tvst

DOI: 10.1167/tvst.6.2.3

Article

Automated Brightness and Contrast Adjustment of Color Fundus Photographs for the Grading of Age-Related Macular Degeneration

Edem Tsikata[1,2*], Inês Laíns[1,3,4,5,6*], João Gil[3,4,5], Marco Marques[3,5], Kelsey Brown[1], Tânia Mesquita[5], Pedro Melo[5], Maria da Luz Cachulo[3,4,5], Ivana K. Kim[1,6], Demetrios Vavvas[1,6], Joaquim N. Murta[3,4], John B. Miller[1,6], Rufino Silva[3,4,5], Joan W. Miller[1,6], Teresa C. Chen[1,2], and Deeba Husain[1,6]

[1] Massachusetts Eye and Ear, Department of Ophthalmology, Harvard Medical School, Boston, MA, USA
[2] Glaucoma Service of Massachusetts Eye and Ear, Harvard Medical School, Boston, MA, USA
[3] University of Coimbra, Faculty of Medicine, University of Coimbra, Coimbra, Portugal
[4] Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal
[5] Association for Biomedical Research and Innovation on Light and Image, Coimbra, Portugal
[6] Retina Service of Massachusetts Eye and Ear, Harvard Medical School, Boston, MA, USA

Correspondence: Deeba Husain, Massachusetts Eye and Ear, Harvard Medical School, 243 Charles Street, 12th Floor, Boston, MA, USA. e-mail: Deeba_Husain@meei.harvard.edu

Received: 25 November 2016
Accepted: 22 January 2017
Published: 13 March 2017

Keywords: image analysis; age-related macular degeneration; automated optimization

Citation: Tsikata E, Laíns I, Gil J, Marques M, Brown K, Mesquita T, Melo P, da Luz Cachulo M, Kim IK, Vavvas D, Murta JN, Miller JB, Silva R, Miller JW, Chen TC, Husain D. Automated brightness and contrast adjustment of color fundus photographs for the grading of age-related macular degeneration. Trans Vis Sci Tech. 2017;6(2):3, doi:10.1167/tvst.6.2.3
Copyright 2017 The Authors

Purpose: The purpose of this study was to develop an algorithm to automatically standardize the brightness, contrast, and color balance of digital color fundus photographs used to grade AMD and to validate this algorithm by determining the effects of the standardization on image quality and disease grading.

Methods: Seven-field color photographs of patients (>50 years) with any stage of AMD and a control group were acquired at two study sites, with either the Topcon TRC-50DX or Zeiss FF-450 Plus cameras. Field 2 photographs were analyzed. Pixel brightness values in the red, green, and blue (RGB) color channels were adjusted in custom-built software to make the mean brightness and contrast of the images equal to optimal values determined by the Age-Related Eye Disease Study (AREDS) 2 group.

Results: Color photographs of 370 eyes were analyzed. We found a wide range of brightness and contrast values in the images at baseline, even for those taken with the same camera. After processing, image brightness variability (brightest image–dimmest image in a color channel) was reduced 69-fold, 62-fold, and 96-fold for the RGB channels. Contrast variability was reduced 6-fold, 8-fold, and 13-fold, respectively, after adjustment. Of the 23% images considered nongradable before adjustment, only 5.7% remained nongradable.

Conclusions: This automated software enables rapid and accurate standardization of color photographs for AMD grading.

Translational Relevance: This work offers the potential to be the future of assessing and grading AMD from photos for clinical research and teleimaging.

Introduction

Age-related macular degeneration is the leading cause of adult blindness in developed countries and the third leading cause worldwide.[1,2] Diagnosis of AMD is based on the presence of the characteristic fundus abnormalities of this condition: drusen and focal pigmentation changes in its early stages, which may progress to retinal atrophy and choroidal neovascularization in the advanced forms of the disease.[3,4] At this time the gold standard for classification of AMD for clinical research purposes is detection of these features on film color photographs and high-resolution digital color images of the macula.[4–8] Despite advances in imaging, including the widespread use of optical coherence tomography (OCT), all currently validated AMD grading systems

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

are based on fundus color photography. Therefore, there is a great need to standardize the color images used for clinical studies.

Fundus images with optimal exposure and color balance are essential for accurate evaluation of retinal features.[9–11] Suboptimal images can lead to inaccurate categorization of drusen and pigment abnormalities,[12,13] which may impede the diagnosis or correct staging of AMD. Fundus photograph quality has several contributing factors, including patient factors (media opacities and pupil size),[9] camera properties (sensor resolution and chromatic response), image-processing software, and photographer techniques (exposure settings and focus). Cameras with the same optics and sensors can produce varying output depending on the color profiles specified by the camera manufacturer or photographers. With so many pitfalls, substandard images have been reported to be as high as 20% in clinical studies.[11,14] Even with best practice, acquisition of consistently high-quality images cannot be guaranteed at the time of capture. Post hoc standardization[15] is therefore necessary for AMD grading.[16]

Previous AMD studies[6,16] assessed digital photographic quality and implemented approaches to improve color balance, brightness, and contrast. Specifically, Hubbard et al. in the Age-Related Eye Disease Study (AREDS) 2 trial, developed an enhancement procedure based on adjustment of the three-color luminance histograms.[16] Though this approach improved the contrast and brightness of the fundus images, it relied on manual procedures applied to individual photographs.

The aim of this study was to develop and test software to automatically standardize the brightness, contrast, and color balance of digital fundus photographs. The effects of the standardization procedure on image quality and the staging of AMD were also investigated. Using the optimal image parameters established by the AREDS 2 group,[16] digital images from two types of cameras were batch-processed to standardize their color characteristics prior to AMD grading.

Methods

This work is part of a prospective, multicenter study on AMD biomarkers. The software was developed at the Massachusetts Eye and Ear (MEE) (Boston, USA). The program was clinically tested at MEE (designated site A) and at the Faculty of Medicine of the University of Coimbra (FMUC) in collaboration with the Association for Innovation and Biomedical Research on Light and Image (AIBILI) and the Coimbra University Hospital (designated site B).

The study was approved by the institutional review boards of MEE, FMUC, and AIBILI and by the Portuguese National Data Protection Committee. All participants provided written informed consent. The study was conducted in accordance with the Health Insurance Portability and Accountability Act requirements and the tenets of the Declaration of Helsinki.

Study Population and Procedures

From January 2015 to July 2016, we recruited patients with a diagnosis of AMD and a control group of subjects without any evidence of the disease in both eyes, aged 50 years or older. For both, the exclusion criteria included diagnosis of any other vitreoretinal disease, active uveitis or ocular infection, significant media opacities that precluded the observation of the ocular fundus, refractive error equal to or greater than 6 diopters of spherical equivalent, past history of retinal surgery, history of any ocular surgery or intraocular procedure (such as laser and intraocular injections) within the 90 days prior to enrollment, and diagnosis of diabetes mellitus, with or without concomitant diabetic retinopathy (due to the remaining study purposes). At site A, participants were consecutively recruited at the Retina Service and the Comprehensive Ophthalmology and Optometry Services when they came for their regular appointments. The study population in site B was derived from a population-based cohort study.[17] All subjects with an established diagnosis of any stage of AMD were invited to participate. Subjects without signs of AMD in the prior evaluation[17] were also invited and were included as controls if they remained without the disease (see AMD Diagnosis and Staging). Those who presented with AMD at the time of the current evaluation were also considered but included in the AMD study group.

All participants underwent a comprehensive eye exam, including measurement of best-corrected visual acuity, current refraction, intraocular pressure, biomicroscopy, and dilated fundus exams. Nonstereoscopic, seven-field fundus photographs were obtained with one of two types of cameras: Topcon TRC-50DX (Topcon Corporation, Tokyo, Japan), with a 35-degree field of view, or Zeiss FF-450 Plus (Carl Zeiss Meditec Inc., Dublin, CA) with a 30-degree field of view. These cameras used charge coupled device (CCD) sensors: Topcon cameras–Pike 11MP CCD

(site A) and Nikon D2H (site B); Zeiss camera–Escalon E5 (site A).

AMD Diagnosis and Staging

For AMD diagnosis and staging, two independent experienced graders analyzed all field 2 CFP (IL, JG), according to the AREDS classification system.[18,19] In cases of disagreement, a senior author (RS or DH) established the final categorization. Images taken with Topcon cameras were evaluated with IMAGEnet 2000 software (version 2.56; Topcon Medical Systems), and those obtained with a Zeiss camera were observed using VISUPAC (version 4.5.1; Carl Zeiss Meditec). We adopted the most recent AREDS 2 definitions,[19] namely that the standard disc diameter equals 1800 μm, which affects the size of the Early Treatment Diabetic Retinopathy Study grid and of the standard drusen circles, and that geographic atrophy (GA) is present if the lesion has a diameter equal or superior to 433 μm (AREDS circle I-2) and at least two of the following features are present: absence of RPE pigment, circular shape, or sharp margins (thus meaning that the involvement of the central fovea is not a requirement). Therefore, briefly, we established the following groups[18,19]: controls–presence of drusen of maximum size $<$ circle C0 and total area $<$ C1; early AMD—drusen of maximum size $\geq$ C0 but $<$ C1 or presence of AMD characteristic pigment abnormalities in the inner or central subfields; intermediate AMD—presence of drusen maximum size $\geq$ C1 or of drusen maximum size $\geq$ C0 if the total area occupied is $>$ I2 for soft indistinct drusen and $>$ O2 for soft distinct drusen; and late AMD—presence of GA according to the criteria described above or evidence of neovascular AMD.

Image Processing

In this study, image brightness is based on the pixel intensities for each color channel. Contrast quantifies the variation in the pixel intensities and is measured as four times the standard deviation of the pixel intensity values in each color channel. Color balance is calculated as the ratio of the green and blue channel brightness values to the brightness of the red channel (the brightest channel).

The software program was written in the C++ language, and the Open Computer Vision library (OpenCV version 2.4.3; Willow Garage, Menlo Park, CA) was used for image processing. The graphical user interface was constructed in the Qt application development framework (Qt 4.8; The QT Company, Oslo, Norway). To analyze each digital photograph, the retinal image was automatically extracted and the brightness and contrast in the red, green, and blue (RGB) color channels measured. In this report, image contrast is synonymous with the standard deviation of the brightness, and the span of a color curve is defined as four times the standard deviation. This definition of the span encompasses 95.5% of the brightness values in the distribution (2 SD above and below the mean). Topcon images had a pure black border, which could readily be distinguished from the retinal picture. For Zeiss images, which had a border that was not completely black, a circular mask of radius 1226 pixels was used to extract the retinal picture. Within the region identified as the retinal image, the pixel values were read by the software, and the mean and standard deviation of the brightness of the RGB channels were calculated.

To enhance image contrast, contrast ratios were calculated by dividing the ideal curve spans by the measured spans. Pixel brightness values were then multiplied by the appropriate contrast ratios to achieve the target image contrast in each color channel. To achieve the target brightness values, constant terms were added to the scaled pixel values to make the overall image brightness equal to the targets.

A mathematical description of these concepts is presented below.

$$\alpha = \Delta_{\text{ideal}}/4\sigma,$$

where α is the contrast ratio,
Δ_{ideal} is the ideal brightness span, and
σ is the standard deviation of the color channel.

Capitalized symbols refer to global image properties. Lowercase symbols denote pixel values.

$$x_{\text{final}} = \alpha x_{\text{initial}} + (X_{\text{ideal}} - \alpha X_{\text{mean}}),$$

where X_{ideal} is the ideal image brightness,
X_{mean} is the mean brightness of the original image,
x_{initial} is an initial pixel brightness, and
x_{final} is the transformed pixel brightness.

Statistical Analysis of Data

We collected and managed study data using REDCap electronic data capture tools hosted at MEE.[20] Mean, standard deviation, and ranges were

Table 1. Demographic Characteristics of the Study Population, Organized by Study Site and Camera Used

Study Site	Site A		Site B	
Camera (sensor)	Topcon TRC-50DX (Pike 11 MP CCD), $n = 114$	Zeiss FF-450 (Escalon E5 CCD), $n = 33$	Topcon TRC-50DX (Nikon D2H CCD), $n = 223$	Total, $n = 370$
Image format and number (%)				
TIFF	102 (89.5)	0 (0)	0 (0)	102 (27.6)
JPEG	11 (9.6)	0 (0)	223 (100)	234 (63.2)
PNG	1 (0.9)	33 (100)	0 (0)	34 (9.2)
Age, y, mean ± SD	71.4 ± 7.8	76.9 ± 6.5	73.1 ± 8.1	72.9 ± 8.0
Gender, number (%)				
Female	71 (62.3)	23 (69.7)	151 (67.7)	245 (66.2)
Male	43 (37.7)	10 (30.3)	72 (32.3)	136 (33.8)
Race and number (%)				
White	102 (89.5)	32 (97.0)	221 (99.1)	355 (96.0)
Black	0 (0)	0 (0)	2 (0.9)	2 (0.5)
Other	2 (1.7)	0 (0)	0 (0)	2 (0.5)
Unknown	10 (8.8)	1 (3.0)	0 (0)	11 (3.0)
Laterality, number (%)				
OD	57 (50.0)	19 (5.1)	114 (30.8)	190 (51.4)
OS	57 (50.0)	14 (3.8)	109 (29.5)	180 (48.6)
Lens and number (%)				
Phakic	92 (25.0)	22 (57.6)	197 (88.3)	311 (84.1)
Pseudophakic	22 (5.9)	11 (32.4)	26 (11.7)	59 (15.9)
AMD stage and number (%)				
Control	26 (22.8)	0 (0.0)	60 (26.9)	86 (23.2)
Early AMD	14 (12.3)	4 (12.1)	47 (21.1)	65 (17.6)
Intermediate AMD	51 (44.7)	12 (36.4)	89 (39.9)	152 (41.1)
Late AMD	20 (17.5)	15 (45.5)	11 (4.9)	46 (12.4)
Not gradable*	3 (2.6)	2 (6.1)	16 (7.2)	21 (5.7)

Unknown data for race refers to included participants who decided not to provide this information. TIFF, Tagged Image File Format; JPEG, Joint Photographic Experts Group; PNG, Portable Network Graphics; SD, standard deviation; OD, right eye; OS, left eye; AMD, Age-related Macular Degeneration.

* Images not gradable even after adjustment with the developed software.

used to summarize image data, and the characteristics of different types of images and cameras were compared with *t*-tests. Two independent experienced graders masked to the diagnosis (IL and JG) evaluated all images before adjustment and classified them as gradable (if the grader felt able to stage AMD with confidence) or nongradable (if this was not the case). After software processing, the process was repeated for the originally nongradable images. All statistical analysis was performed using Stata (version 12.1; StataCorp LP, College Station, TX) and P values < 0.05 were considered statistically significant.

Results

Imaging Properties of the Entire Sample Set before and after Adjustment

We included field 2 color fundus photographs of 370 eyes of 370 subjects with a mean age of 72.9 ± 8.0 years old. The laterality of the eyes was randomly selected using the RAND function in Microsoft Excel (Microsoft Corporation, Redmond, WA). Sixty-six percent of the subjects ($n = 245$) were female. Table 1 presents the demographic characteristics of the

Table 2. Brightness, Standard Deviation of Brightness and Color Balance Ratios for the Analyzed 370 Images before and after Automated Image Enhancement

	Before	After
Brightness parameters		
Red brightness		
Mean ± SD	155.0 ± 40.5	191.7 ± 0.4
Range	52.5–231.1	190.0–192.6
Green brightness		
Green mean ± SD	72.3 ± 22.1	96.1 ± 0.2
Range	15.7–115.5	95.2–96.8
Blue brightness		
Mean ± SD	25.5 ± 12.2	32.0 ± 0.1
Range	0.07–76.8	31.6–32.4
Contrast parameters		
SD red brightness		
Mean ± SD	25.4 ± 6.9	30.7 ± 1.2
Range	11.4–54.5	24.8–32.6
SD green brightness		
Mean ± SD	16.3 ± 5.2	31.1 ± 0.9
Range	4.6–42.8	27.1–32.1
SD blue brightness		
Mean ± SD	7.5 ± 3.3	8.0 ± 0.1
Range	0.9–20.7	6.8–8.3
YCrCb colorspace parameters		
Intensity Y		
Mean ± SD	91.7 ± 24.9	117.3 ± 0.2
Range	28.3–135.8	116.6–118.0
Chrominance red difference (Cr)		
Mean ± SD	173.1 ± 13.2	181.0 ± 0.2
Range	138.2–202.8	180.2–181.3
Chrominance blue difference (Cb)		
Mean ± SD	90.6 ± 11.7	79.9 ± 0.1
Range	64.7–116.3	79.5–80.3
Color balance parameters		
Green/Red ratio		
Mean ± SD	0.466 ± 0.084	0.501 ± 0.001
Range	0.209–0.788	0.498–0.505
Blue/red ratio		
Mean ± SD	0.167 ± 0.082	0.167 ± 0.0005
Range	0.000–0.549	0.165–0.169

included eyes, as well as their AMD staging, organized by study site and camera type.

Table 2 summarizes the image properties before and after software processing. The brightness and contrast values in each channel were larger after adjustment than before, and they were closer to the target values, the ideal AREDS 2 (red brightness = 192, green brightness = 96, blue brightness = 32; red span = 128, green span = 128, blue span = 32), after processing. The range of brightness values (brightest image to dimmest image) decreased 69-fold, 62-fold, and 96-fold for the RGB channels, respectively. The contrast ranges also decreased 6-fold, 8-fold, and 13-fold for the RGB channels, respectively. The range of green/red color balance values decreased 83-fold after adjustment, and the range of blue/red color balance values decreased 137-fold.

Though brightness and contrast were standardized in the RGB color space, the image parameters in the YCrCb color space also exhibited reduced variance after adjustment (Table 2). The YCrCb space describes the pixel properties using an intensity or luma coordinate (Y) and two chrominance coordinates: Cr, for the red difference, and Cb, for the blue difference.[21] YCrCb parameters are derived from the RGB values using linear equations specified in the International Telecommunication Union–Radiocommunication standard ITU-R 601.7.[21] The range of values decreased 77-fold, 59-fold, and 65-fold for the intensity, red difference chrominance, and blue difference chrominance, respectively.

Variations among Cameras and within the Same Camera

Baseline images from the three cameras were underexposed relative to the ideal AREDS 2 settings (Fig. 1, top panel). Automated enhancement increased the overall brightness of the images (Fig. 1, middle panel). The histograms confirmed underexposure, and the effect of the software enhancement, which was stretching of the brightness curves to improve the contrast and displacing them to the right to correct underexposure (Fig. 1 bottom panel).

Quantitative analysis confirmed initial underexposure relative to the AREDS 2 ideal values (red brightness = 192, green brightness = 96, blue brightness = 32; red span = 128, green span = 128, blue span = 32) for images from all three cameras (Table 3). Additionally, images acquired with the Topcon camera at site A (Pike 11MP sensor) were significantly underexposed compared to Topcon camera images from site B (Nikon D2H) (red: 115.2 vs. 175.8; green: 48.9 vs. 83.8; blue: 21.6 vs. 28.8; $P < 0.001$ for all). At the same site (A), Topcon camera images were also significantly underexposed relative to those taken with the Zeiss camera in the red and green channels but had comparable exposure in the blue channel (red: 115.2 vs. 132.7, $P = 0.013$; green: 48.9 vs. 70.2, $P < 0.001$; blue: 21.6 vs. 20.4,

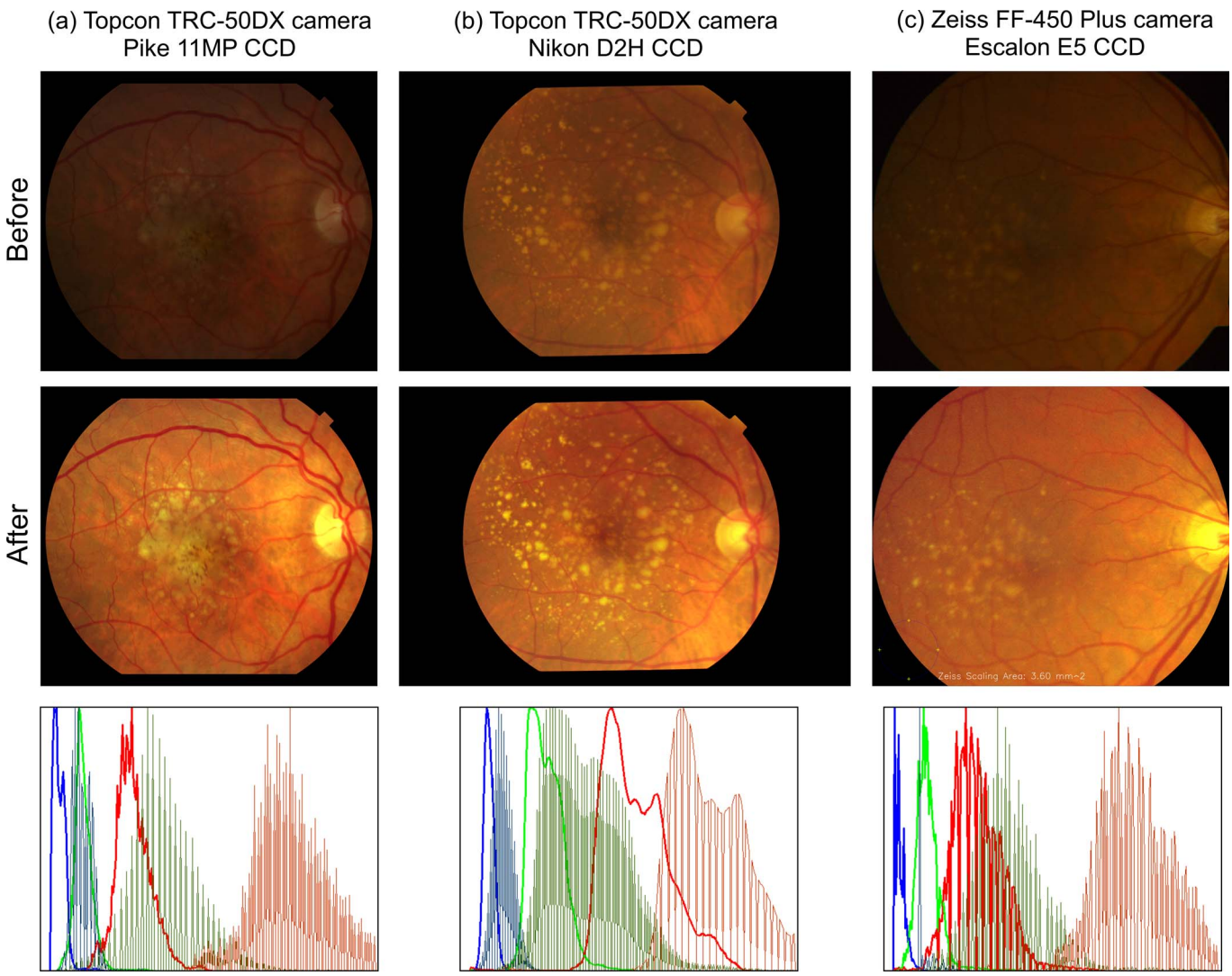

Figure 1. Figure 1a and b (*top panel*) are unmodified fundus photographs acquired with Topcon cameras. Figure 1a was acquired with a Topcon TRC-50DX camera with a Pike 11MP CCD sensor and Figure 1b with a Topcon TRC-50DX camera with a Nikon D2H CCD sensor. Figure 1c was obtained with a Zeiss FF-450 Plus camera with an Escalon E5 CCD sensor. Images in the *middle panel* were produced by processing with the automated enhancement software. In the *bottom row*, the bright *solid lines* represent the color histograms of the original images and the *thinner lines* the histograms of the standardized images.

$P = 0.697$). Before processing, the green to red color balance ratios were also significantly different among the three types of cameras ($P < 0.01$). After adjustment, the green to red color balance ratios for all three cameras were equal to 0.501 ± 0.001. Despite the similarity in the measured values after adjustment, the differences were statistically significant ($P < 0.001$ for all comparisons). The blue to red color balance ratios were significantly different only for the Topcon Pike versus Topcon Nikon D2H sensors before adjustment (0.181 vs. 0.167, $P = 0.024$). After adjustment, the blue to red color balance ratios of images from the two Topcon cameras were similar ($P = 0.424$).

Before processing, the brightness and contrast of the photos from the same camera exhibited variability. Figure 2 illustrates the variation in the brightness of images from a Topcon camera (Topcon TRC-50DX, Nikon D2H sensor, site B, $n = 223$ images). The brightness values were generally below the AREDS 2 target values: 175.8 ± 21.1 versus 192, 83.9 ± 11.0 versus 96 and 28.8 ± 8.0 versus 32 for the RGB channels. After adjustment, the brightness values were consistent with the optimal AREDS 2 values: 191.7 ± 0.3 versus 192, 96.1 ± 0.2 versus 96 and 32.0 ± 0.1 versus 32 for the RGB channels, respectively.

Table 3. Mean Brightness and Color Balance Ratios for Three Types of CCD Sensors Used in This Study

	Before Adjustment		
Camera (sensor type) [number of images]	Topcon (Pike 11MP) [n = 114]	Topcon (Nikon D2H) [n = 223]	Zeiss (Escalon E5 CCD) [n = 33]
Red brightness			
Mean ± SD	115.2 ± 32.7	175.8 ± 21.1	132.7 ± 42.0
Minimum–maximum	52.5–190.8	108.2–231.1	53.6–195.1
Green brightness			
Mean ± SD	48.9 ± 18.2	83.8 ± 11.0	70.2 ± 22.2
Minimum–maximum	15.7–102.1	54.9–115.5	28.9–110.1
Blue brightness			
Mean ± SD	21.6 ± 14.5	28.8 ± 8.0	20.4 ± 18.2
Minimum-–maximum	0.3–64.0	10.6–55.6	0.1–76.8
Green/red ratio			
Mean ± SD	0.424 ± 0.098	0.480 ± 0.057	0.539 ± 0.095
Minimum–maximum	0.209–0.706	0.323–0.716	0.384–0.788
Blue/red ratio			
Mean ± SD	0.181 ± 0.101	0.167 ± 0.053	0.163 ± 0.136
Minimum–maximum	0.004–0.433	0.050–0.434	0.000–0.549

Phakic and Pseudophakic Eyes

For all images obtained with a Topcon camera (n = 337), the average brightness in the red channel was significantly higher for subjects with natural lenses (n = 289) than for those with intraocular lenses (n = 48) (159.2 vs. 145.3, P = 0.034), but comparable in the green channel (73.3 vs. 67.8, P = 0.130). The average

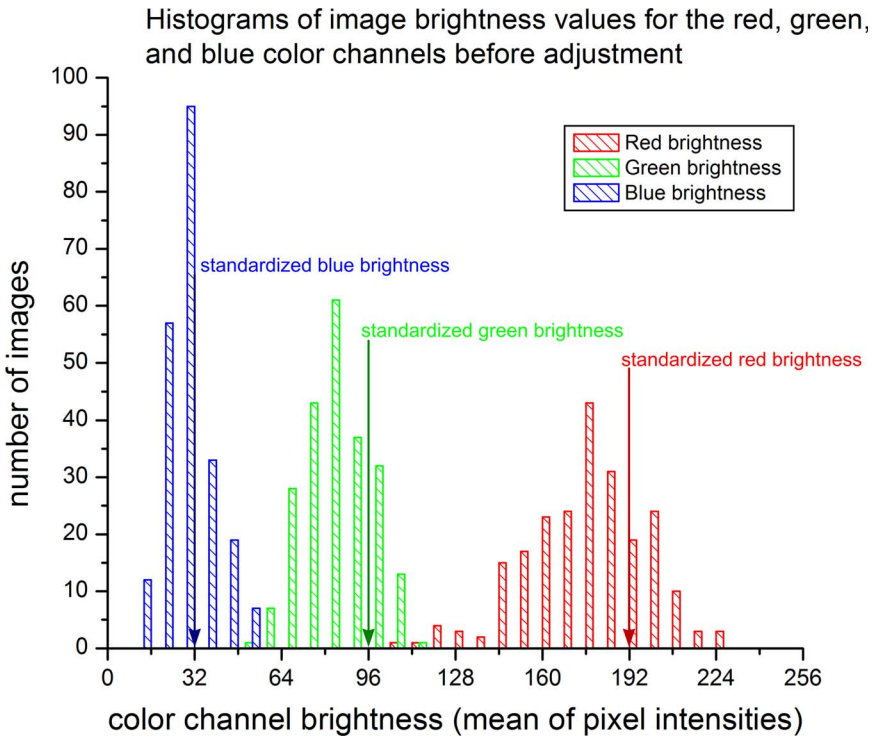

Figure 2. Comparison of the mean brightness of Topcon color photographs in the RGB channels before and after adjustment (223 images, Topcon TRC-50DX, D2H sensor). The target brightness values for the RGB channels were 192, 96, and 32.

Table 3. Extended

Camera (sensor type) [number of images]	After Adjustment: Topcon (Pike 11MP) [$n = 114$]	Topcon (Nikon D2H) [$n = 223$]	Zeiss (Escalon E5 CCD) [$n = 33$]
Red brightness			
Mean ± SD	191.5 ± 0.4	191.7 ± 0.4	191.6 ± 0.3
Minimum–maximum	190.0–192.0	190.6–192.6	190.5–192.2
Green brightness			
Mean ± SD	95.9 ± 0.1	96.1 ± 0.2	96.0 ± 0.1
Minimum–maximum	95.2–96.1	95.7–96.8	95.8–96.4
Blue brightness			
Mean ± SD	32.0 ± 0.1	32.0 ± 0.1	32.0 ± 0.1
Minimum-–maximum	31.6–32.2	31.7–32.4	31.6–32.3
Green/red ratio			
Mean ± SD	0.501 ± 0.001	0.501 ± 0.001	0.501 ± 0.001
Minimum–maximum	0.498–0.505	0.499–0.504	0.499–0.504
Blue/red ratio			
Mean ± SD	0.167 ± 0.001	0.167 ± 0.0004	0.167 ± 0.001
Minimum–maximum	0.165–0.169	0.165–0.169	0.165–0.168

brightness in the blue channel was significantly lower for phakic than for pseudophakic eyes (24.6 vs. 34.0, $P < 0.001$). After adjustment, no significant differences were observed in any color channel ($P \geq 0.188$) (Table 4). As detailed in the same table, before adjustment, the green to red ratios were comparable between phakic and pseudophakic eyes (0.458 vs. 0.461, $P = 0.341$), but pseudophakic eyes showed significantly increased blue to red color balance ratios compared to phakic eyes (0.231 vs. 0.157, $P < 0.001$). The blue to red color balance ratios were not significantly different after software adjustment (0.167 vs. 0.167, $P = 0.136$).

Figure 3 presents an example of an eye with a natural lens (Fig. 3a) and an eye with an intraocular lens (Fig. 3b). In Figure 3a (top panel), the green to red ratio was 0.418, and the blue to red ratio was 0.173. In Figure 3b (top panel), the green to red ratio was 0.501 and the blue to red ratio was 0.343. The tonal differences in the images principally arose from the higher proportion of blue in Figure 3b. After adjustment, the green to red ratio in Figure 3a (middle panel) was 0.500 and the blue to red ratio was 0.168. The green to red ratio in Figure 3b (middle panel) was 0.500, and the blue to red ratio was 0.167. The visibility of the hyperpigmented zone, Figure 3b (inset), is enhanced in the adjusted image.

Processing Time and Grading Discrepancies

The mean processing time per image was 9.47 ± 2.12 seconds (range: 5.2–15.2 seconds) on a computer with a 2.50-GHz Intel i7 processor and 16 gigabytes of memory. In this time, the adjusted images were generated as new files, and the numerical results and histograms were also automatically created by the software. Of the 370 images evaluated, 85 (23.0%) were considered nongradable before software adjustment. After adjustment, 21 images (5.7%) remained nongradable (Table 1). Poor focusing prevented grading in 20 images. In one image, artifacts created by the correction of extreme underexposure precluded grading. Following image adjustment, 11 subjects initially considered controls were classified as having early AMD after adjustment, and seven patients went from early to intermediate AMD. Figure 4 presents an example. Since these color photos were from patients enrolled by us as part of a bigger cross-sectional study on AMD biomarkers, we have detailed clinical exam and OCT imaging on all. All patients that had discrepancy in grading between preadjustment and postadjustment images were evaluated in detail to confirm that the postadjustment grading was consistent with both the clinical exam and the OCT imaging. This was done to verify that the change was real and not a mere artifact.

Table 4. Brightness Values for the RGB Channels of 48 Subjects with Intraocular Lenses and 289 Phakic Subjects

	Before	After
IOL eye red brightness		
Mean ± SD	145.3 ± 40.9	191.7 ± 0.3
Range	52.5–204.5	190.8–192.2
IOL eye green brightness		
Mean ± SD	67.8 ± 23.2	96.1 ± 0.2
Range	18.8–107.0	95.7–96.5
IOL eye blue brightness		
Mean ± SD	34.0 ± 14.7	32.0 ± 0.1
Range	3.57–60.1	31.7–32.2
IOL eye green/red ratio		
Mean ± SD	0.461 ± 0.081	0.501 ± 0.001
Range	0.298–0.716	0.500–0.503
IOL eye blue/red ratio		
Mean ± SD	0.231 ± 0.087	0.167 ± 0.001
Range	0.068–0.434	0.165–0.169
Phakic eye red brightness		
Mean ± SD	159.2 ± 39.2	191.7 ± 0.4
Range	53.7–231.1	190.0–192.6
Phakic eye green brightness		
Mean ± SD	73.3 ± 21.8	96.1 ± 0.2
Range	15.7–115.5	95.2–96.8
Phakic eye blue brightness		
Mean ± SD	24.6 ± 10.1	32.0 ± 0.1
Range	0.3–64.0	31.6–32.4
Phakic eye green/red ratio		
Mean ± SD	0.458 ± 0.080	0.501 ± 0.001
Range	0.209–0.706	0.498–0.505
Phakic eye blue/bed ratio		
Mean ± SD	0.157 ± 0.067	0.167 ± 0.0004
Range	0.005–0.433	0.165–0.169

Discussion

The current gold standard for AMD detection and grading for clinical research and teleimaging is color fundus photography. There is considerable variability in quality of images based on patient factors and camera characteristics. Uniform quality of images is essential to accurate detection and grading of AMD. Therefore, we designed this study, where we present a cross-sectional analysis of 370 color fundus photographs and demonstrate that their brightness, contrast, and color balance can be automatically standardized to conform to a color model optimized for AMD grading.[16] Using image parameters determined by the AREDS 2 group,[16] a software program modified the brightness values of individual image pixels to make the global image parameters match the ideal values. This software was tested on fundus photographs obtained with three types of camera systems. All cameras presented a wide range of luminance characteristics, but after adjustment, variability in brightness, contrast, and color balance was reduced.

Variation in image brightness decreased more than 50-fold for each color channel after processing, and the means of the brightness were within 0.5 units of the defined targets. The observed enhancements in contrast uniformity were less dramatic, and after processing, the means of the contrast measurements were as much as 1.3 units below their targets. Pixel brightness values are constrained to be integers between 0 and 255, and rounding errors (from the integer representation of the numbers) after enhancement led to the observed deviations from perfection. Furthermore, if the adjustments created pixel values larger than 255, these values saturated at 255.

Regardless of the camera system used, measurements in the RGB or YCrCb color spaces indicated that the photos were consistently underexposed relative to the optimal AREDS 2 values. This was corrected by the program. Images acquired with the Topcon camera at study site B were shown to be closer to the ideal targets before adjustment than images acquired with either the Topcon or Zeiss cameras in site A. The appearance of images is influenced by the CCD sensor used to acquire them, the camera image-processing software, and the color calibration settings specified by the manufacturer or operator. Though the IMAGEnet processing software of the Topcon systems was the same, the cameras in the two sites had different sensors: a Nikon D2H sensor was used in site B, while a Pike sensor was used in the site A Topcon camera. The observed differences might therefore be explained by variations in the color balance of sensors (significantly different for the green to red ratios, P value < 0.001, and blue to red ratios, P value $= 0.024$), but might also reflect the preferences of the photographers.

Color balance ratios determine the tonal characteristics of images. Zeiss images subjectively appeared to be overly yellow, and quantitative analysis confirmed the elevated green to red color balance ratio relative to the Topcon camera before adjustment (Zeiss = 0.539 vs. Topcon site A = 0.424, P value $<$ 0.001, and Topcon site B = 0.480, P value < 0.001).

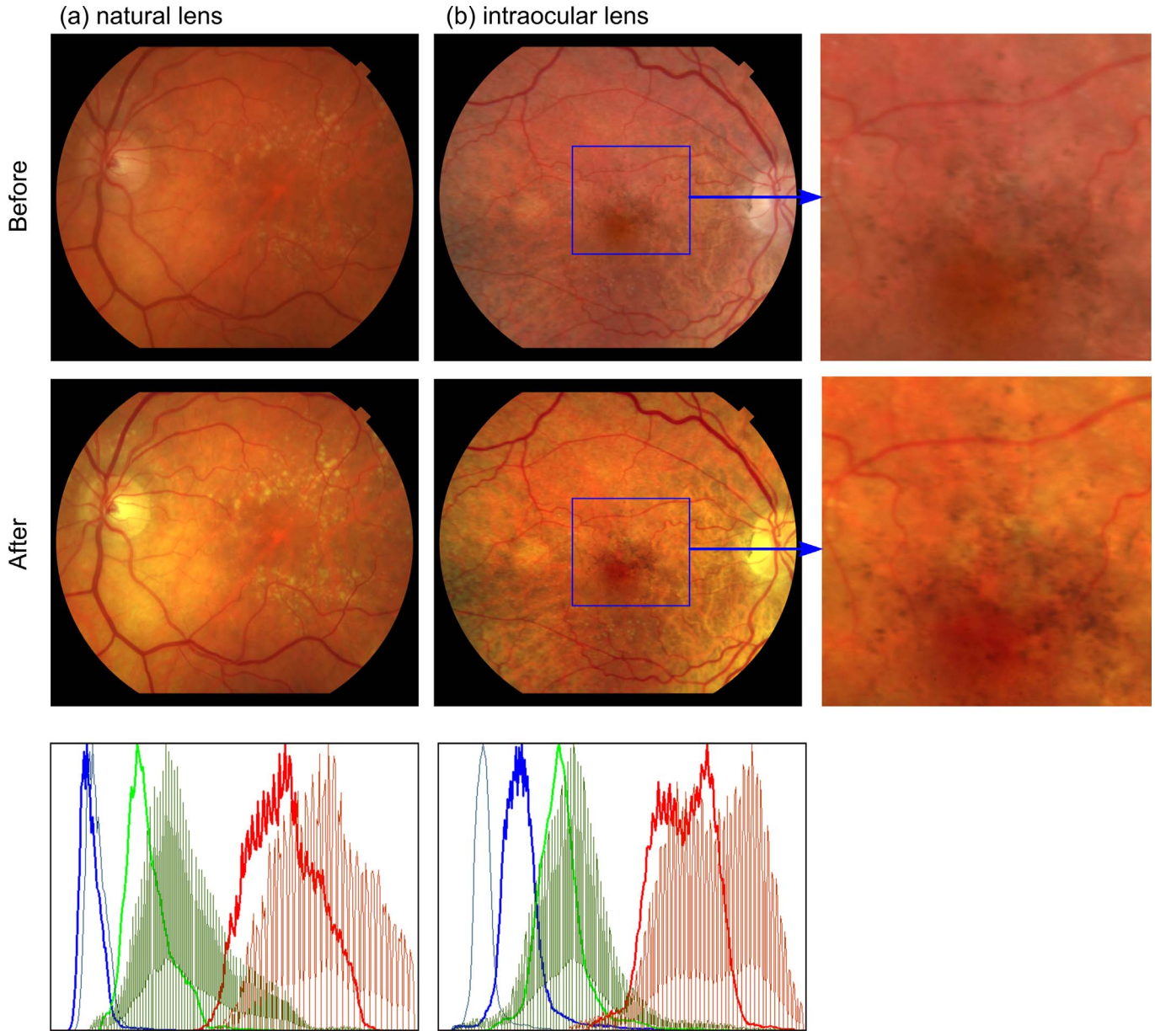

Figure 3. Comparison of the tonal properties of a phakic eye and a pseudophakic eye. Figure 3a (*top panel*) is an unmodified photograph from a subject with a natural lens and Figure 3b (*top panel*) is a photograph from a subject with an intraocular lens. The inset in Figure 3b shows a region of hyperpigmentation. The *middle panels* show the effect of the software enhancement. The histograms on the *bottom* reveal that there is increased transmission of blue light in the eye with the intraocular lens.

After adjustment, the green to red ratios for all cameras were 0.501, and the tonal characteristics of the images were indistinguishable (Fig. 1).

Patient factors, such as lens properties, are known to affect the properties of color fundus photographs. We found that pseudophakic eyes had significantly higher blue to red color balance ratios than phakic eyes (0.231 vs. 0.157, P value < 0.001). This was most likely due to increased transmission of blue light through the intraocular lenses compared to natural lenses.[22–24] To our knowledge, detection of excess blue light in photographs of eyes with intraocular lenses has not been previously reported. The green to red color balance ratios for phakic and pseudophakic eyes measured in this study were statistically similar (0.461 vs. 0.458, P value = 0.341), even though increased green transmission through intraocular lenses has been reported.[23] The effect of an increased blue component in the photos is unclear. The yellowish appearance of drusen is captured well in

Figure 4. Grading of AMD with color fundus photograph enhancement. Before adjustment, this photo was graded as early AMD, and after adjustment, it was changed to intermediate AMD.

the green channel, while the variation in RPE pigment density (hypopigmentation in atrophy and hyperpigmentation in clumping) is captured well in the red channel. The blue channel is a minor component, with a strength three times weaker than the green and six times weaker than the red channels in the ideal AREDS2 color model (32 units in blue vs. 96 in green and 192 in red).[16] However, an increased blue to red color balance ratio in eyes with intraocular lenses may reduce the contrast of hyperpigmentation features (Fig. 3b), leading to inconsistency in grading. The standardization algorithm corrected the high blue to red ratio in the pseudophakic eyes.

The software program can optimize the brightness, contrast, and color balance of substandard images but cannot salvage images where information is lost through poor focus or extremes of under- or overexposure. The fraction of nongradable images recorded in this study before adjustment was 23.0%, which is consistent with previous reports.[11,14] After adjustment, 5.7% of the images remained nongradable. This demonstrates that enhancement of the color characteristics alone can make the majority of substandard images acceptable for AMD grading. Furthermore, adjustment of the images resulted in diagnosis of early AMD in 11 control subjects and change in grading from early to intermediate in seven patients, indicating the usefulness of the software program in diagnosis and grading of AMD.

This study has several limitations. We assessed our data according to the type of sensor of the different cameras used, but we were not able to account for potential variations due to manufacturers and photographers' preferences, which might have affected our results. Additionally, adjusting image parameters to conform to a color model may not be appropriate in all circumstances. Based on images showing all stages of AMD, the AREDS 2 color model is most useful in early to intermediate stage AMD to highlight the appearance of yellow drusen against the retinal background. As mentioned by Hubbard et al.,[16] such adjustment may also distort the appearance of eyes with heavily pigmented choroids. The brightness and contrast parameters adopted in this study were chosen to enhance the visibility of characteristic AMD abnormalities, specifically drusen, hyperpigmentation, and depigmentation. Different settings will be necessary to improve the visibility of abnormalities typical of other pathologies. In all circumstances, it will be crucial to examine the images before and after adjustment to ensure that artifacts do not lead to false-positive detection of disease.

Due to the large proportion of Caucasian subjects (with low amounts of choroidal pigment) in this study, the effect of the AREDS 2 color model on eyes with varying amounts of pigmentation could not be explored. Age-related macular degeneration affects people of Caucasian descent at roughly 9 to 10 times the rate of people of African descent.[2,17,25] The small number of non-Caucasian subjects reflects the disease prevalence in clinical population. While we observed significantly higher blue to red color balance ratios in the 48 pseudophakic eyes, the types of intraocular lenses (blue-blocking or not) were not documented. A further caveat is that the images analyzed in this study had 30- or 35-degree fields of view and exhibited no perceptible vignetting. More sophisticated procedures may be necessary to adjust images with uneven

illumination. A final criticism of this approach is that the image enhancement is meant to be used in conjunction with manual grading of the color photos. Many research groups are pursuing automated drusen detection in color photos,[26] and OCT may provide even more information about AMD in three dimensions.[27–29] However, all currently validated grading systems rely on human evaluation of photographs, and the algorithm developed in this study automates the time-consuming and difficult task of post hoc standardization.

This algorithm was developed to fully automate the standardization of color fundus photographs for AMD classification, thus eliminate problems of manual adjustments, such as subjective variability. Brightness, contrast, and color balance differences due to natural or artificial intraocular lenses or camera type can be rapidly and accurately corrected to give consistent output. Furthermore, the automated enhancement algorithm may improve the visualization of disease features, including drusen and pigmentary changes, and is helpful in the diagnosis, grading, and thus management of AMD. Automated standardization is particularly important when studying biomarkers for AMD progression or the effect of intervention in clinical trials.

Acknowledgments

The authors would like to thank all US and Portuguese ophthalmologists, photographers, technicians, and nurses who performed the required study procedures.

This project was financially supported by the Miller Retina Research Fund (MEE), the Miller Champalimaud Award (MEE), the Portuguese Foundation for Science and Technology/Harvard Medical School Portugal Program (HMSP-ICJ/006/2013), the Fidelity Charitable Fund (Harvard University), the Massachusetts Lions Eye Research Fund, an American Glaucoma Society Mid-Career Award, and a National Institutes of Health award (UL 1 RR025758).

* Edem Tsikata and Inês Laíns equally contributed to this work.

Disclosure: **E. Tsikata**, None; **I. Laíns**, Grant from the Portuguese Foundation of Science and Technology, Harvard Medical School Portugal Program (S); **J. Gil**, None; **M. Marques**, None; **K. Brown**, None; **T. Mesquita**, None; **P. Melo**, None; **M. da Luz Cachulo**, None; **I.K. Kim**, Genentech (C), Allergan (C), Icon Therapeutics (C); **D. Vavvas**, None; **J. Murta**, Alcon (C); **J.B. Miller**, Allergan (C); **R. Silva**, Allergan (C), Novartis (C), Thea (C), Alimera (C), Alcon (C); **J.W. Miller**, Alcon Research Institute (C), Amgen (C), KalVista Pharmaceuticals (C), Maculogix (C), Valeant Pharmaceuticals via Massachusetts Eye and Ear (P, R), ONL Therapeutics (C, P); **T.C. Chen**, None; **D. Husain**, None

References

1. Wong WL, Su X, Li X, et al. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. *Lancet Glob Heal*. 2014;2:e106–e116.
2. Lim LS, Mitchell P, Seddon JM, Holz FG, Wong TY. Age-related macular degeneration. *Lancet*. 2012;379:1728–38.
3. Age-Related Eye Disease Study Research 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.
4. Seddon JM, Sharma S, Adelman RA. Evaluation of the clinical age-related maculopathy staging system. *Ophthalmology*. 2006;113:260–266.
5. Klein R, Davis MD, Magli YL, Segal P, Klein BE, Hubbard L. The Wisconsin age-related maculopathy grading system. *Ophthalmology*. 1991;98:1128–1134.
6. van Leeuwen R, Chakravarthy U, Vingerling JR, et al. Grading of age-related maculopathy for epidemiological studies: is digital imaging as good as 35-mm film? *Ophthalmology*. 2003;110:1540–1544.
7. Somani R, Tennant M, Rudnisky C, et al. Comparison of stereoscopic digital imaging and slide film photography in the identification of macular degeneration. *Can J Ophthalmol*. 2005;40:293–302.
8. Klein R, Meuer SM, Moss SE, Klein BE, Neider MW, Reinke J. Detection of age-related macular degeneration using a nonmydriatic digital camera and a standard film fundus camera. *Arch Ophthalmol*. 2004;122:1642–1646.

9. Rasta SH, Partovi ME, Seyedarabi H, Javadzadeh A. A comparative study on preprocessing techniques in diabetic retinopathy retinal images: illumination correction and contrast enhancement. *J Med Signals Sens*. 5:40–48.
10. Teng T, Lefley M, Claremont D. Progress towards automated diabetic ocular screening: a review of image analysis and intelligent systems for diabetic retinopathy. *Med Biol Eng Comput*. 2002;40:2–13.
11. Fleming AD, Philip S, Goatman KA, Olson JA, Sharp PF. Automated assessment of diabetic retinal image quality based on clarity and field definition. *Invest Ophthalmol Vis Sci*. 2006;47: 1120–1125.
12. Neelam K, Muldrew A, Hogg R, Stack J, Chakravarthy U, Beatty S. Grading of age-related maculopathy: slit-lamp biomicroscopy versus an accredited grading center. *Retina*. 2009;29:192–198.
13. Ishiko S, Akiba J, Horikawa Y, Yoshida A. Detection of drusen in the fellow eye of Japanese patients with age-related macular degeneration using scanning laser ophthalmoscopy. *Ophthalmology*. 2002;109:2165–2169.
14. Bartling H, Wanger P, Martin L. Automated quality evaluation of digital fundus photographs. *Acta Ophthalmol*. 2009;87:643–647.
15. Patton N, Aslam TM, MacGillivray T, et al. Retinal image analysis: concepts, applications and potential. *Prog Retin Eye Res*. 2006;25:99–127.
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. Cachulo M da L, Lobo C, Figueira J, et al. Prevalence of age-related macular degeneration in Portugal: The Coimbra Eye Study– report 1. *Ophthalmologica*. 2015;233:119–127.
18. 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.
19. 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.
20. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42:377–381.
21. ITU. International Telecommunication Union–Radiocommunication (ITU-R) BT.601 standard. Geneva, Switzerland. 1982.
22. Said FS, Weale RA. The variation with age of the spectral transmissivity of the living human crystalline lens. *Gerontologia*. 1959;3:213–231.
23. Boettner EA, Wolter JR. Transmission of the ocular media. *Invest Ophthalmol Vis Sci*. 1962;1: 776–783.
24. Tanito M, Okuno T, Ishiba Y, Ohira A. Transmission spectrums and retinal blue-light irradiance values of untinted and yellow-tinted intraocular lenses. *J Cataract Refract Surg*. 2010; 36:299–307.
25. Friedman DS, Katz J, Bressler NM, Rahmani B, Tielsch JM. Racial differences in the prevalence of age-related macular degeneration: the Baltimore Eye Survey. *Ophthalmology*. 1999;106:1049–1055.
26. Mora AD, Vieira PM, Manivannan A, Fonseca JM. Automated drusen detection in retinal images using analytical modelling algorithms. *Biomed Eng Online*. 2011;10:59.
27. Folgar FA, Chow JH, Farsiu S, et al. Spatial correlation between hyperpigmentary changes on color fundus photography and hyperreflective foci on SDOCT in intermediate AMD. *Invest Ophthalmol Vis Sci*. 2012;53:4626–4633.
28. Yi K, Mujat M, Park BH, Sun W, Miller JW, Seddon JM, Young LH, de Boer JF, Chen TC. 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.
29. Chiu SJ, Izatt JA, O'Connell RV, Winter KP, Toth CA, Farsiu S. Validated automatic segmentation of AMD pathology including drusen and geographic atrophy in SD-OCT images. *Invest Ophthalmol Vis Sci*. 2012;53:53–61.