

Evaluation of Optimized Digital Fundus Reflex Photographs for Lens Opacities in the Age-Related Eye Disease Study 2: AREDS2 Report 7

Amitha Domalpally,¹ Ronald P. Danis,¹ Emily Y. Chew,² Traci E. Clemons,³ Susan Reed,¹ John Paul SanGiovanni,² and Frederick L. Ferris III² for the Age-Related Eye Disease Study 2 Research Group⁴

¹Fundus Photograph Reading Center, Department of Ophthalmology and Visual Science, Madison, Wisconsin

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

³EMMES Corporation, Rockville, Maryland

Correspondence: Ronald P. Danis, Department of Ophthalmology and Visual Science, 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 Age-Related Eye Disease Study 2 Research Group

Submitted: April 26, 2013

Accepted: July 22, 2013

Citation: Domalpally A, Danis RP, Chew EY, et al. Evaluation of optimized digital fundus reflex photographs for lens opacities in the Age-Related Eye Disease Study 2: AREDS2 Report 7. *Invest Ophthalmol Vis Sci.* 2013;54:5989-5994. DOI:10.1167/iops.13-12301

PURPOSE. We described the system for grading lens opacities using stereoscopic digital fundus reflex photographs in the Age-Related Eye Disease Study 2 (AREDS2) and compared reproducibility with the AREDS lens grading system, which used retroillumination film images.

METHODS. Stereoscopic fundus reflex photographs were acquired in a standardized fashion at baseline and annually. Images were enhanced and evaluated in the red channel at a central reading center. Percentage involvement of cortical and posterior subcapsular (PSC) lens opacities within the central 5 mm diameter zone of a modified AREDS lens grid was estimated. Reproducibility was assessed for contemporaneous variability (ongoing, monthly regrade on 5% of submissions, $n = 777$ eyes) and temporal drift (regrading a subset of baseline photographs annually, $n = 88$).

RESULTS. In the contemporaneous exercise, the agreement for presence of cortical opacities was 93% ($\kappa = 0.86$) and for PSC opacities it was 97% ($\kappa = 0.83$). Intraclass correlation (ICC) for area of central zone involvement was 0.95 for cortical and 0.99 for PSC opacities. Historic data for contemporaneous regrading of film-based images in AREDS showed an ICC of 0.94 for cortical and 0.82 for PSC. The final annual temporal drift exercise had a reproducibility of 95% for cortical and PSC opacities.

CONCLUSIONS. Digital grading using fundus reflex images with image enhancing tools has reproducibility comparable to film-based retroillumination images, and may be useful for centralized objective lens opacity assessment in clinical trials using widely available fundus cameras. Red reflex images limit evaluation to cortical and PSC opacities, and do not permit assessment of nuclear opacities. (ClinicalTrials.gov number, NCT00345176.)

Keywords: cataract, lens opacity, AREDS2

Cataract is the leading cause of reversible visual loss on a global scale¹ and in the United States.² Age-related cataracts account for 40% to 51% of visual impairment in the United States.³ Because of the largely subjective nature of assessing severity of lens opacities in a clinical setting, photographic documentation of lens opacities has been used in epidemiologic studies and clinical trials. Imaging methods have involved slit-lamp and Scheimpflug photography for grading nuclear cataracts, and retroillumination techniques for grading cortical and posterior subcapsular (PSC) opacities.⁴ Various standardized systems have been developed for evaluation of lens opacities using images. These include the Lens Opacity Classification System (LOCS), Hopkins system, Wisconsin Cataract Grading System, Age-Related Eye Disease Study Grading system, and World Health Organization (WHO) cataract grading system.⁵⁻⁹ There is no established gold standard for the evaluation of lens opacities, and the choice of subjective or objective system depends on the availability of capture systems and the outcomes.¹⁰ Clinical evaluation of

cataract has been shown to be comparable to centralized objective evaluation at reading center.¹¹ However, an advantage of centralized reading of images is the possibility of independent, masked assessment to minimize bias in clinical trials.

The Age-Related Eye Disease Study (AREDS) was a randomized clinical trial designed to evaluate the effects of high-dose antioxidant vitamins and zinc on the incidence and progression of age-related macular degeneration (AMD) and lens opacities.¹² Slit-lamp film images were obtained for evaluation of nuclear sclerosis using a decimal scale. Stereoscopic anterior and posterior retroillumination film images were captured using Neitz cameras (Kowa Optical, Torrance, CA) and evaluated for cortical and PSC opacities. The study used a modification of the Wisconsin Cataract Grading system for evaluation of lens opacities.^{5,6} The Age-Related Eye Disease Study 2 (AREDS2) assessed the effects of dietary xanthophylls and omega 3 fatty acids on progression of AMD and cataract. Because AREDS2 included 82 clinical sites at academic and community centers with retinal specialists, it was beyond the scope of the study to

evaluate lens opacities using specialized cameras. However, fundus reflex photographs taken with standard fundus cameras could be incorporated for grading the severity of cortical and PSC lens opacities. The grading of nuclear sclerosis requires specialized equipment, such as a slit-lamp camera and cannot be done reliably with fundus reflex photographs. Because of the absence of data on nuclear cataract, which is a major cause of cataract surgery, the main cataract outcomes in AREDS2 were: (1) progression to cataract surgery or a 10% absolute increase in the area of cortical opacity within the central 5 mm of the lens, or a 5% increase in the area of PSC opacity within the central 5 mm of the lens; and (2) progression to cataract surgery or a 20% absolute increase in the area of either opacity within the central 5 mm of the lens. We describe the methodology for evaluating cortical and PSC opacities in the central 5 mm of the lens and the reproducibility of gradings using digital stereoscopic fundus reflex images with centralized masked assessment of the lens opacities.

METHODS

Photography

At each annual visit, stereoscopic fundus reflex photographs of both eyes were taken by certified photographers with certified equipment. Dilation of the pupils to at least 6 mm was attempted to obtain good stereoscopic images and to capture the region of interest. Diagnostic contact lens examination was avoided before imaging to prevent corneal epithelial haze. A fixation target was positioned to direct the participant's gaze towards the primary (straight ahead) position, so that the optic nerve did not appear directly behind the lens and alter the red reflex. Image focus was maintained on the pupillary margin. The camera was moved laterally approximately 3 mm between exposures to obtain good stereo effect.

Image Optimization

For lens grading, digital images were displayed in a viewing platform (IMAGENet system, version 2.56; Topcon, Paramus, NJ). The brightness, contrast, and color balance were standardized for all digital images using a histogram-based enhancement procedure, similar to the one used for retinal images.¹³ Before evaluation, software tools were used to display the image in the red color channel (of the red, green, and blue [RGB] image) with a luminance histogram. The color channel is represented as a luminance curve with the *x*-axis showing the intensity values on a 256-step scale, where 0 is the darkest point and 255 the brightest point. The *y*-axis shows the number of pixels in a given intensity value. The brightness and contrast of the image were optimized by adjusting the luminance curve so that the lower end of the spectrum was near the 0 pixel space. The upper end of the curve was expanded until the image of the light reflex in the red channel just began to enlarge (saturate). The enhanced image in the red channel was saved separately so the grader had access to the original and optimized images in the rare event that the optimization process was ineffective.

Image Quality

Each submission was assessed subjectively for image quality according to the degree to which important lesion features could be discerned. Images deemed ungradable received a formal assessment of technical quality by imaging specialists. Technical or operator-dependent issues for ungradable quality included poor focus, stereo, or illumination; incorrect magnification; and poor resolution of images. Inadequate pupil size

was considered a patient-related issue with the assumption that this was the best pharmacologically achieved dilation for the patient. In some eyes with severe retinal pathology, red reflex could not be obtained and was considered ungradable. Photographer and site image quality reports were compiled semiannually. Photographers with more than 5% ungradable images that appeared to be due to operator-dependent issues (such as focus, illumination, color balance, stereo, small pupil, or a dry cornea) had a sample of their work reviewed in detail by a reading center photographer along with the photographer to provide feedback assistance and help improve the quality of images.

Grading

All images were evaluated by a certified ocular disease evaluator (grader) using Topcon IMAGENet software for display and a handheld stereoscope (Screen-Vu, Portland, OR). Images were graded longitudinally, that is, with access to images and evaluation data from previous visits. Images from all visits were grouped in the same patient folder. Based on iris color and pattern, the grader confirmed that all visits belonged to the same patient.

The status of the lens was classified as phakic, pseudo-phakic, or aphakic. The visibility of lens capsule, a haptic, or a positioning hole was clear evidence of pseudophakia. In the absence of such evidence, the distinction between a clear lens and an intraocular lens was based upon the central light reflex. A single central light reflex usually indicates a phakic eye and presence of more than one reflex indicates pseudophakia. The light reflex also helps identify right-left stereo orientation. With appropriately placed stereoscopic photograph pairs, the reflex is expected to be in the posterior region of the lens. Further detailed evaluation was performed in phakic eyes only.

In phakic eyes, a digital grid was overlaid on the left stereoscopic photograph pair. The grid had been modified from AREDS by the addition of a fourth outermost circle used to align the grid with the limbus. The diameters of the 4 circles were 2, 5, 8, and 12 mm. The area within the 2 mm circle was called the central subfield (or circle), between the 2 and 5 mm rings was the inner zone, and between 5 and 8 mm was the outer zone. The central subfield and inner zone together were referred to as the central zone, the primary area of evaluation in AREDS2. The space between the 8 and 12 mm rings was not evaluated. Eight equally spaced radial lines divided the inner and outer zones at the following clock-hour positions 10:30, 12:00, 1:30, 3:00, 4:30, 6:00, 7:30, and 9:00 creating 16 subfields. Each meridian had a 0.5 mm hatch mark. A lens grid drawing tool was used to draw a horizontal line from nasal to temporal across the horizontally visible iris diameter to calibrate the grid and center it. Using the horizontally visible iris diameter rather than the pupillary margin for grid centering is advantageous in asymmetric dilation. Horizontal and vertical pupil diameters were measured in mm using line tools from the inner edge of the pupillary margin. In case of lid obstruction, the horizontal measurement was duplicated for vertical diameter also.

Areas of cortical and PSC opacity involvement were evaluated using percentage assessment in each of the 16 subfields and the central circle. Typical cortical opacities are linear or wedge-shaped radially oriented spokes, originating in the periphery of the lens and pointing towards the center. The appearance can vary as a string of vacuoles or dense black opacities. Anterior subcapsular cataracts, although rare, were graded as cortical opacities. The grader estimated and recorded the percentage, to the nearest whole number, of area covered by cortical opacities in each subfield. In estimating area involved by stippling, the grader mentally swept the opacities

together and estimated the area they would cover if contiguous. Image quality and pupillary dilation can have a significant effect on identification of cortical opacities. For the outer subfields, presence of significant lens opacity, that is >50% of visible subfield involvement, was documented. The outer zone was considered ungradable if presence or absence of opacity could not be documented in at least half of the zone.

Longitudinal grading can help reduce noise due to variable image quality between visits. The graders had access to previous and baseline visit measurements in the electronic forms. For ease of data entry, the measurements for previous and baseline visits were displayed alongside pertinent fields. Since images of all visits also were grouped together, the grader could review the images to confirm changes at any time point. If definite cortical opacity was seen at a visit, but disappeared at a follow-up visit due to issues with focus, the grader labeled the subfield as ungradable. A subfield also was considered to be ungradable if less than two-thirds of that subfield was visible, for example, if a subfield was limited in size by the extent of pupillary dilation or ptosis.

Identification of PSC opacities was based on their location in the posterior central zone of the lens. The peripheral extent of opacity was used for the area measurement without any attempt to subtract clear spaces, even if PSC was lacy with small open areas. The light reflex was included within the estimation if the PSC surrounded it. If a lens had cortical and PSC opacities with overlapping borders, the grader attempted to distinguish the boundaries, so that the area covered by each type of opacity was estimated separately. Other opacities, such as white anterior cortical opacities (WACO), Mittendorf's dot, and pseudoexfoliation of lens, also were documented, but not included in lens opacity measurements.

Quality Control

Reproducibility was assessed in terms of temporal drift and contemporaneous agreement. For temporal drift exercises, images were selected from baseline. The eligibility criteria required relatively clear lenses at baseline to enable retinal imaging. A random selection from baseline dataset most probably would have given a low yield of eyes with lens opacities. To avoid this, images were selected stratified by area of cortical opacity, and graded annually by all the graders participating in the study. For contemporaneous quality control, 5% of images from the previous month's grades were selected randomly and circulated in a masked fashion for regrade. Regrades also were done using longitudinal data. Case reviews were held with Reading Center graders every month to review the evaluation protocol and discuss difficult-to-grade images with a reading center ophthalmologist.

Statistical Analysis

The original grade and the regrade were compared for reproducibility. Categorical variables (presence/absence) were assessed using percentage agreement and κ statistics. Continuous variables (measurements) were compared using intraclass correlation (ICC). Bland Altman plots also were created with mean difference between the grades and 95% confidence limits (CL).

RESULTS

Image Quality

Images were deemed ungradable in 0.7% (318 eyes) of the submissions. Of these, technical (operator-dependent) issues

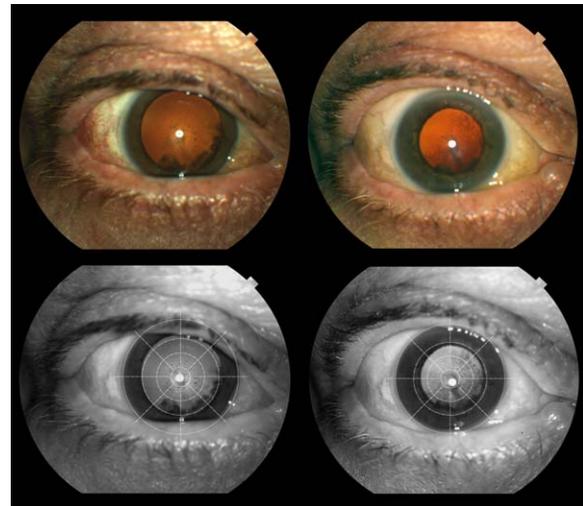


FIGURE 1. Fundus reflex images showing progression of cortical and posterior subcapsular cataract at baseline and five years. *Top row* shows original images and *bottom row* shows enhanced red channel images with the modified AREDS grid.

were found in 63%, patient-related issues in 6% (e.g., ptosis, small pupils), and a combination of the two in 29%.

Grading Quality

Fundus reflex images from 88 eyes were evaluated by 6 graders on an annual basis to test for temporal drift (Fig. 1). Agreements for presence of lens opacities between the first and last temporal drift exercise are shown in the Table. The ICC for the percentage of cortical opacity in the central zone was 0.74 in the first annual exercise with a mean difference of 2.46% (95% CL $-20.67, 15.75$). For the temporal drift exercise at the end of the study, the ICC was 0.86 for cortical opacities with a mean difference of 0.47% (95% CL $-15.18, 14.24$). Agreement for presence of PSC opacity was 95%. For PSC opacity in the central zone, the ICC was 90% in the first year and 95% in the final year. The sample with measurable PSC was insufficient for further analysis.

To assess contemporaneous agreement, 1293 images were reevaluated. Of these 777 eyes that were phakic, there was 93% (κ 0.86) agreement for cortical opacities and 97% (κ 0.83) agreement for PSC opacities. The ICC for percentage of cortical opacity within the central zone was 0.95, with a mean difference between the two grades of 0.1% (95% CL, $-5.36-5.56$, Fig. 2A). For PSC opacities within the central zone, the ICC was 0.99 with a mean difference of 0.16% (95% CL, $-3.82-4.14$, Fig. 2B).

DISCUSSION

Standardized image capture is of great benefit in multicenter clinical trials of age-related cataract, because it allows the pooling of data across clinical sites, and provides the potential for an objective and quantifiable metric for analysis. Compared to a site-based interpretation of images, a centralized masked independent assessment of outcomes often is preferred in clinical trials in the regulatory environment.¹⁴ Our study described one kind of methodology for objective centralized evaluation of lens opacities, and reported its feasibility and reproducibility. The strength of this methodology is its feasibility in a large scale trial without the use of

TABLE. Intergrader Agreement for Presence or Absence of Cortical and PSC Lens Opacities

Variables	Temporal Drift Y 1, n = 88 Eyes		Temporal Drift Y 6, n = 88 Eyes		Contemporaneous, n = 777 Eyes	
	Agreement, %	κ (95% CL)	Agreement, %	κ (95% CL)	Agreement, %	κ (95% CL)
Presence of any lens opacity	90%	0.67 (0.47, 0.87)	89%	0.62 (0.42, 0.82)	89%	0.75 (0.7, 0.8)
Presence of cortical opacities	95%	0.89 (0.79, 1)	95%	0.88 (0.77, 0.98)	93%	0.86 (0.82, 0.89)
Presence of posterior subcapsular opacities	90%	0.35 (0.03, 0.67)	95%	0.64 (0.35, 0.93)	97%	0.83 (0.76, 0.9)

specialized cameras. The limitation is that evaluation is restricted to cortical and posterior subcapsular opacities within the central 5 mm diameter of the lens. Nuclear opacification is not assessed in this system. Therefore, a complete assessment of cataract necessarily would include

clinician assessment of nuclear sclerosis and peripheral lens opacities in combination with this methodology. This simplified methodology of lens assessment can be useful in large scale studies where it is important to find group differences.

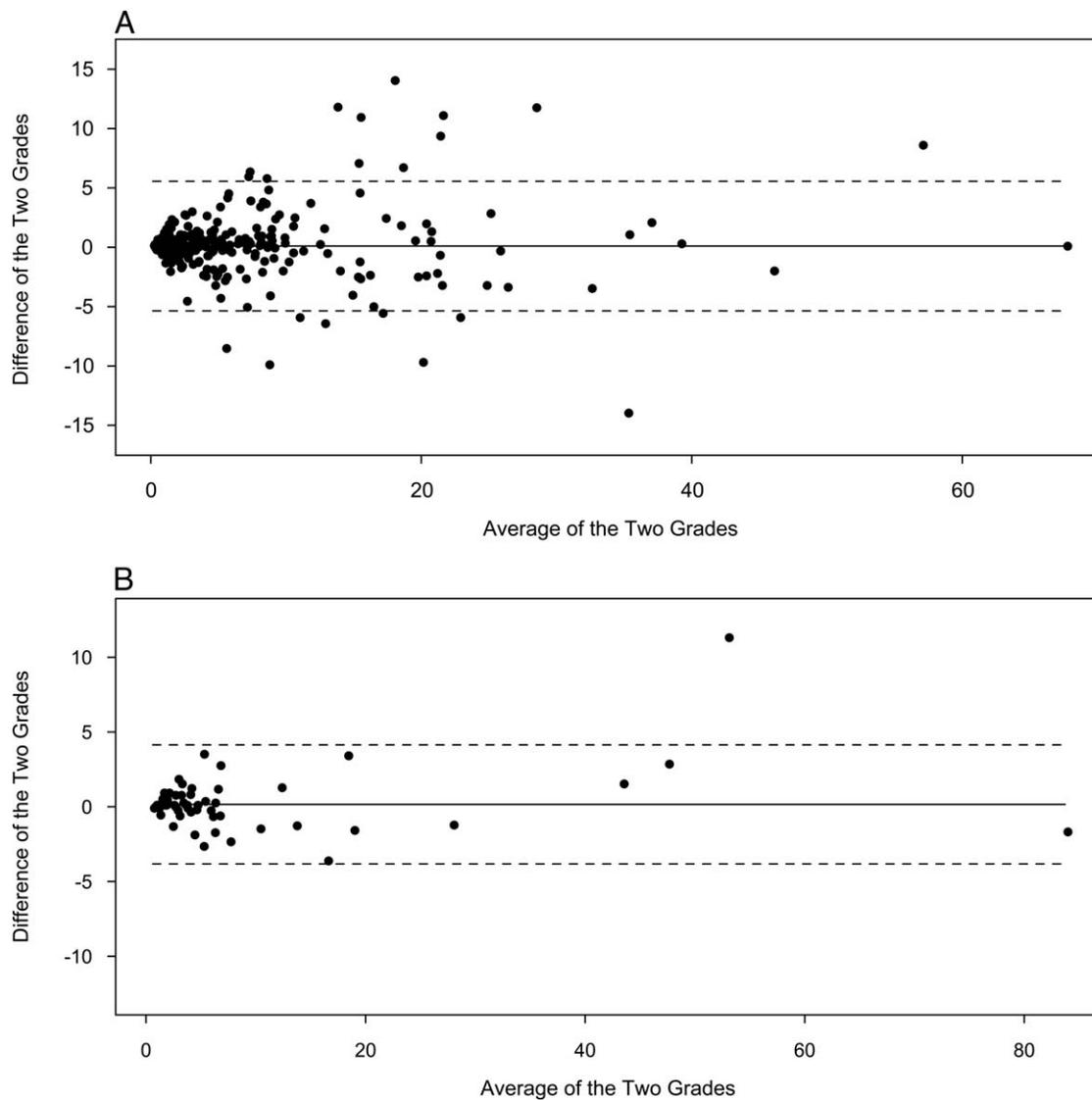


FIGURE 2. (A) Bland Altman plots for contemporaneous reproducibility of percentage involvement of central zone by cortical cataract showing a mean difference (solid line) of 0.1%, and Bland Altman limits (dashed lines) of +5.56 and -5.36. (B) Bland Altman plots for contemporaneous reproducibility of percentage involvement of central zone by posterior subcapsular cataract showing a mean difference (solid line) of 0.16%, and Bland Altman limits (dashed lines) of +4.14 and -3.82.

The AREDS included standardization of image capture across 11 clinical sites in the United States. Image capture using specialized cameras, such as the Neitz retroillumination camera (for assessing cortical and PSC opacities; Kowa Optical) and modified slit-lamp cameras (for assessing nuclear sclerosis), was possible for the relatively small number of sites involved in the AREDS. The large number of clinical sites in AREDS2 (87 sites) could not be equipped with the specialized cameras that had been used in AREDS. However, AREDS2 did require certification of fundus cameras for acquiring retinal images for evaluation of AMD. Fundus reflex images also could be acquired during the same sitting using the same camera. As a result, photographic evaluation of cortical and PSC opacities was performed using fundus reflex images in AREDS2. To accommodate the lack of photographic documentation for nuclear sclerosis, which is a major cause of cataract surgery, progression to cataract surgery also was used as a clinical outcome.

The secondary outcome of the study included progression of lens opacities within the central 5 mm diameter of the lens. At baseline, cortical opacities were visible in approximately 40% of eyes and posterior subcapsular opacities in 8%.¹⁵ For the area between 5 and 8 mm diameter, 78% of eyes did not have a significant cataract, 19% had a significant cataract, and 3% were ungradable for this feature.

A comparison between gradings with film-based Neitz retroillumination images and fundus reflex images captured with a Zeiss fundus camera (Carl Zeiss Meditech, Dublin, CA) in AREDS showed comparable areas of measurement for cortical opacities (Ansary SE, et al. *IOVS* 1997;38:ARVO Abstract 878) PSC opacities visible in Neitz retroillumination photographs were visible only approximately half of the time in film-based fundus reflex photographs. A similar evaluation using digitized fundus reflex images showed a 100% agreement with Neitz retroillumination images for presence/absence of PSC (Ansary S, Armstrong J, Osterby K, et al., unpublished data, 1997). The area of PSC opacity was smaller in the fundus reflex images, with a mean difference of 6% (1.18 mm²). The images used for this comparison were not optimized and evaluated in red channel, which potentially could improve detection of opacities. In addition, the longitudinal system used in AREDS2 can be expected to reduce the noise due to image quality and subjectivity of evaluation. Using longitudinal evaluation rather than single image independent grading reduced apparent PSC regression by more than half compared to independent grading (i.e., no access to previous visit or grading data; Pak JW, et al. *IOVS* 2008;49:ARVO E-Abstract 1924).

Nearly all of the images were digital submissions in AREDS2, compared to AREDS where all submissions were film images. To standardize the image evaluations in AREDS2, the few film images submitted were digitized at the reading center and displayed in the same software environment as the natively digital images. In film images, interclinic variability in tonal characteristics was controlled by specifying the film emulsions. In digital images, the color balance is affected by the type of camera and the capture settings of the camera software. Optimizing the images for evaluation in the red channel helped standardize the balance, and provided the most contrast between opacities and the background. A comparison between film images and digital red reflex images showed that the cortical opacity evaluations are comparable. The study did not have sufficient numbers of PSC to draw conclusions (Harris S, et al. *IOVS* 2007;46:ARVO E-Abstract 3170).

The agreements for cortical and PSC opacities were better in the contemporaneous than the temporal drift exercises. The study eligibility required adequate photo quality of retinal photographs as assessed by the reading center. To accommodate this, eyes with visually important cataracts at baseline

were excluded. The temporal drift images were selected from the baseline dataset and, therefore, had minimal lens opacities. The contemporaneous exercise sample was selected throughout the course of the study and included eyes with significant lens opacities, which may explain the apparently better reproducibility. The reproducibility for cortical opacities is comparable to that in AREDS (0.95 in AREDS2 and 0.94 in AREDS) and is improved for PSC (0.99 in AREDS2 and 0.82 in AREDS), perhaps because of digital grading.

To conclude, fundus reflex images can be used to provide reproducible evaluation of cortical and PSC opacities in large multicenter trials using widely available fundus cameras. This methodology, however, limits evaluation to cortical and posterior subcapsular cataracts within the central 5 mm of the lens, and does not include evaluation of nuclear opacities. The grading procedure adapted from AREDS for centralized evaluation of digital images for lens opacities can be used for evaluation and is adequately reproducible.

Acknowledgments

Supported by NIH Grant U10-EY013018 (RD) and intramural program funds from NIH Contract No. HHS-N-260-2005-00007-C, and Administrative Data Base (ADB) Contract No. N01-EY-5-0007.

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

References

1. Resnikoff S, Pascolini D, Etya'ale D, et al. Global data on visual impairment in the year 2002. *Bull World Health Organ*. 2004; 82:844-851.
2. Congdon N, Vingerling JR, Klein BE, et al. Prevalence of cataract and pseudophakia/aphakia among adults in the United States. *Arch Ophthalmol*. 2004;122:487-494.
3. Robin AL, Thulasiraj RD. Cataract blindness. *Arch Ophthalmol*. 2012;130:1452-1455.
4. West SK, Taylor HR. The detection and grading of cataract: an epidemiologic perspective. *Surv Ophthalmol*. 1986;31:175-184.
5. Klein BE, Klein R, Linton KL, Magli YL, Neider MW. Assessment of cataracts from photographs in the Beaver Dam Eye Study. *Ophthalmology*. 1990;97:1428-1433.
6. Age-Related Eye Disease Study Research Group. The age-related eye disease study (AREDS) system for classifying cataracts from photographs: AREDS report No. 4. *Am J Ophthalmol*. 2001;131:167-175.
7. Wong WL, Li X, Li J, et al. Cataract conversion assessment using lens opacity classification system III and Wisconsin cataract grading system. *Invest Ophthalmol Vis Sci*. 2013;54: 280-287.
8. Chylack LTJ, Wolfe JK, Friend J, et al. Validation of methods for the assessment of cataract progression in the Roche European-American Anticataract Trial (REACT). *Ophthalmic Epidemiol*. 1995;2:59-75.
9. Thylefors B, Chylack LT Jr, Konyama K, et al. A simplified cataract grading system. *Ophthalmic Epidemiol*. 2002;9:83-95.
10. Taylor HR, Lee JA, Wang F, Muñoz B. A comparison of two photographic systems for grading cataract. *Invest Ophthalmol Vis Sci*. 1991;32:529-532.
11. Chew EY, Kim J, Sperduto RD, et al. Evaluation of the age-related eye disease study clinical lens grading system AREDS report No. 31. *Ophthalmology*. 2010;117:2112-2119.
12. Age-Related Eye Disease Study Research Group. A randomized, placebo-controlled, clinical trial of high-dose supplementation

- with vitamins C and E and beta carotene for age-related cataract and vision loss: AREDS report No. 9. *Arch Ophthalmol*. 2001;119:1439-1452.
13. 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.
 14. FDA. *FDA Guidance for Industry—Standards for Clinical Trial Imaging Endpoints*. Rockville, MD: Food and Drug Administration; 2011:1-26.
 15. Chew E, SanGiovanni JP, Ferris FL, et al. Lutein/zeaxanthin for the treatment of age-related cataract: AREDS2 Randomized Trial Report No. 4. *JAMA Ophthalmol*. 2013;131:843-850.