Impact of Binarization Thresholding and Brightness/Contrast Adjustment Methodology on Optical Coherence Tomography Angiography Image Quantification

NIHAAL MEHTA, KEKE LIU, A. YASIN ALIBHAI, ISAAC GENDELMAN, PHILLIP X. BRAUN, AKIHIRO ISHIBAZAWA, OSAMA SOROUR, JAY S. DUKER, AND NADIA K. WAHEED

• PURPOSE: **Binarization is a critical technique in optical coherence tomography angiography (OCTA) image analysis, but there is no consistency in the method used in published OCTA studies. This study assessed whether differences in OCTA binarization and brightness and contrast adjustments affect quantification measurements.**

• DESIGN: **Prospective cross-sectional validity study.**

• METHODS: **This was a single-center study examining 21 eyes of 11 healthy individuals. All eyes were imaged using a swept-source OCTA (Zeiss), and quantitative measurements resulting from five binarization thresholding and five brightness/contrast adjustment methods were compared. All measurements were calculated for the superficial plexus and choriocapillaris (CC), as well as unaveraged and averaged en face OCTA images.**

• RESULTS: **There were statistically significant differences between measurements from different binarization thresholding methods ($P < 0.0001$), as well as measurements from different histogram adjustments ($P < 0.0001$). The binarization thresholds yielded different measurements when combined with variable brightness/contrast adjustments. The method of analysis also affected the directionality of trends in imaging measurements between unaveraged and averaged CC images.**

• CONCLUSIONS: **The method of OCTA image binarization thresholding and histogram adjustment significantly alters quantitative measurements and the directionality of trends. Results obtained from different OCTA binarization methods should be seen as valid only for that given method. This has significant consequences for clinical trials using OCTA measurements as outcome measurements. A consensus is needed across the research community for a consistent method for OCTA image quantification and greater attention paid to fully describing methods in published studies. (Am J Ophthalmol 2019;205:54–65. © 2019 Published by Elsevier Inc.)**

SINCE ITS INCEPTION IN 1991, OPTICAL COHERENCE TOMography (OCT) has become a mainstay in ophthalmology, allowing micrometer-level resolution imaging of ocular structures.[1] In the past several years, OCT angiography (OCTA) has emerged as a means of providing detailed images of retinal vasculature.[2,3] Currently available OCTA devices are capable of generating high-quality en face images of retinal plexi. Particularly for research purposes, the ability to reproducibly and objectively quantify OCTA images is critical, as it allows images to be compared. Such analyses have applications in clinical trials that use OCTA to monitor disease progression. OCTA image quantification of the superficial plexus, deeper layers, and most recently the choriocapillaris (CC) are widely deployed in research settings.[4–9] The most widely used OCTA measurement is vessel area density (VAD), which uses the binarized image to quantify the percentage of the image occupied by "flow" information (conventionally displayed in white) as a fraction of the total image area. Another common measurement applied to skeletonized images of the retinal vasculature is vessel length (VL). In the CC, the number and size of nonflow or "flow deficit" areas are often assessed. Other measurements, such as total vessel area and vessel diameter index, are easily calculated from the aforementioned ones. As imaging technology and ocular therapeutics continue to improve, it is likely that OCTA quantification of all vascular layers will expand into clinical use.

In some cases, OCTA instruments have built-in software that can calculate imaging measurements. The most recent software available for the RTVue-XR Avanti SD-OCTA system (Optovue Inc., Fremont, California, USA) and the Nidek RS-3000 Advance 2 (Nidek Inc., Fremont, California, USA), for example, can automatically calculate

AJO.com Supplemental Material available at AJO.com.

Accepted for publication Mar 6, 2019.

From the New England Eye Center (N.M., K.L., A.Y.A., I.G., P.X.B, A.I., O.S., J.S.D., N.K.W.), Tufts Medical Center, Boston, Massachusetts, USA; The Warren Alpert Medical School of Brown University (N.M.), Providence, Rhode Island, USA; University of Hawai'i John A. Burns School of Medicine (K.L.), Honolulu, Hawai'i, USA; Yale School of Medicine (P.X.B.), New Haven, Connecticut, USA; Department of Ophthalmology (A.I.), Asahikawa Medical University, Asahikawa, Japan; Department of Ophthalmology (O.S.), Tanta University, Egypt.

Inquiries to Nadia K. Waheed, Department of Ophthalmology, Tufts Medical Center, 800 Washington Street, Box 450, Boston, Massachusetts 02111 USA; e-mail: nwaheed@tuftsmedicalcenter.org

0002-9394/$36.00
https://doi.org/10.1016/j.ajo.2019.03.008

VAD in different regions of its images. However, at the time of this study, this capability was not universally available across different devices. For example, presently, the ability of the built-in software in the Cirrus HD-OCT 5000 with Angioplex and the PLEX Elite 9000 (Carl Zeiss Meditec, Dublin, California, USA) to calculate measurements is not yet available in the United States. As such, it has become common for researchers to use quantification methods on en face OCTA images after exporting from the machine.

All these methods require applying an image threshold in order to produce binarized OCTA images. It is important to note the distinction between thresholding of OCTA signal data to generate an OCTA or flow image, which is an essential step in OCTA image formation and has been discussed in detail by our group,[10] and thresholding of resulting OCTA images for quantitative analysis purposes, which is the process examined here. This latter process is perhaps better termed "binarization thresholding" to clearly distinguish it from OCTA signal thresholding. In short, binarization thresholding involves applying an algorithm to already generated OCTA images to assign all image pixels to be either black (a grayscale value of 0 in conventional 8-bit images) or white (a grayscale value of 255), resulting in a binary black-and-white image. The "threshold" level is set such that all pixels with a grayscale value above the threshold are assigned to be white and vice versa. The simplest form of binarization thresholding is global, which applies the same threshold value across the entire image.[11] Another approach is to perform local, or adaptive, binarization thresholding. Instead of assigning one global threshold value across the image, these local methods assign different threshold values to various areas of the image, in essence "adapting" to the local histograms. An obvious advantage of local binarization thresholding is to account for variable illumination or contrasting. There are many different methods available for binarization thresholding. Partly as a result, there is little consistency in the method deployed for OCTA image quantification to calculate measurements.

Moreover, image histogram adjustments can be made by altering brightness and/or contrast (brightness/contrast). The ability to adjust brightness/contrast during image acquisition is built into the software of some OCT systems (including the PLEX Elite; Zeiss). These adjustments may also be made during image analysis in order make the image appearance qualitatively consistent (eg, to appear to have the same brightness) and/or optimize analysis. This is often necessary because OCTA images can have variable appearances; for example, within 1 acquisition from the same subject, an en face angiogram of the deep capillary plexus can have lower contrast than that of the superficial plexus. As with binarization thresholding, there is no clear consensus as to the best method for standardizing image histograms.

Because OCTA quantification is becoming increasingly common, there is a pressing need to understand how different methods can affect measurements and how well varying methods perform in both producing accurate measurements and minimizing variability. If differences in methods can change measurements, this would be consequential for any research or clinical trial that uses OCTA measurements as an outcome measurement. To the authors' knowledge, no published study has compared the effects of different binarization thresholding methods, both local and global, and brightness and contrast adjustments, on quantification measurements of OCTA images of the retinal vasculature and CC. The present study examined the effects of variable binarization thresholding and brightness/contrast adjustments on quantification of the superficial plexus and CC OCTA images. Whether different methods alter the effect that en face averaging has on quantification was then assessed.

SUBJECTS AND METHODS

THIS PROSPECTIVE STUDY WAS PERFORMED AT THE NEW England Eye Center (Boston, Massachusetts, USA), evaluating normal eyes between May and August 2018. The study protocol was approved by the Tufts Medical Center Institutional Review Board and adhered to the tenets of the Declaration of Helsinki and the Health Insurance Portability and Accountability Act of 1996. Written informed consent was obtained from all participants in accordance with the Tufts Medical Center institutional review board.

• **IMAGE ACQUISITION:** Twenty-one eyes from 11 healthy individuals with no retinal pathology and a refractive error of <7.0 diopters (D) were imaged using a 3- × 3-mm scan protocol centered on the fovea with the Plex Elite 9000 SS-OCTA (Zeiss). The Plex Elite 9000 (Zeiss) uses a 1060-nm central wavelength and images at 100,000 A-scans per second with an axial resolution of 6.3 μm and a transverse resolution of 20 μm. The image cubes are acquired with 300 A-scans per B-scan and 300 B-scans per volume for a 3- × 3-mm view. The final superficial plexus and CC images are resized by the built-in software to an output size of 1,024 × 1,024 pixels during exporting. Three successive images were obtained in 8 eyes for averaging. All images were obtained by trained ophthalmic photographers, and acquisitions were repeated multiple times, if necessary, to obtain high-quality images.

• **IMAGE PROCESSING:** Automated segmentation of retinal and choroidal layers was performed by software in the OCTA machine. The superficial plexus slab was defined as the distance from the internal limiting membrane to the inner plexiform layer and the CC from 29 μm to 49 μm posterior to the automatically fitted retinal pigment epithelium line. En face angiograms were exported for analysis and all image analysis was performed using version 2.0.0

of the Fiji distribution of ImageJ software (National Institutes of Health, Bethesda, Maryland, USA). For averaging, linear registration followed by elastic registration was performed in ImageJ software by using the plug-in "Register Virtual Stack Slices" on the 3 superficial plexus images. Features were extracted using the scale-invariant feature transform and multiscale-oriented patches algorithms, followed by random sample consensus.[12–14] The linear registration was conducted using a rigid registration model which translates and rotates the images in order to most closely match feature points. The elastic registration model performs B-spline based elastic deformation to best match the images.[15] The same transformations were then applied to the CC images from the same acquisition. This method was used because the superficial plexus has more distinct features for registration and the method was based on a prior published CC analysis.[16] After registration, the 3 images of both the superficial plexus and the CC from each subject were averaged to yield 1 image of each layer.

• **QUANTITATIVE ANALYSIS:** Our analysis was limited to a central 2.5-mm (853-pixel) diameter circle in order to exclude artifacts from averaging, which occurs at the periphery of images during the registration process. Each image was analyzed separately with 5 binarization thresholds: the global default, the global mean, the global Otsu, the local mean, and the local Phansalkar. These specific methods were chosen because they have been applied in the recent OCTA published studies, and they represent distinct conceptual and mathematical approaches to the problem of image binarization. The default ImageJ parameters were not changed; for all local binarization thresholds, a radius of 15 pixels (43.9 μm) was used. The mean global binarization threshold method takes the average grayscale value across the image as a threshold value. The default global method in ImageJ software is a modified version of the IsoData binarization thresholding algorithm, which divides the image into foreground and background pixels and iteratively tries different threshold values until one larger than the mean value is found.[17] The IsoData approach was designed to improve upon a simple mean binarization threshold and to function even in a lower-contrast image without a clearly bimodal histogram. Finally, among global binarization thresholds, the Otsu method is often applied to OCTA analysis. This algorithm determines a threshold value that minimizes the variance in grayscale values within each class and maximizes variance among the classes. The mean method can also be applied locally through a local mean binarization threshold. More recently, the Phansalkar local binarization threshold has been used on en face OCTA images of the CC and superficial plexus.[6,7,18] The Phansalkar algorithm incorporates the mean and SD of the grayscale values in the local area. It was designed for images with a variable appearance and particularly to optimize binarization thresholding in low-contrast images.[19]

Brightness/contrast adjustments were also applied. The initial pixel values were coded as 8-bit values, ranging from 0-255 pixels. Following transformation, pixel values were clipped to the 0-to 255-pixel range. Five image brightness/contrast adjustments were separately applied, as follows: 1) no change, 2) increased brightness [$f_B(p) = p + 50$], where $f(p)$ = transformed pixel value and p = initial pixel value; 3) increased contrast [$f_C(p) = 255 \times (p - 50)/155$)]; 4) increased brightness and contrast, [$f_{BC}(p) = 255 \times (\{p + 50\} - 50)/155$)]; and 5) contrast-limited adaptive histogram equalization (CLAHE) (the default ImageJ parameters are retained: where block size is 127, histogram bins are 256, maximum slope is 3.00, with fast analysis). CLAHE performs histogram equalization on local areas of the image and limits increases in contrast to prevent noise amplification. VAD was computed from the binarized images using the "Measure" feature. For the CC images, the "Analyze Particles" tool was used to measure the number of flow deficits, the total flow deficit area, and the average flow deficit area size. Finally, the superficial plexus images were skeletonized using the plug-in "Skeletonize" feature, and the total vessel length was calculated using the "Analyze Skeleton" feature. This analysis was performed over the 21 unaveraged and 8 averaged images on both the superficial plexus and the CC slabs. The entire process was automated using an ImageJ software macro.

• **COMPARISONS:** To analyze the effect of binarization thresholding, measurements were compared using the 5 different binarization thresholds with no brightness/contrast changes. To analyze the effects of brightness/contrast changes, a fixed binarization threshold (local Phansalkar) was used. To examine the combined effects of binarization thresholding and brightness/contrast adjustments, only VAD in the superficial plexus and CC were compared across all 5 binarization thresholds and brightness/contrast changes. Finally, for analyzing the effect of en face averaging in the CC, measurements in the unaveraged and the 8 averaged images were compared using all 5 binarization thresholds and brightness/contrast adjustments.

• **STATISTICAL ANALYSIS:** All statistical analyses were performed using Stata/SE version 15.0 software (StataCorp, College Station, Texas, USA). To examine whether differences in quantitative variables were significant between the binarization threshold and the brightness/contrast adjustment methods, a mixed-effects analysis of variance (ANOVA) model was used to account for the correlation between eyes from the same individual. All pairwise comparisons were performed using the Tukey honest significant difference method to adjust for multiple comparisons. For analyzing the effects of thresholding and brightness/contrast adjustments on en face averaging, the element of interest was the directionality of trends rather than significant differences and, as such, statistical hypothesis testing was not applied to these values.

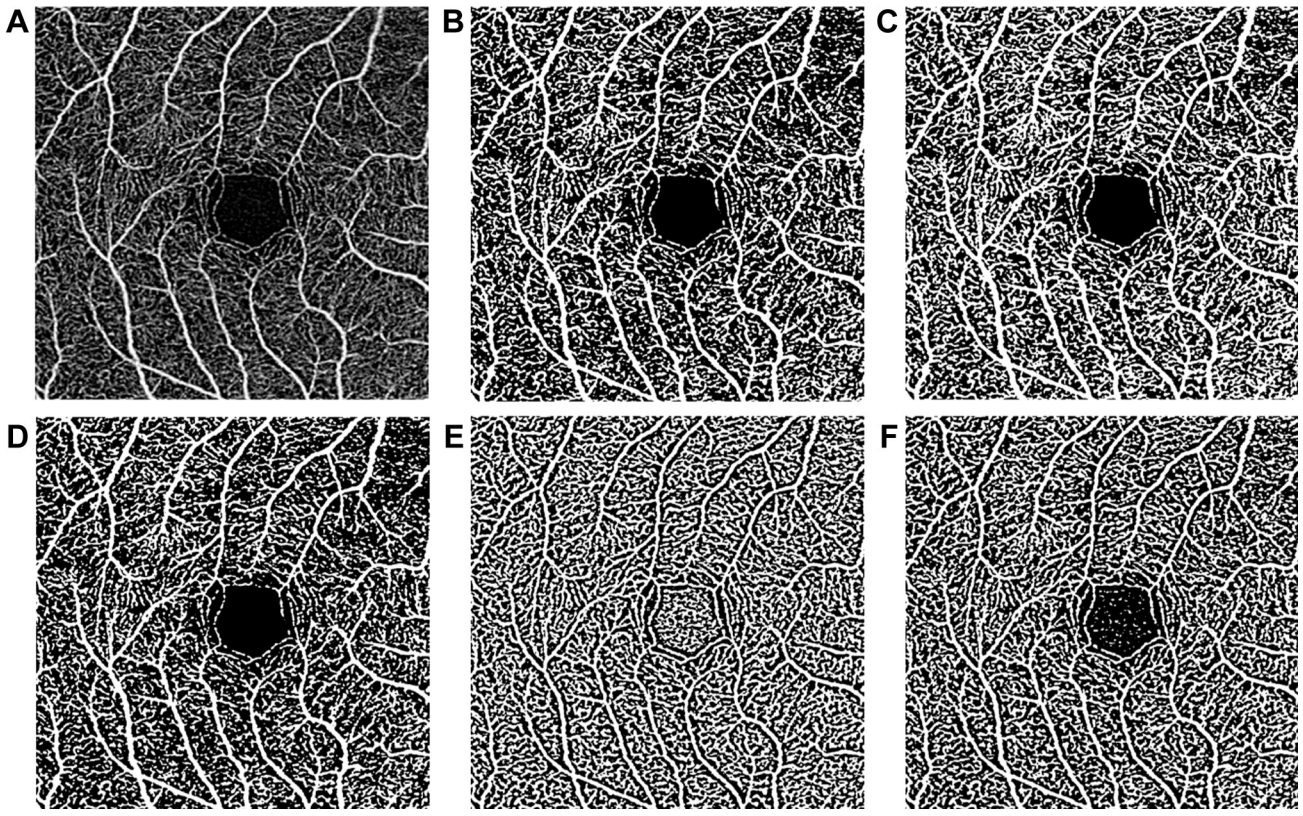

FIGURE 1. Qualitative effect of binarization thresholding on superficial plexus images (A). An original superficial plexus image is shown. (B) The image has been binarized using the global default binarization threshold (with a threshold grayscale value of 59). (C) The global mean binarization threshold (threshold value of 52). (D) The global Otsu binarization threshold (threshold value of 61). (E) Local mean binarization threshold (F) and local Phansalkar binarization threshold.

RESULTS

• **BINARIZATION THRESHOLDING:** Qualitatively, variations in binarization thresholding had a significant effect on the resulting binarized images, as shown in Figures 1 and 2. In the superficial plexus (Figure 1), the global binarization thresholds showed minimal noise in the foveal avascular zone (FAZ), whereas the local mean binarization threshold created a large amount of FAZ noise.

In the CC (Figure 2), the differences between local and global binarization thresholds were even more pronounced. All global binarization thresholds produced binarized images that qualitatively appeared identical. The local mean binarization threshold resulted in a more homogenous appearance. The Phansalkar local binarization threshold noticeably segmented most of the CC image into flow areas (white) with more numerous and smaller flow deficits (black).

The mean values for quantitative measurements across different binarization thresholds with no brightness/contrast changes are shown in Table 1. All differences for both the superficial plexus and CC were statistically significant using the ANOVA model ($P < 0.0001$). Significant pairwise comparisons are summarized in Supplemental Figure 1 (Supplemental Material available at www.ajo.com). In the superficial plexus, local binarization thresholds tended to produce higher VAD and VL numbers. In the CC, the Phansalkar binarization threshold uniquely produced a significant increase in VAD and resulted in many more flow deficits with a smaller average flow deficit size.

• **BRIGHTNESS AND CONTRAST ADJUSTMENTS:** The qualitative differences between images subjected to different brightness/contrast adjustments are shown in Figure 3 (superficial plexus) and Figure 4 (CC). Increasing brightness resulted in translation of the image histogram with no change in its distribution. Increasing contrast stretched the histogram across the x-axis. CLAHE broadened and flattened the histogram.

The mean values for quantitative measurements across different brightness/contrast adjustments with the same binarization thresholding method (local Phansalkar) are summarized in Table 2. All differences for both the superficial plexus and the CC were statistically significant using the ANOVA model ($P < 0.0001$). Significant pairwise comparisons are summarized in Supplemental Figure 2.

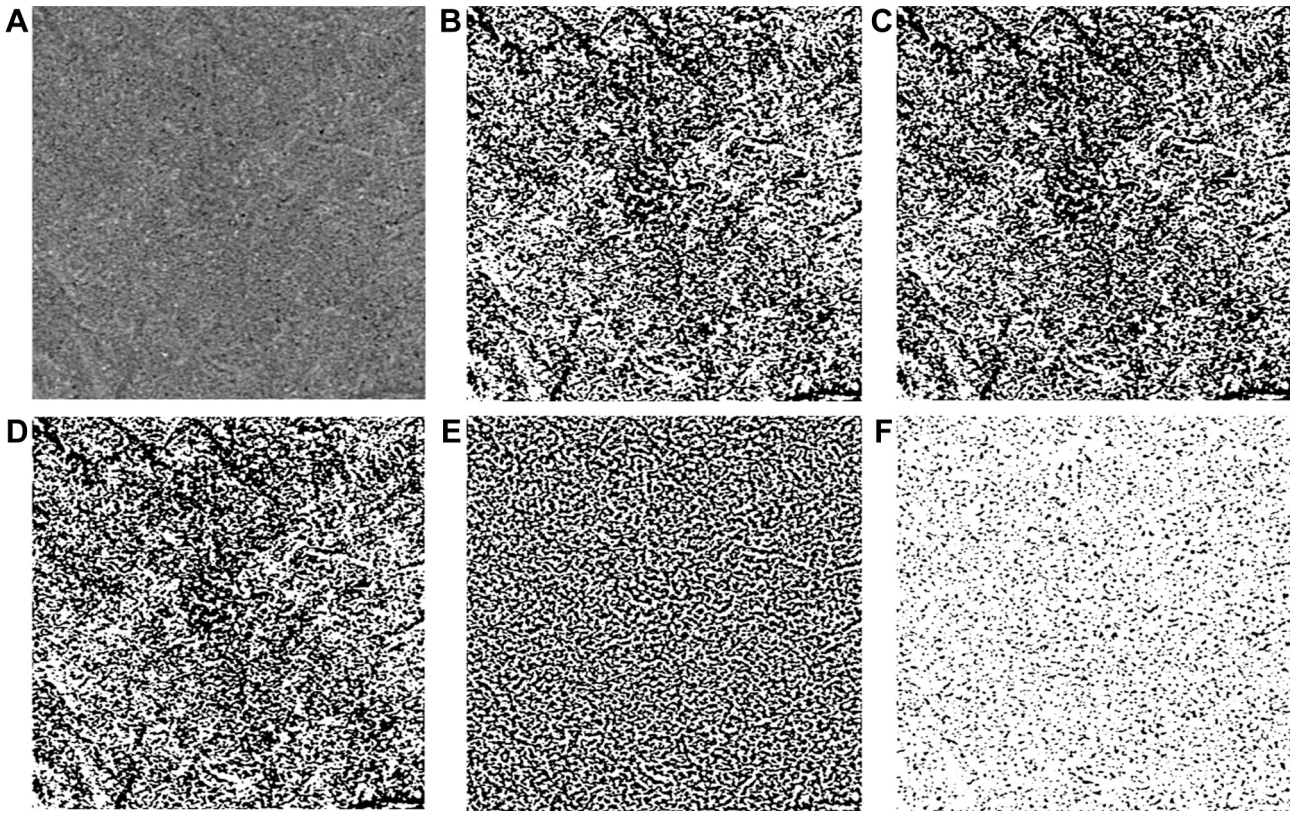

FIGURE 2. Qualitative effects of variable binarization thresholding on choriocapillaris images. (A) An original choriocapillaris image is shown. (B) The original image has been binarized using the global default binarization threshold (threshold grayscale value of 107). (C) The global mean binarization threshold (threshold value of 108) (D) The global Otsu binarization threshold (threshold value of 108) (E) The local mean binarization threshold (F) and local Phansalkar binarization threshold.

Increasing brightness alone had a relatively small effect. Increasing contrast decreased VAD and VL in the superficial plexus. In the CC, increasing contrast decreased VAD while decreasing number and increasing average size of flow deficits. Increasing both the brightness and the contrast increased superficial plexus VL but otherwise had a relatively small effect. Finally, CLAHE increased VAD and VL in the superficial plexus, decreased VAD in the CC, and decreased number while increasing average size of CC flow deficits.

• **COMBINED EFFECT OF VARIABLE BINARIZATION THRESHOLDING AND BRIGHTNESS/CONTRAST CHANGES ON VAD:** In the superficial plexus (Figure 5), increasing brightness and increasing brightness/contrast did not change the distribution of the VAD values between or within binarization thresholds. Increasing contrast, however, caused a wider distribution between binarization thresholds and decreased overall VAD values. CLAHE decreased the distribution between and within binarization thresholds and slightly increased most VAD values. In the CC (Figure 6), increasing contrast and applying CLAHE caused the Phansalkar local binarization threshold to yield VAD measures similar to the other binarization thresholds.

• **EN-FACE AVERAGING:** The mathematical differences between mean measurements from the 8 averaged images and all unaveraged (averaged minus unaveraged) superficial plexus and CC images across different binarization thresholds and brightness/contrast changes are shown in Table 3. Without making brightness or contrast changes, in the superficial plexus the en face averaging method caused an increase in VAD with all binarization thresholds. VL increased with averaging by using all binarization thresholds but was unchanged using the local mean. In the CC images, VAD increased with averaging using all binarization thresholds but was unchanged using the local mean. The number of flow deficits decreased with averaging using all binarization thresholds except for the global mean (580-1,216). The average size of flow deficits decreased with averaging using all binarization thresholds except for the local mean (5,797-6,004 μm^2).

Across different brightness and contrast changes applied to both the unaveraged and the averaged images while using a fixed binarization threshold (local Phansalkar), superficial plexus VAD and VL increased with averaging unless contrast was increased, in which case VAD and VL were unchanged. In the CC images, VAD increased, and average flow deficit size decreased with averaging for all brightness

TABLE 1. Methods for OCTA Quantification: Quantitative Effects of Binarization Thresholding

	Mean (±SD) Values				
	Superficial Plexus		Choriocapillaris		
Binarization Threshold	VAD (%)	VL (mm)	VAD (%)	Number of Flow Deficits	Average Flow Deficit Size (μm^2)
Global default	39.3 (±3.2)	120.9 (±14.6)	56.7 (±3.7)	872 (±226)	2,656 (±963)
Global mean	43.3 (±1.6)	130.9 (±10.1)	50.9 (±1.8)	580 (±105)	4302 (±880)
Global Otsu analysis	38.2 (±3.4)	118.1 (±15.2)	52.9 (±3.7)	682 (±190)	3,723 (±1391)
Local mean	47.9 (±0.8)	145.2 (±8.1)	51.0 (±0.7)	428 (±76)	5,797 (±1059)
Local Phansalkar analysis	44.6 (±3.6)	135.6 (±14.5)	84.3 (±4.1)	2029 (±123)	382 (±2052)

Mean (±SD) values are shown for each measurement. There are significant differences between quantitative measurements in the superficial plexus and choriocapillaris OCTA images between binarization thresholding methods.

OCTA = optical coherence tomography angiography; SD = standard deviation; VAD = vessel area density; VL = vessel length.

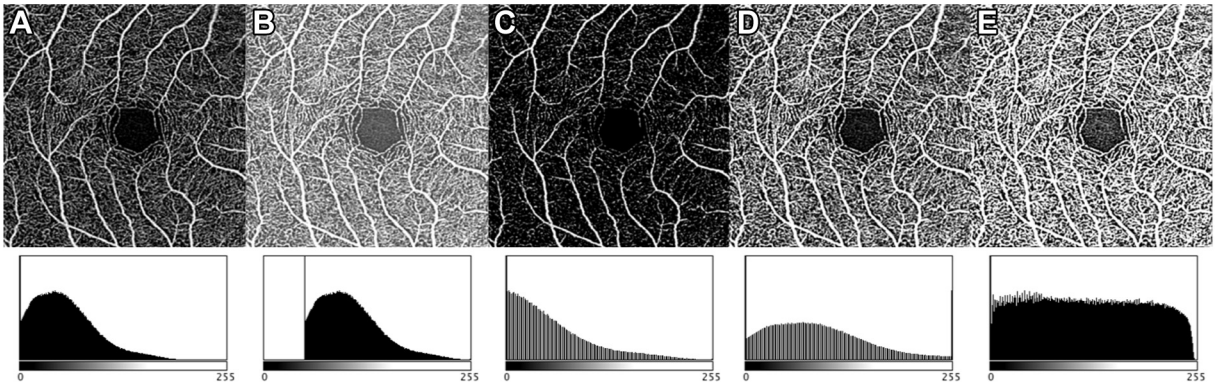

FIGURE 3. Qualitative effect of brightness and contrast adjustments on superficial plexus images. (A) An original superficial image is shown with no change in brightness or contrast. (B) An image with increased brightness. (C) An image with increased contrast. (D) Image with increased brightness and contrast. (E) Image with contrast-limited adaptive histogram equalization.

and contrast changes. The number of flow deficits decreased with averaging unless contrast was increased or CLAHE was applied. In both of these cases, the number of flow deficits increased (1,426-1,714 and 1,219-1,569, respectively).

DISCUSSION

OCTA QUANTIFICATION HAS BECOME COMMONPLACE, BUT there is no consistent method for analyzing images in published reports. Variations of all the binarization thresholding algorithms included in this analysis have been previously applied. An IsoData algorithm was recently used by Kim and associates,[20] and the ImageJ Default method, specifically, was used by Cicinelli and associates[21] on choroidal images. The Otsu algorithm was used by Al-Sheikh and associates[4] for the superficial plexus, Ohayon and associates[22] for an outer retina to CC slab, and Chidambara and associates[23] for superficial, deep, outer retina, and choroidal slabs. Park and associates[24] used a mean-based method on full-thickness retinal slab angiograms. As previously mentioned, Spaide,[6] Uji and associates,[7] and Al-Sheikh and associates[4] recently used the Phansalkar local binarization threshold for CC images, and Tang and associates[18] for superficial plexus images. In addition, there are a wide variety of other binarization methods applied in the studies. Wang and associates[8] used a customized global binarization threshold for superficial plexus images and a separate global binarization threshold for the CC. Uji and associates[16] used a combination of a median local binarization threshold and Huang's global binarization threshold, whereas Kim and associates[25] used a Hessian filter along with the median local binarization threshold. Many studies have also simply used automated quantification from the OCTA machine itself.[26–28] Still other studies do not, as far as these authors can tell, specify the exact binarization thresholding method.[29–31] For adjusting brightness/contrast of images, Uji and

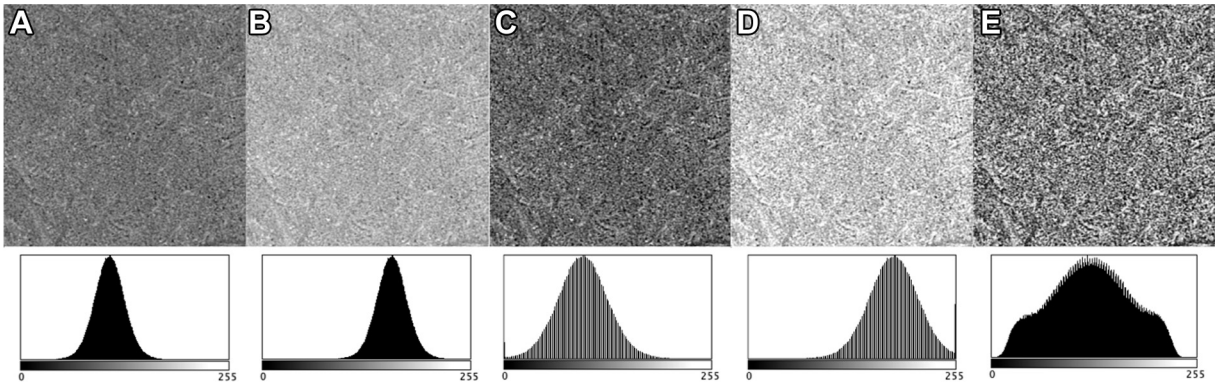

FIGURE 4. Qualitative effect of brightness and contrast adjustments on choriocapillaris images. (A) An original choriocapillaris image is shown with no change in brightness or contrast. (B) Image shown with increased brightness. (C) Image shown with increased contrast. (D) Image shown with increased brightness and contrast. (E) Image with contrast-limited adaptive histogram equalization.

TABLE 2. Methods for OCTA Quantification: Quantitative Effects of Brightness and Contrast Changes

	Mean (±SD) Values				
	Superficial Plexus		Choriocapillaris		
Brightness/Contrast Adjustments	VAD (%)	VL (mm)	VAD (%)	Number of Flow Deficits	Average Flow Deficit Size (μm^2)
No change	44.6 (±3.6)	135.6 (±14.5)	84.3 (±4.2)	2,029 (±123)	382 (±107)
Increased brightness	46.7 (±2.9)	140.6 (±12.9)	84.6 (±4.0)	2,032 (±118)	374 (±101)
Increased contrast	28.2 (±2.7)	93.1 (±13.1)	65.4 (±4.8)	1,426 (±403)	1,308 (±445)
Increased brightness and contrast	49.8 (±2.5)	148.1 (±11.7)	84.2 (±4.0)	2,043 (±121)	382 (±102)
CLAHE	53.3 (±1.3)	156.7 (±8.9)	62.1 (±2.2)	1,219 (±249)	1,591 (±328)

Mean (±SD) values are shown for each measurement. There are significant differences between the quantitative measurements in the superficial plexus and the choriocapillaris and the brightness/contrast adjustment methods. All values were obtained using the local Phansalkar binarization threshold.

CLAHE = contrast-limited adaptive histogram equalization; OCTA = optical coherence tomography angiography; SD = standard deviation; VAD = vessel area density; VL = vessel length.

associates[7] recently used CLAHE to optimize low-contrast CC images, and Spaide[32] similarly used CLAHE on posterior vitreous OCT B-scans with intraimage variations. Moreover, as previously mentioned, adjustments to image brightness and contrast can be made during image acquisition itself by using the OCT system's built-in software and may not be reported in the method of published studies. As such, there is a need to better understand the impact of variations in methods.

The various binarization thresholds examined in this study produced different binarized images. Global thresholds tended to appear similar because, although these methods use different mathematical algorithms, the threshold values they generate can be similar and are consistently applied across the image. Local thresholds produce a more homogenous appearance because they mathematically account for regional variations in image appearance; the threshold value for darker areas will tend to be scaled up, for example, resulting in a binarized image that appears more uniform.

In this study, the effect of binarization thresholding was less pronounced, although still significant, in the superficial plexus, likely because imaging features are relatively distinct. In the CC, there were statistically significant differences from variable binarization thresholding, particularly when applying the local Phansalkar binarization threshold. The inevitable question of which binarization threshold is "best" must be approached from a number of perspectives. The ideal binarization threshold will accurately binarize an image, preserving true blood flow information, and produce quantitative measurements that reflect anatomic reality. The amount of noise introduced into the FAZ is a useful means of comparing different methods, as the FAZ is a fairly large avascular area that is easy to visualize; however, the issue of noise introduction by binarization thresholding is not limited to the FAZ, as

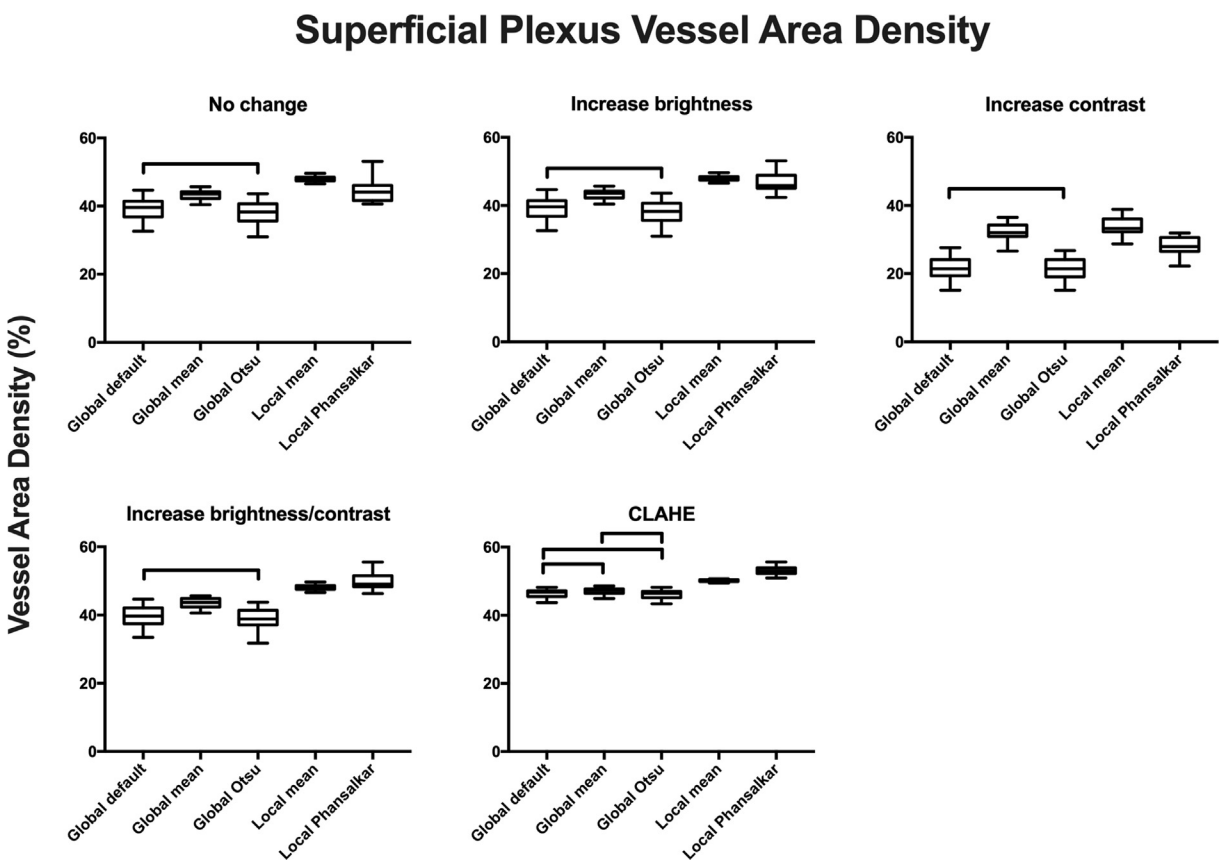

FIGURE 5. Interaction between binarization thresholding and brightness/contrast changes in the superficial plexus. Boxplots of superficial plexus vessel area density (VAD) across 5 binarization thresholding and brightness/contrast adjustment methods show changes in intra- and interclass variability. Individual plots show median VAD (line) with 25th-75th percentiles (box) and ranges (brackets). All pairwise comparisons are significant ($P < 0.05$), except for those indicated by horizontal brackets. VAD = vessel area density.

other nonperfusion or low-intensity image areas would be subjected to the same issues. As shown in Figure 1, local binarization methods subject the FAZ to a low local threshold value, causing thresholding of values to be white, whereas they should be black. The local Phansalkar method better minimizes FAZ noise than the local mean method. Although excluding the FAZ in analysis is possible, doing so in an automated way without also excluding nearby vessels is difficult. Another potential option would be to apply a local binarization threshold with a larger local radius.

In the CC, qualitative comparisons between binarized images are challenging as there are few distinguishing features. One logical approach is to use histopathological studies to provide references for the anatomical size of CC structures. Calculating based on studies of post mortem eyes, the diameter of a CC flow deficit in a normal eye is in the magnitude of 10^1 to 10^2 μm^2.[33,34] By this standard, only average flow deficit values produced by the local Phansalkar binarization threshold fall within this range (385 μm^2 in the present analysis), whereas other binarization thresholds produce values on the order of 1 magnitude larger. However, there are many confounding effects in making this comparison. For example, the voids may not be captured in direct cross-section on OCTA, the resolution of OCTA does not compare to that of electron microscopy, and histopathologic studies are post mortem, whereas OCTA is performed in vivo.

Given the lack of a clear way of assessing the quality of binarization thresholding algorithms in OCTA analysis, it is difficult to conclude that any one method is clearly superior for either superficial plexus or CC images. It is clear that variations in binarization methods produce different quantitative results, and it is important to note that none of the methods examined produced an entirely accurate binarized image. This is consequential for research studies and clinical trials using OCTA measurements as outcome measures and provides a strong argument in favor of software built in to OCT systems that can produce quantitative measurements using a consistent method. However, each OCT system incorporates distinct image processing steps and likely generates measurements using different

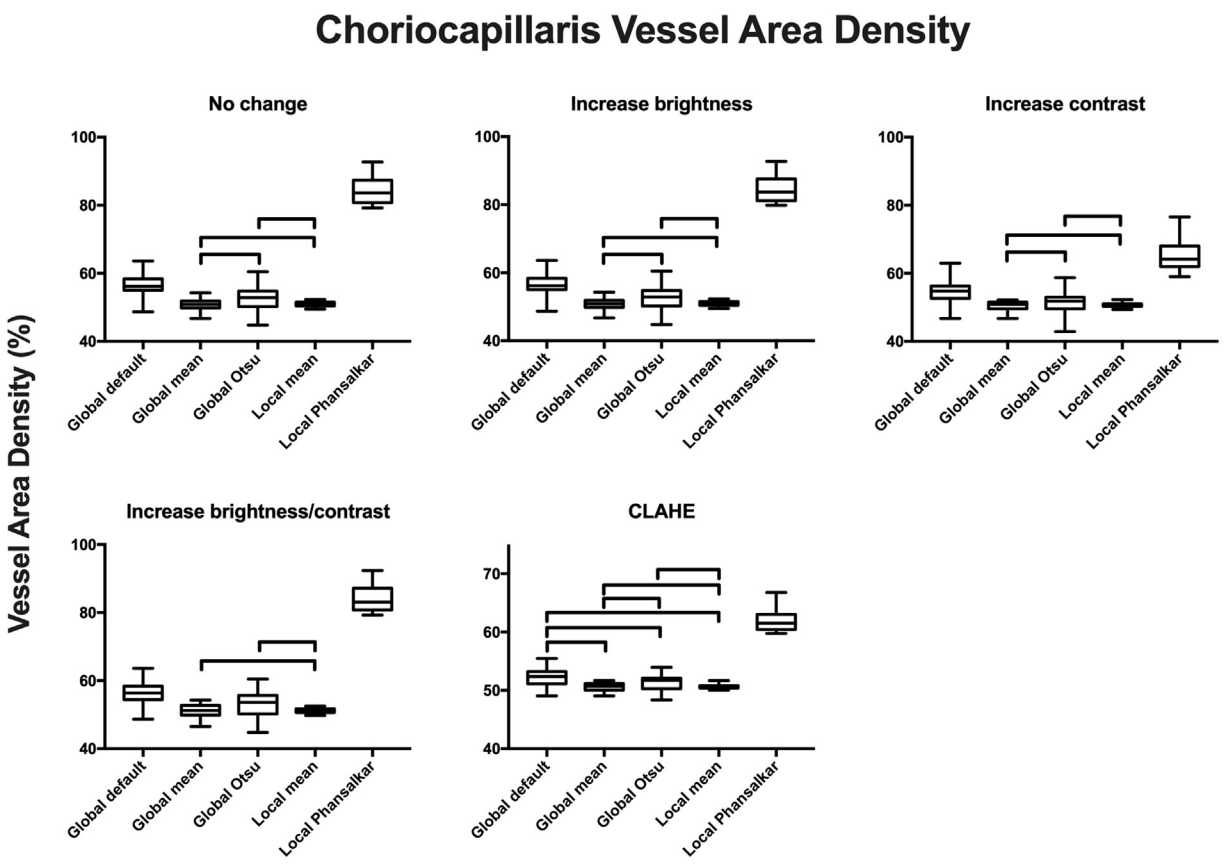

FIGURE 6. Interaction between binarization thresholding and brightness/contrast changes in the choriocapillaris. Boxplots of choriocapillaris VAD across five binarization thresholding and brightness/contrast adjustment methods show changes in intra- and inter-class variability. Individual plots show median VAD (line) with 25th-75th percentiles (box) and ranges (brackets). All pairwise comparisons are significant ($P < 0.05$) except for those indicated by horizontal brackets. Vessel area density (VAD).

quantification algorithms. Moreover, even built in software measurements cannot be guaranteed to be accurate until further studies compare various methods to known anatomical ground truth to verify whether they produce meaningful measurements. It would be beneficial to develop a measurement of accuracy for binarization methods compared to ground truth, as well as a consensus on the degree of accuracy needed for meaningful vascular quantification.

Histogram adjustments to alter brightness and contrast can also be easily applied during OCTA image analysis. In the present study, brightness and contrast adjustments significantly altered measurements. Adjusting brightness alone has the effect of translating the image histogram left or right on the x-axis, scaling all grayscale values uniformly (Figure 2). Because most binarization thresholding methods rely primarily on the mean of the histogram values, it makes sense that brightness adjustments alone did not drastically alter measurements; applying a mean-based binarization thresholding algorithm to a scaled histogram will result in an analogous scaling of the threshold value. The exception, however, is that, if brightness is increased or decreased to the point where grayscale values from the original histogram are either clipped to less than 0 or greater than 255, the pixels corresponding to these values will be desaturated or saturated, resulting in loss of imaging information and altered measurements. Contrast adjustments change the histogram's shape by either stretching or contracting the distribution of grayscale values, thereby altering mean-based calculations performed on the image. When contrast needs to be adjusted for analysis purposes (eg, low-contrast or variably illuminated images), it is difficult to apply a consistent contrast change across different images as the extent of the change depends on each original image; stretching a histogram by a set scalar will cause a relatively large change for a narrow histogram but a relatively small change for a wider one. CLAHE offers a potential means of applying a single algorithm consistently across images but also clearly alters quantitative measurements and thus must be used with caution.

In this analysis, contrast adjustments also interact with binarization thresholding, causing changes in the intra- and interclass variation in VAD measurements (Figures 5

TABLE 3. Methods for OCTA Quantification: Differences in Measurements Between En-face Averaged and Unaveraged Images

		Difference in Mean Values				
		Superficial Plexus		Choriocapillaris		
Binarization Threshold	Brightness/Contrast Change	VAD (%)	VL (mm)	VAD (%)	Number of Flow Deficits	Average Flow Deficit Size (μm^2)
Global default	No change	7	13	42	−594	−2,458
Global mean		6	10	20	636	−2,892
Global Otsu		6	12	42	−236	−3,425
Local mean		1	0	0	−18	207
Local Phansalkar		6	11	10	−957	−127
Local Phansalkar	No change	6	11	10	−957	−127
	Increased brightness	5	7	10	−961	−119
	Increased contrast	0	0	12	288	−583
	Increased brightness/contrast	5	6	10	−963	−124
	CLAHE	3	3	9	350	−602

Values are mathematical differences between measurements from the unaveraged versus the averaged images (averaged minus unaveraged). Variations in binarization thresholding and brightness/contrast adjustment methods cause differences in the trends between unaveraged and averaged choriocapillaris images.

CLAHE = contrast-limited adaptive histogram equalization; SD = standard deviation; VAD = vessel area density; VL = vessel length.

and 6). Increasing contrast caused increased variation between binarization thresholding methods in the superficial plexus but did not appear to change the distribution within each method. CLAHE, however, decreased variation both between and within binarization thresholds. Notably, values produced by the local Phansalkar binarization threshold were most susceptible to changes in image contrast, likely because the algorithm is designed for low-contrast images and is susceptible to small changes; as such, the effects of any alterations in contrast are more likely to affect the binarized image.

Finally, this study examined whether variations in binarization thresholding and brightness/contrast changes would impact the directionality of changes in measurements between unaveraged and averaged CC images. It was found that the method of binarization thresholding did change the effect of averaging (Table 3). For example, in the CC, the number of flow deficits increased with averaging using a global mean binarization threshold but decreased with all other binarization thresholds. Brightness/contrast changes also mattered. Increasing contrast or using CLAHE uniquely caused an increase in the number of flow deficits between unaveraged and averaged images. Reviewing published studies, there have similarly been conflicting findings around the effects of en face averaging on flow deficit quantification. Uji and associates[7] found that CC averaging decreased the average flow deficit size and increased the number of flow deficits. More recently, Chu and associates[35] found the opposite: that averaging increased the average size of flow deficits and decreased the number of flow deficits. Based on the present findings, it is possible that the differences in methods used by those 2 studies may partly explain the disparate results. Uji and associates[7] used both CLAHE and a local Phansalkar binarization threshold; in the present study, this combination yielded a decrease in average flow deficit size and an increase in number of flow deficits (Table 3), the same pattern found by Uji and associates.[7] Chu and associates[35] used a more complex image processing method involving global followed by local thresholding, which they have described in some detail previously.[36,37] In the present study, certain combinations of methods (eg, a local mean binarization threshold along with no brightness/contrast changes) did yield results analogous to those from the study by Chu and associates.[35] Overall, our findings refute one argument against the importance of consistent method in comparative OCTA studies, that as long as the same method is applied to the images in each study arm, the results should be the same. This study demonstrates that this is not the case. The analysis method can alter the directionality of trends and not just the absolute values, further underscoring how critical it is to have some consistency in method between studies.

There were several limitations to our study. First, only normal 3- × 3-mm swept-source OCTA images from a single machine were used. The results and conclusions may be different for OCTA images acquired from a different manufacturer's OCT instrument, with eyes characterized by pathology, larger images, or spectral domain OCTA images. However, the fact that image analysis method impacts results is clearly demonstrated. Also, only 5 binarization thresholding methods were used, as discussed earlier. There is an even wider range of methods available and deployed in OCTA research. Finally, the analysis was intentionally limited in scope; only the overlapping effects of

binarization thresholding and brightness/contrast changes on VAD were assessed. A future study could fully examine the combined effect of binarization thresholding and brightness/contrast changes on all commonly used measurements across multiple plexuses.

In conclusion, the method of OCTA image binarization thresholding and brightness/contrast adjustment significantly alters quantitative measurements and can even affect the directionality of trends. Contrast adjustments, including CLAHE, significantly alter imaging measurements. Perhaps most importantly, this study suggests that results obtained from different OCTA binarization methods should be seen as valid only for that given method. This has significant consequences for clinical trials using OCTA measurements as outcome measurements. The ophthalmology research community needs to come to some consensus around a consistent method for analyzing OCTA images.

ACKNOWLEDGEMENTS

ALL AUTHORS HAVE COMPLETED AND SUBMITTED THE ICMJE FORM FOR DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST and none were reported. This work was supported by the Macula Vision Research Foundation, West Conshohocken, Pennsylvania; the Massachusetts Lions Clubs, Belmont, Massachusetts; the National Institutes of Health, Bethesda, Maryland (grant 5-R01-EY011289-31); the US Air Force Office of Scientific Research, Wright-Patterson Air Force Base, Ohio (grant FA9550-15-1-0473); the Champalimaud Vision Award, Lisbon, Portugal; the Beckman-Argyros Award in Vision Research, Irvine, California; and the Yale School of Medicine Medical Student Fellowship, New Haven, Connecticut. The authors were supported by Topcon Medical Systems,Inc. (A.I., N.K.W.), (S), Nidek Medical Products,Inc. (G.A., N.K.W.) (S); Macula Vision Research Foundation (N.K.W.) (F), Optovue,Inc. (N.K.W., J.S.D.) (C), and Carl Zeiss Meditec. (N.K.W., J.S.D.). The other authors have no proprietary/financial interest.

REFERENCES

1. Huang D, Swanson EA, Lin CP, et al. Optical coherence tomography. *Science* 1991;254:1178–1181.
2. Spaide RF, Fujimoto JG, Waheed NK, Sadda SR, Staurenghi G. Optical coherence tomography angiography. *Prog Retin Eye Res* 2018;64:1–55.
3. de Carlo TE, Romano A, Waheed NK, Duker JS. A review of optical coherence tomography angiography (OCTA). *Int J Retina Vitreous* 2015;1.
4. Al-Sheikh M, Phasukkijwatana N, Dolz-Marco R, et al. Quantitative OCT angiography of the retinal microvasculature and the choriocapillaris in myopic eyes. *Invest Ophthalmol Vis Sci* 2017;58:2063–2069.
5. Ferrara D, Waheed NK, Duker JS. Investigating the choriocapillaris and choroidal vasculature with new optical coherence tomography technologies. *Prog Retin Eye Res* 2016;52:130–155.
6. Spaide RF. Choriocapillaris flow features follow a power law distribution: implications for characterization and mechanisms of disease progression. *Am J Ophthalmol* 2016;170:58–67.
7. Uji A, Balasubramanian S, Lei J, Baghdasaryan E, Al-Sheikh M, Sadda SR. Choriocapillaris imaging using multiple en face optical coherence tomography angiography image averaging. *JAMA Ophthalmol* 2017;135:1197–1204.
8. Wang Q, Chan S, Yang JY, et al. Vascular density in retina and choriocapillaris as measured by optical coherence tomography angiography. *Am J Ophthalmol* 2016;168:95–109.
9. Zhang Q, Shi Y, Zhou H, et al. Accurate estimation of choriocapillaris flow deficits beyond normal intercapillary spacing with swept source OCT angiography. *Quant Imaging Med Surg* 2018;8:658–666.
10. Cole ED, Moult EM, Dang S, et al. The definition, rationale, and effects of thresholding in OCT angiography. *Ophthalmol Retina* 2017;1:435–447.
11. ImageJ. Auto Threshold, 2017. Available at: https://imagej.net/Auto_Threshold. Accessed May 30, 2019.
12. Lowe DG. Distinctive image features from scale-invariant keypoints. *Int J Comput Vis* 2004;60:91–110.
13. Brown M, Szeliski R, Winder S. Multi-image matching using multi-scale oriented patches, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, CA, USA: IEEE; 2005.
14. Fischler MA, Bolles RC. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun ACM* 1981;24:381–395.
15. Arganda-Carreras I, Sorzano COS, Marabini R, et al. Consistent and elastic registration of histological sections using vector-spline regularization. In: Beichel RR, Sonka M, eds. Computer Vision Approaches to Medical Image Analysis. Berlin: Springer; 2006:85–95.
16. Uji A, Balasubramanian S, Lei J, Baghdasaryan E, Al-Sheikh M, Sadda SR. Impact of multiple en face image averaging on quantitative assessment from optical coherence tomography angiography images. *Ophthalmology* 2017;124:944–952.
17. Ridler T, Calvard S. Picture thresholding using an iterative selection method. *IEEE Trans Syst Man Cybern* 1978;8:630–632.
18. Tang FY, Ng DS, Lam A, et al. Determinants of quantitative optical coherence tomography angiography measurements in patients with diabetes. *Sci Rep* 2017;7:2575.
19. Phansalkar N, More S, Sabale A, Joshi M. Adaptive local thresholding for detection of nuclei in diversity stained cytology images. *Proc IEEE Int Conf Comm Signal Process* 2011;218–220.
20. Kim T-H, Son T, Lu Y, Alam M, Yao X. Comparative optical coherence tomography angiography of wild-type and rd10 mouse retinas. *Transl Vis Sci Technol* 2018;7:42.

21. Cicinelli MV, Rabiolo A, Marchese A, et al. Choroid morphometric analysis in non-neovascular age-related macular degeneration by means of optical coherence tomography angiography. *Br J Ophthalmol* 2017;101:1193.
22. Ohayon A, Sacconi R, Semoun O, Corbelli E, Souied EH, Querques G. Choroidal neovascular area and vessel density comparison between two swept-source optical coherence tomography angiography devices. *Retina* 2018 [E-Pub ahead of print].
23. Chidambara L, Gadde SGK, Yadav NK, et al. Characteristics and quantification of vascular changes in macular telangiectasia type 2 on optical coherence tomography angiography. *Br J Ophthalmol* 2016;100:1482–1488.
24. Park J-H, Yoo C, Girard MJA, Mari J-M, Kim YY. Peripapillary vessel density in glaucomatous eyes: comparison between pseudoexfoliation glaucoma and primary open-angle glaucoma. *J Glaucoma* 2018;27:1009–1016.
25. Kim AY, Chu Z, Shahidzadeh A, Wang RK, Puliafito CA, Kashani AH. Quantifying microvascular density and morphology in diabetic retinopathy using spectral-domain optical coherence tomography angiography. *Invest Ophthalmol Vis Sci* 2016;57:362–370.
26. Ashraf M, Nesper PL, Jampol LM, Yu F, Fawzi AA. Statistical model of optical coherence tomography angiography parameters that correlate with severity of diabetic retinopathy. *Invest Ophthalmol Vis Sci* 2018;59:4292–4298.
27. Milani P, Montesano G, Rossetti L, Bergamini F, Pece A. Vessel density, retinal thickness, and choriocapillaris vascular flow in myopic eyes on OCT angiography. *Graefe's Arch Clin Exp Ophthalmol* 2018;256:1419–1427.
28. Mastropasqua R, Toto L, Di Antonio L, et al. Optical coherence tomography angiography microvascular findings in macular edema due to central and branch retinal vein occlusions. *Sci Rep* 2017;7:40763.
29. Kaizu Y, Nakao S, Sekiryu H, et al. Retinal flow density by optical coherence tomography angiography is useful for detection of nonperfused areas in diabetic retinopathy. *Graefe's Arch Clin Exp Ophthalmol* 2018;256:2275–2282.
30. Iafe NA, Phasukkijwatana N, Chen X, Sarraf D. Retinal capillary density and foveal avascular zone area are age-dependent: quantitative analysis using optical coherence tomography angiography. *Invest Ophthalmol Vis Sci* 2016;57:5780–5787.
31. Durbin MK, An L, Shemonski ND, et al. Quantification of retinal microvascular density in optical coherence tomographic angiography images in diabetic retinopathy. *JAMA Ophthalmol* 2017;135.
32. Spaide RF. Visualization of the posterior vitreous with dynamic focusing and windowed averaging swept source optical coherence tomography. *Am J Ophthalmol* 2014;158:1267–1274.
33. Spraul CW, Lang GE, Lang GK, Grossniklaus HE. Morphometric changes of the choriocapillaris and the choroidal vasculature in eyes with advanced glaucomatous changes. *Vision Res* 2002;42:923–932.
34. Yoneya S, Tso MO. Angioarchitecture of the human choroid. *Arch Ophthalmol* 1987;105:681–687.
35. Chu Z, Zhou H, Cheng Y, Zhang Q, Wang RK. Improving visualization and quantitative assessment of choriocapillaris with swept source OCTA through registration and averaging applicable to clinical systems. *Sci Rep* 2018;8:16826.
36. Chu Z, Lin J, Gao C, et al. Quantitative assessment of the retinal microvasculature using optical coherence tomography angiography. *J Biomed Opt* 2016;21:66008.
37. Reif R, Qin J, An L, Zhi Z, Dziennis S, Wang R. Quantifying optical microangiography images obtained from a spectral domain optical coherence tomography system. *Int J Biomed Imaging* 2012;2012:11.