increased from 0.3% for participants classified as step 1 to 21% in participants classified as step 8. Central GA risk increased from 0.5% in step 4 to 43.5% in step 9.⁴ Reproducibility of the scale was

satisfactory (agreement within two steps was 93.6%, weighted Kappa statistic [SE] 0.73 [0.013]). A simplified scale was also developed, in which presence of large drusen and presence of any pigment abnormality in an eye were each counted as one risk factor; the sum of factors across the two eyes provided a five-step (0–4) person-based scale.⁵ Although persons with lesion characteristics defining the lower steps of these scales are excluded from AREDS2, we hope to evaluate the instrument in independent cohorts.

Use of digital media to represent exposures and outcomes in clinical trials affords an efficiency of data transmission, storage, and processing—as well as ease of portability. Before large-scale transition to such media may take place, it is important to determine whether reproducibility of ascertainment may be reasonably attained at equivalence to that from archived nondigital media. This has been demonstrated in clinical trials for diabetic retinopathy⁶ and for cytomegalovirus and other opportunistic infections in human immunodeficiency syndrome.⁷ This report describes the imaging methodology and grading procedures for color fundus photographs employed in AREDS2, which are modified from those employed in AREDS, particularly with respect to the use of digital imaging in AREDS2, and the reproducibility of classifying eyes according to the AREDS AMD severity scale.

METHODS

Subjects

The AREDS2 study enrolled 4203 participants aged 50 to 85 years after evaluating for eligibility criteria: (1) bilateral large drusen or (2) large drusen in one eye and advanced AMD (neovascular AMD or central GA) in the fellow eye. Further details on subject characteristics are available in AREDS2 Report 1.¹

Photography

All AREDS2 photographers and clinical site digital camera systems are certified by the reading center. Color stereoscopic fundus photographs are obtained using three photographic fields of the macula and optic nerve with 30° or 35° fundus cameras, as in AREDS. The imaging protocol specifies field position and stereoscopic technique. Red reflex photographs with a fundus camera are also obtained, which are helpful in determining the causes for suboptimal-quality image sets. Seven models of digital fundus cameras were permitted for use in AREDS2. All had a minimum resolution specification of 3 megapixels. For baseline image collection, 20 of 82 clinical sites did not have approved digital fundus cameras and were allowed to use Ektachrome color slide film (Eastman Kodak Co., Rochester, NY) for photography. Subsequently, all clinical sites transitioned to digital color photography. Film photograph submissions are digitized at the reading center using a Nikon Super CoolScan 5000 (Nikon, Inc., Melville, NY) to produce 3400 × 2300 pixel images at 8 bits/channel (which produces a 22.4 megabyte image) so that both film-based and natively digital images may be viewed and managed in the same display environment. Digital images are exported and saved in formats specified in the study procedures and are uploaded to the reading center by secure Web site or mailed to the reading center on computer discs. At the reading center, all camera exports are converted to lossless jpeg images (compression to reduce file size without a visible effect on image detail) and are imported into a database for viewing with Topcon IMAGENet software (Topcon Medical Systems, Paramus, NJ), using image scale calibration factors determined for each digital camera

from the certification process. Image quality is monitored throughout the course of the study; and graders evaluate the image quality of each submission according to the degree to which quality affects the ability to discern important lesion features. All ungradable image sets are then evaluated with a detailed technical assessment. Image quality reports are compiled by photographer and by site semiannually, and photographers with more than 5% ungradable images that appear to be due to operator-dependent issues (such as focus, illumination, color balance, field definition, and not related to cataract, small pupil or other media opacity) have a sample of their work reviewed in detail by a reading center photographer to provide assistance in improving the quality of images.

Grading

Grading Tools. Prior to disease evaluation, the tonal properties of the digital image are assessed by optimizing red/green/blue luminance curves through a method developed by Hubbard et al.⁸ Briefly, luminance histograms of image submissions are examined with tools in IMAGENet 2000 software (version 2.56; Topcon Medical Systems). If the tonal characteristics are determined to be not acceptable according to standardized criteria, individual color channel curves are adjusted to match an optimal tonal balance model (if possible). In this manner, the variability in illumination and color contrast between digital images is minimized and more closely resemble film quality. Evaluation is performed using both the original and optimized images. Images are accessed from the database using IMAGENet software (Topcon Medical Systems) for display and measurement on calibrated HP400 monitors (Hewlett Packard, Palo Alto, CA) in calibrated ambient lighting and viewed with a handheld stereoscope (Screen-Vu; PS Manufacturing, Portland, OR). Graders are allowed to use limited zoom features in the display software. An electronic Early Treatment Diabetic Retinopathy Study (ETDRS) macular grid, appropriately sized for the magnification of the digital fundus image, is overlaid to specify the location of some macular lesions by grid subfield, similar to the methodology used in AREDS with acetate overlays on color slides.⁹ Drusen area circles as employed in AREDS are also scaled to the magnification of the photograph (determined at the time of camera system certification) and overlaid on the digital image as needed (Fig. 1). The display software includes mouse-driven planimetry tools that are used for measurements with outputs calibrated to the magnification scale.

Grading Procedure. The grader first examines the image for characteristics of neovascularization using five photographic features: subretinal fluid, serous pigment epithelial detachment, or retinal edema; intraretinal, subretinal pigment epithelium or subretinal blood characteristic of neovascular AMD; intraretinal lipid exudates (hard exudates); subretinal fibrin or fibrosis; and fibrovascular pigment epithelial detachment. If one out of these five features is definitely present, the eye is graded as “questionable neovascular AMD.” Questionable cases are pooled with the corresponding nonneovascular scale step (the severity for the eye if the single neovascular feature was not present) for data analysis purposes. The presence of two or more such features categorizes an eye as “definite neovascular AMD.” Because in current retinal clinical practice, eyes developing neovascular AMD are usually promptly treated with intravitreal antivascular endothelial growth factor agents such as ranibizumab, bevacizumab, or aflibercept,¹⁰ eyes with neovascular AMD may have no evident signs of neovascularization on color photographs obtained at an annual AREDS2 study visit. The AREDS2 coordinating center (EMMES Corporation, Rockville, MD) collects information from

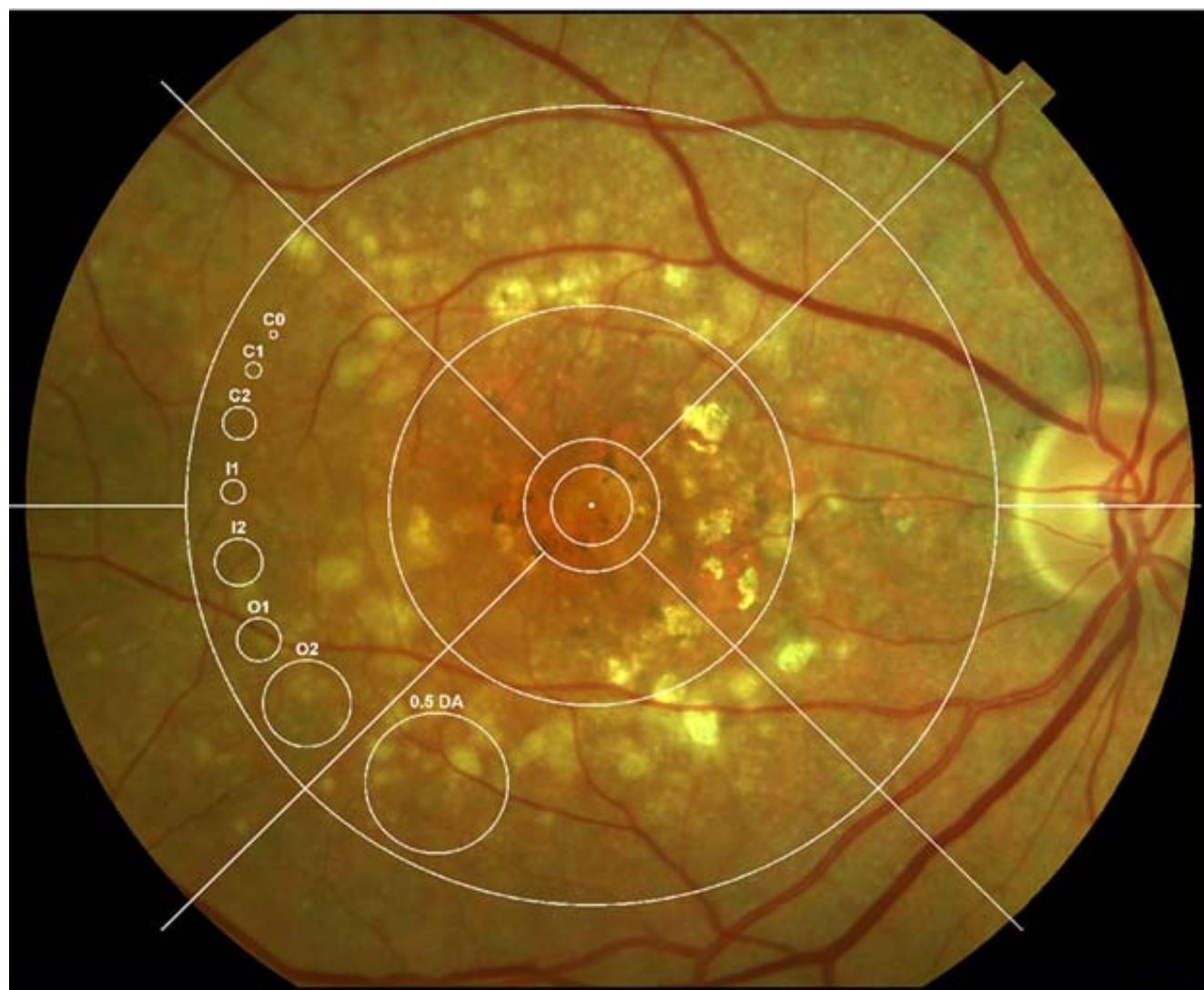


FIGURE 1. Screen capture with digital AREDS macular grid and drusen circles on color fundus photograph in IMAGENet. See Table 1 for area definitions of grading circles.

the clinical sites regarding the treatment of CNV, which is used to classify eyes with neovascular AMD.

The nonneovascular or atrophic AMD features documented in the image evaluation process are similar to the grading used in AREDS. Drusen size, type, and area and the presence and severity of hyper- and hypopigmentation are evaluated as in AREDS (Table 1). Maximum drusen size is estimated using standard drusen circles⁹ scaled to the magnification of the digital image. Area of drusen within the grid is estimated by mentally moving together all drusen present into the appropriate drusen measurement circle, as described in AREDS. Area of GA is directly measured using computerized planimetry with measurements scaled to the magnification of the digital photograph. GA is considered definitely present if the lesion is at least 433 μm in diameter (greater than or equal to the AREDS circle I-2) and if at least two of the following features are present: absence of RPE pigment, circular shape, or sharp margins.⁴ If noncentral, the proximity of the border of the atrophy closest to the center (umbo of the foveola) is documented. In addition, the configuration and severity of the GA are also evaluated. Because drusen and pigmentation changes tend to be small areas and multifocal, planimetry is less useful for quantitative area measurement of these features;

in the digital environment, standard drusen circles and Macular Photocoagulation Study measurement circles are used for area estimations.

Quantitative evaluation is limited to the area within the grid (Fig. 1), which is modified from AREDS. AREDS2 has adopted the modern assumption that the standard disc diameter (DD) equals 1800 μm , rather than retaining the classic assumption that it equals 1500 μm .¹¹⁻¹³ This change is required to accommodate digital imaging in AREDS2, since most cameras are calibrated with 1800 μm equivalency. Consequently, the grid used in AREDS2 is modified from AREDS. The diameters of the grid circles remain as 1/3 DD, 1 DD, and 2 DD, but are now equivalent to 600 μm , 1800 μm , and 3600 μm , respectively (Table 1, Fig. 1).⁹ AREDS2 and AREDS data may be compared by applying a scaling constant to the AREDS data. For linear measurements, this is accomplished by multiplying by 1.20, and for area measurements by multiplying by 1.44.

Baseline AREDS2 images were graded by two independent graders. Grading results were assessed by a software processor, and discrepancies on major questions (component questions for the AREDS2 severity scale) were adjudicated by a third, senior grader (JA). If no grading discrepancies were identified, the first grade was submitted as the grade of record. For annual

TABLE 1. Lesions Graded From Color Photographs in AREDS2

Macular Feature Graded	Grading Data Collected
Area of hyperpigmentation	4-step scale: Absent, Questionable,* Definite < C-2,† Definite ≥ C-2
Area of hypopigmentation	4-step scale: Absent, Questionable, Definite < I-2, Definite < 1/2 DA, Definite ≥ 1/2 DA
Maximum drusen size	6-step scale: Absent, Questionable, Definite < C-0 μm, Definite < C-1 μm, Definite < C-2 μm, Definite ≥ C-2 μm
Area of drusen within the grid	8-step scale: <C-0, <C-1, <C-2, <I-2, <O-2, <1/2 DA, <1 DA, ≥1 DA
Calcified drusen, drusenoid PED, reticular drusen	Presence/absence
Area of geographic atrophy within the grid	DA (disc areas) by planimetry (computer mouse)
Involvement of the center of the macula by geographic atrophy	Presence/absence
Configuration of geographic atrophy	Small single patch < 1 DA, multifocal, horseshoe, ring, solid not center, solid center involved, indeterminate ³⁰
Neovascular AMD	Presence/absence

PED, pigment epithelial detachment.

* Questionable is used as a distinct grading step in AMD classification.

† Drusen circles for measurement are conserved between AREDS and AREDS2. Using the convention that one disc area is 2.54 mm², the drusen circles have the following areas in square millimeters (diameter in microns): C-0 = 0.004 mm² (76 μm), C-1 = 0.017 mm² (150 μm), C-2 = 0.069 mm² (300 μm), I-2 = 0.146 mm² (433 μm), O-2 = 0.487 mm² (395 μm), 1/2 DA = 1.27 mm² (1273 μm), 1 DA = 2.5 mm² (1800 μm). See also Figure 1.

follow-up images, the grading process consists of single-step grading, independent of prior visit and fellow eye images and data.

Grading Quality Control

The AREDS2 grading quality control program consists of several elements. A temporal drift sample of 88 stratified baseline images is regraded annually by the entire grading group; the results are compared to original grades for the same sample. The temporal drift reproducibility exercises allow monitoring the shift due to grader experience, change in grading personnel, and technological advances, particularly in studies with long follow-up such as AREDS2. The contemporaneous quality control includes monthly regrade of a random sample of 5% of submissions. These images are duplicated and passed through the grading process with fictitious identifiers for masked replicate grading. The reproducibility of grading is assessed by calculating percentage agreement and weighted Kappa statistics for ordinal variables and correlation coefficients for continuous area measurements for the entire group. Inspection of scatter plots for outliers and potential systematic bias is performed for temporal drift or problematic questions for targeted retraining if necessary. Regular training exercises are held for the entire grading group with review of difficult cases and reaffirmation of the grading protocol. Reproducibility statistics are also examined for individual graders, and

targeted individual retraining is performed if the grader has reproducibility for specific questions below a set threshold. All graders are encouraged to seek out a reading center ophthalmologist for “second opinions” for assistance with unusual presentations or confounding ocular abnormalities. On an ongoing basis, any eyes meeting the study endpoint are reviewed by a reading center ophthalmologist to confirm the endpoint.

Outcomes

For categorical variables on an ordinal scale, such as drusen area or AREDS severity scale, agreement is measured in terms of percentage agreement and weighted Kappa. For continuous variables such as area of GA, Bland-Altman plots have been presented.

RESULTS

Table 2 shows the results of the temporal drift quality control from the fourth year of the AREDS2 study, contemporaneous quality control data over 4-year period, and published data for AREDS temporal drift. The reproducibility values of grading for AMD lesion components from optimized digital photographs in AREDS2 were essentially equivalent to those from film images in AREDS (Table 2). For AREDS2-AREDS concordance of

TABLE 2. AREDS2 Grading Reproducibility for Select Ordinal Variables in Temporal Drift and Contemporaneous Regrading Samples Compared With Published Data From AREDS

Feature	AREDS2 Temporal Drift Regrade Year 4 Compared to BL, N = 88		Contemporaneous Regrades, N = 1335		Historical AREDS Temporal Drift (AREDS Report 6 and 17), ^{4,9} N = 119	
	% Agree	wKappa (SE)	% Agree	wKappa (SE)	% Agree	wKappa (SE)
Increased pigment within grid	69	0.63 (0.04)	70	0.54 (0.03)	58	0.60 (0.05)
Hypopigmentation within grid	60	0.56 (0.03)	69	0.56 (0.02)	63	0.46 (0.06)
Drusen area within grid	60	0.71 (0.02)	69	0.74 (0.01)	65	0.77 (0.04)
Presence of geographic atrophy	88	0.85 (0.03)	95	0.79 (0.08)	90	0.73 (0.08)
Presence of neovascular AMD*	90	0.79 (0.03)	94	0.86 (0.02)	—	—
AREDS severity scale (two-step agreement)	92	0.73 (0.02)	96	0.76 (0.01)	94	0.73 (0.01)

BL, baseline; wKappa, weighted Kappa.

* Component features of neovascular AMD graded in AREDS; cannot be compared with a summary grade in AREDS2.

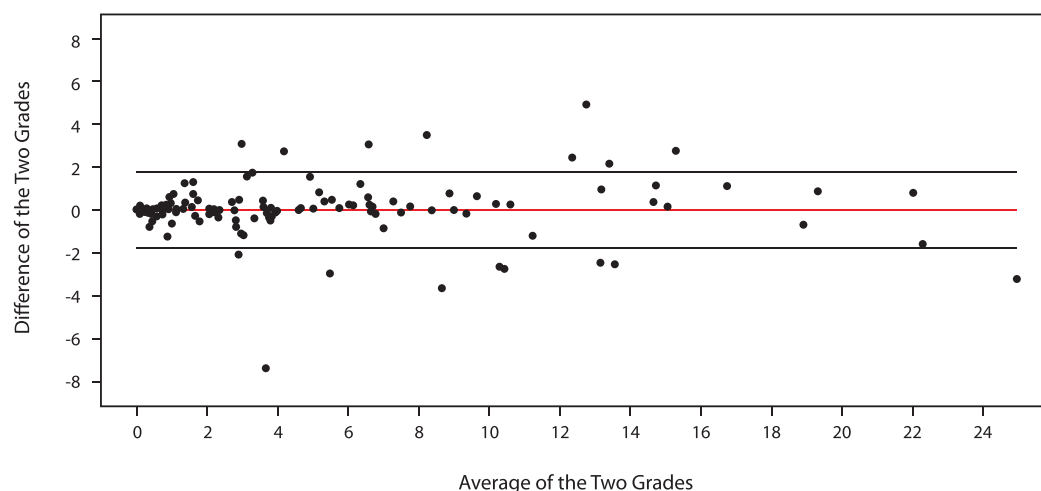


FIGURE 2. Bland-Altman plot for reproducibility of area of geographic atrophy ($n = 288$ eyes). Area measured by two graders using planimetry had a mean difference of 0.02 mm^2 and Bland-Altman limits of $+1.8$ and -1.76 mm^2 .

temporal drift on the two-step agreement in the AREDS severity scale (91.2% vs. 93.6%), an exact test of binomial proportions indicated no difference between AREDS2 and AREDS or in the grading of the component variables of the scale. When masked contemporaneous regrading quality control statistics were compared to the temporal drift samples, the same values for reproducibility were found. Geographic atrophy area measured by planimetry had a mean difference of 0.02 mm^2 with Bland-Altman limits of $+1.8$ and -1.76 mm^2 (Fig. 2). No systematic bias was identified.

DISCUSSION

The grading of features of AMD from optimized digital color fundus photographs in AREDS2 has reproducibility equivalent to historical grading from color slides in AREDS. This indicates continuity of AMD lesion classification from AREDS to AREDS2, which is important for comparison between data sets. A secondary outcome of AREDS2 is to evaluate the utility of the AREDS AMD severity scale for disease staging and possibly to predict outcomes. Because the severity scale was developed using color film images, there was question whether the macular lesion characteristics graded with digital photography were sufficient to reasonably justify use of the same grading steps. The agreement within two steps on the AREDS severity scale between the AREDS and AREDS2 contemporaneous sample is 93.6% and 95.6%, respectively.⁴ Our results indicate that overall there are no gross differences (Table 2). The use of post hoc standardized image optimization for illumination and color balance of digital fundus photographs is important to help overcome the increased variability in image quality of digital compared to film images.⁸

The grading schema employed in AREDS2 is focused on quantification of three key features to arrive at the nine-step AREDS severity scale: drusen area, hyperpigmentation, and hypopigmentation. The individual lesion components are graded using categorical scales (as opposed to a count of number of drusen or area of pigment changes). The reproducibility of grading some macular features, such as area of hypopigmentation, is only moderate (weighted Kappa 0.56, Table 2). The grading of hypopigmentation from color photographs is difficult and is highly dependent upon image quality. Despite this, the reproducibility of grading on the AREDS severity scale is substantial (91% agreement, weighted Kappa 0.73, Table 2). Two-step change along the severity scale

may be a useful outcome for eyes enrolled in clinical trials with early or moderate atrophic AMD, similar to how the ETDRS diabetic retinopathy scale is employed in many clinical studies. When all data have accrued at the end of the clinical trial, the AREDS2 severity levels and change in levels will be evaluated with advanced AMD outcomes in order to test this hypothesis.

An advantage of grading color images in a digital environment is that it allows contemporaneous masked regrading for quality control. In AREDS, color film slides were individually labeled and set as stereo pairs in appropriate anatomic relations into plastic sheets, which were also labeled. The effort involved in relabeling slides for quality control was prohibitive in AREDS. Duplication of color slides under the best of circumstances resulted in images that were different enough from the film originals that an experienced grader could detect them and thereby become unmasked to the reproducibility exercise. Therefore, in AREDS, all quality control was performed unmasked to the type of grading, which theoretically could bias the reproducibility results. The duplication of images in a digital environment and assigning fictitious identifiers for evaluator masking are a relatively easy process in AREDS2, in which all reproducibility grading is masked.

An important endpoint in the color photograph grading for atrophic AMD is the enlargement of GA and the development of GA in the center of the macula. At present, there are multiple alternative image types for this assessment, including fundus autofluorescence^{14,15} and spectral domain optical coherence tomography.¹⁶ Each imaging type has its unique advantages and disadvantages, and the superiority of one image type over another requires comparative grading in the same set of eyes and correlation with clinically important outcomes. In AREDS2, the reproducibility of grading the area of GA from color fundus photographs is high. Of note, because the unit of measurement used in both AREDS and AREDS2 is disc area (DA), the grading area measurements are directly comparable. This is in spite of the revision of the assumed dimensions of the DA from 1.80 mm^2 to the current convention of 2.54 mm^2 .

Automated lesion detection has been employed with some success to measure drusen extent from fundus photographs in eyes with AMD.^{17–19} Advances in technology may make automated grading from color photographs feasible, which is desirable not only to decrease costs but also because machine grading usually has higher reproducibility than human grading. In addition, there has been developing interest in use of

technology alternative to color photographs for the measurement and classification of drusen, such as spectral/Fourier domain optical coherence tomography (OCT).^{20–25} The classification of drusen and measurement of drusen area and volume from OCT is especially promising because it appears well within the technical capabilities of current technology. Because these OCT variables are fundamentally different measurements than those from color photographs, it may not be surprising if the results are not directly comparable to those from color photograph grading. To date, no large data sets evaluated by both methodologies have been reported. OCT measurement of drusen is an ancillary study in AREDS2 that will hopefully shed light upon the relationship between drusen variables measured by both imaging types and relationships between OCT measurements and clinical outcomes. Classification of AMD by fundus autofluorescence imaging also has been proposed.^{14,26–29} Fundus autofluorescence patterns may provide prognostic information for AMD outcomes. Fundus autofluorescence may be particularly useful for the semi-automated measurement of area of GA,¹⁵ although measurements are, again, sometimes discrepant with color photographs.¹⁴ Autofluorescence imaging is another ancillary study within AREDS2 that should help clarify its predictive value. Advantages to the classification of AMD from color photographs, apart from its validation in epidemiologic studies and clinical trials, include the fact that it does not require technology beyond common clinical practice (such as a scanning laser ophthalmoscope for autofluorescence or a spectral domain OCT) and that it is the image type most similar to the view of the physician examining a patient. This makes extrapolation from clinical trials to clinical practice somewhat easier.⁵ As potential therapeutics target earlier stages of AMD, a classification system based upon outcomes validated with a large prospective cohort is increasingly important.

Acknowledgments

Supported by National Eye Institute Grant HHS-N-260-2005-00007-C (TEC).

Disclosure: **R.P. Danis**, None; **A. Domalpally**, None; **E.Y. Chew**, None; **T.E. Clemons**, None; **J. Armstrong**, None; **J.P. SanGiovanni**, None; **F.L. Ferris III**, None

References

- Chew EY, Clemons T, Sangiovanni JP, et al. The Age-Related Eye Disease Study 2 (AREDS2): study design and baseline characteristics (AREDS2 report number 1). *Ophthalmology*. 2012;119:2282–2289.
- Age-Related Eye Disease Study Group. A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C and E, beta carotene, and zinc for age-related macular degeneration and vision loss: AREDS report no. 8. *Arch Ophthalmol*. 2001;119:1417–1436.
- Klein R, Davis MD, Magli YL, Segal P, Klein BE, Hubbard L. The Wisconsin age-related maculopathy grading system. *Ophthalmology*. 1991;98:1128–1134.
- Davis MD, Gangnon RE, Lee IY, et al. The Age-Related Eye Disease Study severity scale for age-related macular degeneration: AREDS Report No. 17. *Arch Ophthalmol*. 2005;123:1484–1498.
- Ferris FL, Davis MD, Clemons TE, et al. A simplified severity scale for age-related macular degeneration: AREDS Report No. 18. *Arch Ophthalmol*. 2005;123:1570–1574.
- Hubbard LD, Sun W, Cleary PA, et al. Comparison of digital and film grading of diabetic retinopathy severity in the Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications Study. *Arch Ophthalmol*. 2011;129:718–726.
- Gangaputra S, Won Pak J, Peng Q, et al. Transition from film to digital fundus photography in the Longitudinal Studies of the Ocular Complications of AIDS. *Retina*. 2012;32:600–605.
- 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.
- The Age-Related Eye Disease Study Research Group. The Age-Related Eye Disease Study system for classifying age-related macular degeneration from stereoscopic color fundus photographs: the Age-Related Eye Disease Study Report Number 6. *Am J Ophthalmol*. 2001;132:668–681.
- Ip MS, Scott IU, Brown GC, et al. Anti-vascular endothelial growth factor pharmacotherapy for age-related macular degeneration: a report by the American Academy of Ophthalmology. *Ophthalmology*. 2008;115:1837–1846.
- Jonas JB, Gusek GC, Guggenmoos-Holzmann I, Naumann GO. Variability of the real dimensions of normal human optic discs. *Graefes Arch Clin Exp Ophthalmol*. 1988;226:332–336.
- Danis RP, Scott IU, Qin H, et al. Association of fluorescein angiographic features with visual acuity and with optical coherence tomographic and stereoscopic color fundus photographic features of diabetic macular edema in a randomized clinical trial. *Retina*. 2010;30:1627–1637.
- Davis MD, Bressler SB, Aiello LP, et al. Comparison of time-domain OCT and fundus photographic assessments of retinal thickening in eyes with diabetic macular edema. *Invest Ophthalmol Vis Sci*. 2008;49:1745–1752.
- Sunness JS, Ziegler MD, Applegate CA. Issues in quantifying atrophic macular disease using retinal autofluorescence. *Retina*. 2006;26:666–672.
- Holz FG, Bindewald-Wittich A, Fleckenstein M, Dreyhaupt J, Scholl HP, Schmitz-Valckenberg S. Progression of geographic atrophy and impact of fundus autofluorescence patterns in age-related macular degeneration. *Am J Ophthalmol*. 2007;143:463–472.
- Yehoshua Z, Rosenfeld PJ, Gregori G, Penha F. Spectral domain optical coherence tomography imaging of dry age-related macular degeneration. *Ophthalmic Surg Lasers Imaging*. 2010;41(suppl):S6–S14.
- Shin DS, Javornik NB, Berger JW. Computer-assisted, interactive fundus image processing for macular drusen quantitation. *Ophthalmology*. 1999;106:1119–1125.
- Smith RT, Sohrab MA, Pumariega NM, et al. Drusen analysis in a human-machine synergistic framework. *Arch Ophthalmol*. 2011;129:40–47.
- Quelleg G, Russell SR, Scheetz TE, Stone EM, Abramoff MD. Computational quantification of complex fundus phenotypes in age-related macular degeneration and Stargardt disease. *Invest Ophthalmol Vis Sci*. 2011;52:2976–2981.
- Yi K, Mujat M, Park BH, et al. Spectral domain optical coherence tomography for quantitative evaluation of drusen and associated structural changes in non-neovascular age-related macular degeneration. *Br J Ophthalmol*. 2009;93:176–181.
- Khanifar AA, Koreishi AE, Izatt JA, Toth CA. Drusen ultrastructure imaging with spectral domain optical coherence tomography in age-related macular degeneration. *Ophthalmology*. 2008;115:1883–1890.
- Jain N, Farsi S, Khanifar AA, et al. Quantitative comparison of drusen segmented on SD-OCT versus drusen delineated on color fundus photographs. *Invest Ophthalmol Vis Sci*. 2010;51:4875–4883.
- Iwama D, Tsujikawa A, Ojima Y, et al. Relationship between retinal sensitivity and morphologic changes in eyes with

- confluent soft drusen. *Clin Experiment Ophthalmol*. 2010;38:483–488.
24. Freeman SR, Kozak I, Cheng L, et al. Optical coherence tomography-raster scanning and manual segmentation in determining drusen volume in age-related macular degeneration. *Retina*. 2010;30:431–435.
 25. Leng T, Rosenfeld PJ, Gregori G, Puliafito CA, Punjabi OS. Spectral domain optical coherence tomography characteristics of cuticular drusen. *Retina*. 2009;29:988–993.
 26. Bindewald A, Bird AC, Dandekar SS, et al. Classification of fundus autofluorescence patterns in early age-related macular disease. *Invest Ophthalmol Vis Sci*. 2005;46:3309–3314.
 27. Bindewald A, Schmitz-Valckenberg S, Jorzik JJ, et al. Classification of abnormal fundus autofluorescence patterns in the junctional zone of geographic atrophy in patients with age related macular degeneration. *Br J Ophthalmol*. 2005;89:874–878.
 28. Delori FC, Fleckner MR, Goger DG, Weiter JJ, Dorey CK. Autofluorescence distribution associated with drusen in age-related macular degeneration. *Invest Ophthalmol Vis Sci*. 2000;41:496–504.
 29. Spaide RF. Fundus autofluorescence and age-related macular degeneration. *Ophthalmology*. 2003;110:392–399.
 30. Sunness J, Bressler N, Tian Y, Alexander J, Applegate C. Measuring geographic atrophy in advanced age-related macular degeneration. *Invest Ophthalmol Vis Sci*. 1999;40:1761–1769.