Cornea & External Disease

Automated Image Threshold Method Comparison for Conjunctival Vessel Quantification on Optical Coherence Tomography Angiography

William W. Binotti[1,2], Daniel Saukkonen[1,2], Yashar Seyed-Razavi[1], Arsia Jamali[1], and Pedram Hamrah[1,2]

[1] Center for Translational Ocular Immunology, Tufts Medical Center, Tufts School of Medicine, Boston, MA, USA
[2] Cornea Department, New England Eye Center, Tufts Medical Center, Tufts School of Medicine, Boston, MA, USA

Correspondence: Pedram Hamrah, New England Eye Center, Tufts Medical Center, 800 Washington St., Boston, MA 02111, USA. e-mail: phamrah@tuftsmedicalcenter.org

Purpose: To determine the impact of image binarization and the best thresholding method for conjunctival optical coherence tomography angiography (OCTA).

Methods: Vessel density (VD) of 14 OCTA conjunctival images (nine nasal and five temporal conjunctivas, and eight right and six left eyes) from normal subjects was analyzed. The binarization of gold-standard images, created by removing pixels that do not represent vessels on ImageJ software, was assessed by three masked graders to determine consistency of VD for images. Various thresholding methods on ImageJ, including manual, 1-, 2- and 3-step processes, were performed on unprocessed images for comparison. Interclass correlation coefficient (ICC) ≥ 0.750 were classified as good reliability and selected for calculation of the performance of the pixel location in the binarized images of each method.

Results: Analysis of the gold-standard threshold method achieved an ICC of 0.816 with excellent agreement ($R^2 = 0.965$, $P < 0.001$). From a total 28 different methods and variations performed, only nine methods performed with good reliability, including two 1-step thresholds, six 2-step thresholds, and one 3-step threshold method. Overall, 2-step threshold methods were more reliable than 3-step threshold methods. The 2-step method of Bandpass filter + Phansalkar local threshold (LT) showed the best performance with mean pixel accuracy of 86.9% ± 6.8%, area under the curve of 0.826, sensitivity of 79.0%, and specificity 86.1%.

Conclusions: Bandpass filter + Phansalkar LT was the best method for VD measurement in conjunctival OCTA. Most commonly reported threshold methods showed unsatisfactory agreement. There is a need in the OCTA field for a standardized method to allow comparison between different studies.

Translational Relevance: The proposed threshold method using a widely accessible and commonly used software provides an accurate VD measurement for future OCTA studies.

Introduction

The eye receives a high volume of blood flow, where many microangiopathic pathologies can manifest. Quantifying capillaries and vessels in ocular tissue has been an important biomarker to identify healthy from diseased eyes.[1–4] Optical coherence tomography angiography (OCTA) has become a mainstay in research and is growing for clinical practice and clinical trials.[5–9] This technology allows noninvasive flow detection of blood vessels and is able to generate volumetric angiograms at micrometer resolution.[10,11] In recent years, OCTA has expanded to the anterior segment and has been shown to be useful in a variety of diseases, such as corneal neovascularization, limbal stem cell deficiency, ocular tumors, and pterygium.[12–20] The bulbar conjunctival vessels have been the focus of

many studies for in vivo and noninvasive assessment of vascular parameters, because of its ease of access and capacity of reflecting ocular and systemic diseases.[21–24]

Despite the advances in vessel imaging technology, there are a wide range of image thresholding methods for vessel measurement without a consistent and standardized method for processing OCTA images and quantifying vessel parameters. Image thresholding or binarization is the process that converts a gray value image into pixels above a determined threshold value in "white" (1) and below the value in "black" (0). There are many methods in determining the threshold value of the image, for example a global threshold applies a single value over the entire image.[25,26] Another method is using a local or adaptive threshold that applies different values according to regional variations of the image.[26,27] Also, various image processing techniques have been used to remove background noise and improve image contrast, whether it is smoothing the image or applying different filters to remove small particles and noise.[25,28] Moreover, many authors propose a combination of the different methods for thresholding.[25,28]

The most commonly used vessel parameter in OCTA is vessel density (VD), which quantifies the percentage of pixels that represent blood flow compared to the total pixels in the entire image. Consistency and reproducibility of these measurements are important for research studies. However, the method of thresholding has been shown to significantly alter their measurements in retinal and choroidal OCTA.[29,30] Similar to the posterior segment, the anterior segment lacks a standardized threshold method for quantitative OCTA assessment. With increasing studies using OCTA for analysis, there is a need to better understand the impact of different threshold methods on vessel measurements and to establish a consistent and accurate method for binarization of vessels in OCTA images.[29,30] We hypothesize that different threshold methods will affect the VD measurements in anterior segment OCTA. Accordingly, the aim of this study was to assess the agreement and accuracy of different threshold methods in quantifying VD on OCTA images of the conjunctiva, by comparing to their reference gold standard images.

Methods

Images from healthy volunteers were selected from the OCT database at the Cornea Department of New England Eye Center, Department of Ophthalmology, Tufts Medical Center, Boston, MA. Subjects were retrospectively reviewed, and images were de-identified following the tenets of the Declaration of Helsinki and the study was approved by Tufts' Institutional Review Board/Ethics committee (IRB no. 12530). Inclusion criteria was subjects with conjunctival images who performed the OCTA protocol adapted for anterior segment in the OCT database. Exclusion criteria were any ocular diseases, including dry eye disease, systemic diseases (i.e., diabetes, hypertension), ocular infection, history of recent ocular surgery, and images of poor quality with excessive noise and low clarity that would impede conjunctival vessel visualization. In contrast to previous methods established on a single OCTA image (generally the best image example), multiple images were used for threshold comparison to better represent the variability of gray value that occurs between OCTA images.

OCTA Imaging Technique

The OCTA images were acquired by the spectral domain OCTA system, Avanti XR AngioVue (Optovue Inc., Freemont, CA, USA; version 2018.0.0.10). The 6 × 6 mm HD Retina scan mode was used with the Optovue anterior segment lens (Long Cornea Adaptive Module) and manual adjustments of Z motor, P motor, and focus to image the conjunctival surface. The system acquires volumetric scans of 400 × 400 A-scans at 70,000 A-scans per second, using a light source centered on 840 nm and a bandwidth of 45 nm. Flow is detected through motion contrast of repeated A-scans at the same location, and motion artifact was removed by 3D orthogonal registration and merging of two scans using the SSADA algorithm, thus producing a 5 μm axial resolution volumetric angiogram with the motion of the red blood cells within vessels.

Each image was exported in their 400 × 400 pixel format, analyzed and processed on open source software FIJI from ImageJ (National Institutes of Health, Bethesda, MD, USA). Because OCTA is a novel and adapted technology from the retinal to the anterior segment, images of the conjunctiva can present with variable background noise as compared to retinal images. To focus the analysis on the conjunctiva, the eyelids, eyelashes and cornea were carefully removed from all OCTA images before thresholding.

Establishing Gold Standard Images

Pixels that represented noise or imaging artifacts and clearly not a vascular component were manually removed in the selected images by an experienced OCTA grader (Figs. 1A, 1B). These images were then used as reference for subsequent thresholding
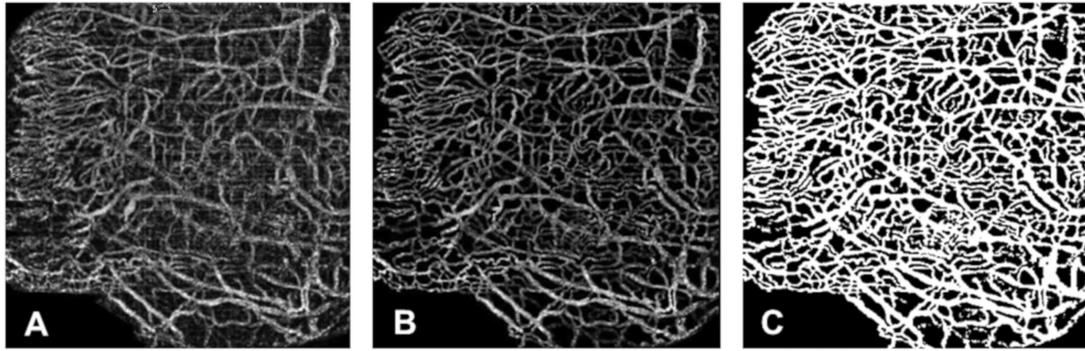
Figure 1. (**A**) OCTA RAW image of the nasal conjunctiva. (**B**) Processed image removing background noise and pixels that do not represent vessels. (**C**) Manually thresholded OCTA binary image.

processes. Next, the manual threshold was determined by three blinded and independent graders (W.B., D.S., Y.S.R.) with the aim of matching their respective reference image (Fig. 1C). After binarization (thresholding), the number of foreground pixels representing blood flow divided by the total pixels were represented as VD (% area). Each image generated from the mean threshold value of the three graders and its VD was considered the gold standard for the methodology comparison.

Thresholding Methods

The following thresholding methods to acquire binary images for vessel quantification were selected based on previously published methods and applied to the unprocessed RAW images. The available methods consist of a variation of algorithms that transform a gray-value image into a binary image through a 1-step threshold, or a combination of image pre-processing and filters before the threshold (2-step or 3-step threshold methods). The additional steps before thresholding seek to reduce background noise, to remove pixels that do not represent vessels, or to create continuity of neighboring vessel pixels. Each VD from the thresholded images were compared to their respective gold standard image VD.

1-Step Threshold Methods

Manual Threshold

The same three blinded graders independently set the threshold on the RAW images according to their interpretation of the best representation of the vessels. Their threshold was applied to the entire (global) image. Their averaged VD for each image was used to compare to the gold standard.

Global Threshold

Global threshold applies a single threshold value to all pixels (global) within the entire image into either foreground pixels (1) or background pixels (0). The threshold value can be achieved by many methods (e.g., mean gray value of the image or Otsu algorithm), which compares the gray value variance of each pixel from the average gray value to determine the optimal cutoff value for the entire image (global) by minimizing the intraclass intensity variance.[26]

Local Threshold

Local threshold applies different local thresholds based on the neighboring gray values of each individual pixel, accounting for focal high-variability gray value in the image. Also known as an adaptive threshold method. Algorithms may use the mean or median gray values, Otsu (as previously described) or Phansalkar, which applies different local thresholds based on a combination of mean, standard deviation, and normalization of gray value within a 15-pixel radius. The latter is a method particularly used in low contrast images.[27]

2-Step Threshold Methods

With the goal of creating threshold images that best represent the structure of interest (vessels in this case), many advocate a combination of filters or processing methods to enhance the structure of interest and minimize noise.[25,28,31–33] Therefore a selection of commonly used combined methods were performed and are schematically represented in Figures 2 and 3.

Close Morphology With a One- or Two-Pixel Radius ± Local Threshold

First, by applying a closing morphology (image dilation followed by erosion), the algorithm "fills" the
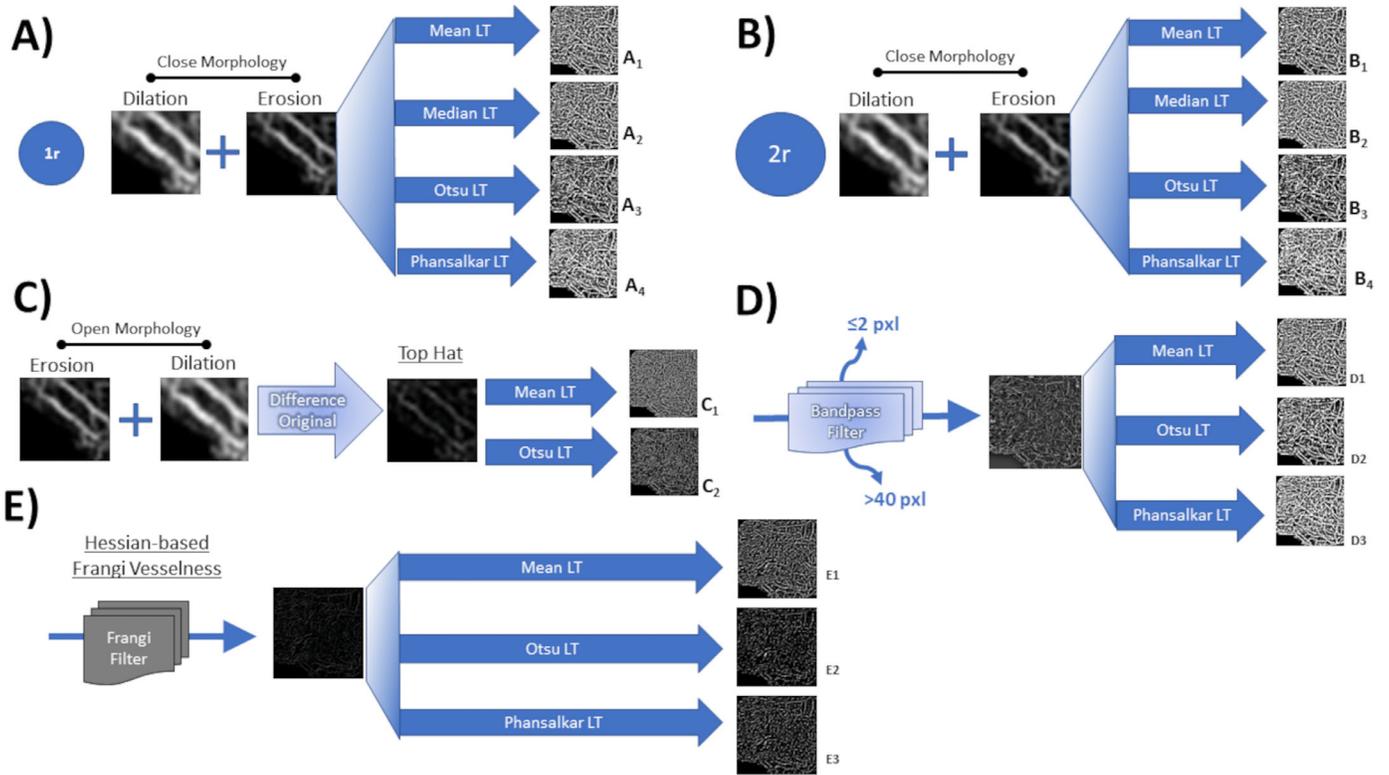
Figure 2. Diagram illustrating the different combined threshold methods. (A_{1-4}) A closing morphology filter of a 1-pixel radius was applied; then, 4 different local threshold methods were analyzed, mean, median, Otsu and Phansalkar. (B_{1-4}) A closing morphology filter of a two-pixel radius was applied; then, four different local threshold methods were analyzed, mean, median, Otsu, and Phansalkar. (C_{1-2}) Top hat filter followed by a mean or an Otsu local threshold was applied. (D_{1-3}) A bandpass filter that removes pixels ≤ 2 pixels and >40 pixels followed by a mean or Otsu or Phansalkar threshold was applied. (E_{1-3}) A Frangi filter to identify vessel-like structures was applied, followed by mean or Otsu or Phansalkar local threshold.
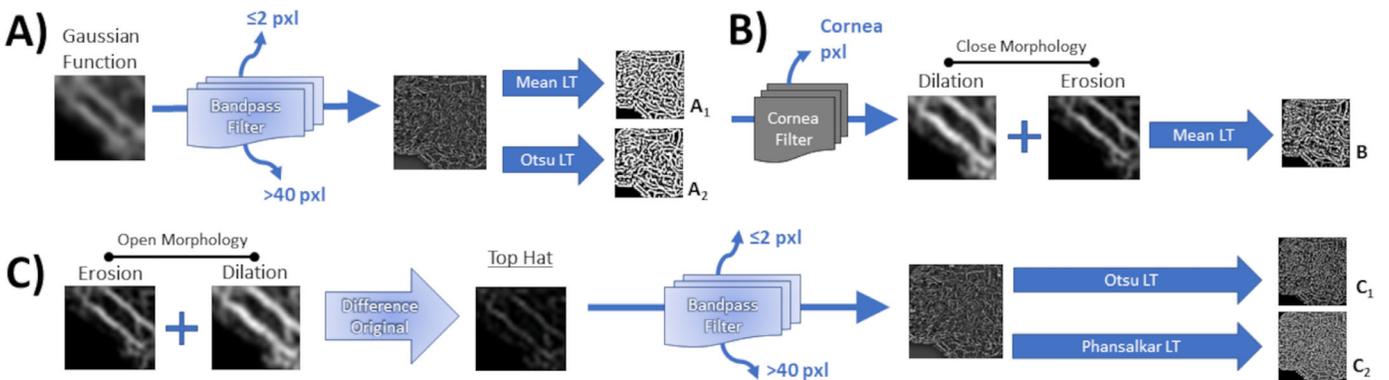
Figure 3. Diagram illustrating the different combined threshold methods. (A_{1-2}) Gaussian smoothing filter with a sigma radius of four pixels was applied; then, a bandpass filter that removes pixels ≤ 2 pixels and >40 pixels followed by a mean or an Otsu local threshold was applied. (B) The average gray value of three corneal measurements was used to threshold image; then a close morphology followed by a mean threshold was applied. (C_{1-2}) A top hat filter, then a bandpass filter followed an Otsu or a Phansalkar threshold was applied.

spaces between noncontiguous pixels within a one- or two-pixel radius of the neighboring pixels (removing "holes"). Then, four different local threshold methods were analyzed, mean, median, Otsu and Phansalkar (Fig. 2A_{1-4} and Fig. 2B_{1-4}, respectively), based on previ-

ous descriptions.[25,28,31] Close morphology is typically used in vascular imaging to improve pixel connectivity of vessel structures on threshold. One-pixel and two-pixel radius algorithm were compared to assess whether the size of the radius improved the method.

Top Hat Filter ± Local Threshold

Top hat filter consists of performing a morphological opening (image erosion followed by dilation) and then subtracting from the original image. This process removes particles that are smaller (pixels) than the structuring element and darker (lower gray value) than their surroundings; intended to remove the image noise, correct nonuniform lighting, and improve background contrast. Last, a mean and an Otsu local threshold was applied (Fig. 2C[1-2], respectively).[28,32,34]

Bandpass Filter ± Local Threshold

A bandpass filter removes structures ≤2 pixels and ≥50 pixels, which removes the image noise and improved background contrast. No morphological or top hat processing was applied. Then, a local threshold method was applied, mean, Otsu, and Phansalkar (Fig. 2D[1-3], respectively), as previously described. This combination was included to compare the effect of not performing the commonly used top hat filter for image analysis.[23,32]

Frangi Vesselness Filter ± Local Threshold

Hessian-based Frangi vesselness filter is an algorithm that is designed to enhance visualization of vessels on images by identifying multiscale pixels with vessel- or tube-like characteristics and enhancing vessel contrast and suppressing background noise. This method is commonly used in ultrasound, computerized tomography, magnetic resonance, and retinal photography vessel imaging.[35,36] Last, local thresholds was applied as previously described (Fig. 2E[1-3]).

3-Step Threshold Methods

Bandpass Filter ± Gaussian Blur ± Local Threshold

First, the image is processed through a Gaussian blur (aka smoothing) with a sigma radius of four pixels, which is used to reduce overall image noise.[37] Then, a Bandpass filter is applied that removes background pixels. Last, a mean and an Otsu local threshold was applied (Fig. 3A[1-2], respectively).[17,34,38]

Cornea Global Filter ± Close Morphology ± Mean Local Threshold

Based on studies that use the foveal avascular zone to calculate the signal to noise ratio and determine the global threshold, the cornea (avascular in healthy subjects) was used for the global image filter.[16,32,39] The average gray value of three measures of the cornea was used to filter pixels below that gray value. Subsequently, a close morphology was performed to fill "holes" within a 2-pixel radius, as previously described. Last, a mean local threshold was applied (Fig. 3B).[32]

Bandpass Filter ± Top Hat Filter ± local Threshold

A Bandpass filter was applied to remove small particles and background noise, then a top hat filter, as described previously. Last, an Ostu or Phansalkar local threshold was performed for binarization (Fig. 3C[1-2]).[23,32,39]

Performance of the Methods

The final step of establishing the best threshold method is to determine not only the number of pixels thresholded as vessels in an image, i.e. vessel density (VD), but additionally the correct location of these pixels representing vessels, compared to the gold standard image. Therefore pixels of vascular flow (foreground pixels) that were present both on the image of interest and its gold standard were calculated as the true positives. Pixels that were included on the image of interest, but not on its gold standard, were calculated as false positives. Pixels that were included on the gold standard, but not on the image of interest, were calculated as false negatives. Pixels that were counted as non-vessels (background pixels) in both the image of interest and its gold standard were considered true negatives. Therefore the accuracy of the correct location of thresholded pixels were represented by adding the true-positive and the true-negative values (%) of each image then averaging their accuracy.

Statistical Analysis

The Statistical Package for the Social Sciences software (ver. 17; SPSS Inc., Chicago, IL, USA) was used to analyze the data. Shapiro-Wilk test was performed to determine if the data was normally distributed. For the gold standard validation, the average between grader 2 and 3 were compared to grader 1 and their difference used to construct Bland-Altman plots and ICC test was performed to demonstrate the variability of the graders' measures. Then, a linear regression was calculated and R^2 for goodness-of-fit determined to demonstrate the association of the VD calculated from the new gold standard images with the VD calculated from manually-removed pixel images. Shapiro-Wilk normality test was performed on the mean standardized residuals to determine bias of the linear regression model. Bias was considered if $P < 0.05$.

Similarly, for the manual thresholding on RAW images, the average between graders 2 and 3 were compared to grader and the average of all 3 graders were compared to the gold standard VD. Furthermore, the VDs of each method were compared to the gold standard through ICCs with their 95% confidence

intervals (lower bound − upper bound) and analyzed for absolute agreement. Only methods with good reliability, established as ICC ≥ 0.750,[40] were considered for the pixel-location accuracy assessment. Mean accuracy was calculated by adding true positive and true negative values of all images represented with their 95% confidence intervals. Sensitivity was the calculated by dividing true positive by the sum of true positives and false negatives. Specificity was calculated by dividing the true negatives by the sum of true negatives and false positives. Area under the curve (AUC) was obtained by the average of sensitivity and specificity values.

Results

A total of 14 images of 8 subjects were included and carefully selected to represent the variations of the conjunctival vessel location (i.e., nasal and temporal) and image characteristics, as shown in Table 1. The images that were meticulously assessed to remove the non-flow (background) pixels for the gold standard validation underwent binarization by 3 graders, their ICC was 0.816 [0.519–0.937] and the Bland-Altman plot showed a good agreement with less than ±10% variability between graders' measurements and no significant bias (Fig. 4A). When comparing the new gold standard images created from the grader's threshold and the manually-removed pixel images, the VD agreement was 96.5% ($R^2 = 0.965$, $P < 0.001$) with no significant bias (mean error <0.01 ± 0.96, $P = 0.605$), as shown in Figure 4B.

The ICC values and their 95% confidence interval of the methodologies are summarized in Table 2. The repeatability of manual thresholding performed by 3 graders showed a poor agreement between them (ICC = 0.482 [0.032–0.798]) and a moderate agreement compared to the gold standard (ICC = 0.525 [−0.090 to 0.839]; Table 2). Overall, the local (adaptive) threshold methods showed better agreement compared to the global threshold and the mean threshold better agreement compared to Otsu threshold methods. Of interest, the mean and Phansalkar local threshold methods showed a good agreement and met the criteria for further accuracy assessment (ICC = 0.836 and ICC = 0.831, respectively). Conversely, the mean global threshold showed very poor agreement (ICC = 0.245).

The two-step methods that met the criteria for further accuracy assessment were the close morphology with 1-pixel radius and mean local threshold (ICC = 0.838; Fig. 2A$_1$), 1-pixel radius and Phansalkar local threshold (ICC = 0.822; Fig. 2A$_4$), close morphology with 2-pixel radius and mean local threshold (ICC = 0.841; Fig. 2B$_1$), close morphology with 2-pixel radius and Phansalkar local threshold (ICC = 0.763; Fig. 2B$_4$), Bandpass filter with mean local threshold (ICC = 0.835; Fig. 2D$_1$), Bandpass filter with Phansalkar threshold (ICC = 0.832; Fig. 2D$_3$), and Bandpass filter with Gaussian filter and Otsu local threshold (ICC = 0.762; Fig. 3A$_2$). Conversely, the top hat filter with Otsu local threshold (Fig. 2C$_2$) and the top hat filter, bandpass filter with Otsu local threshold (Fig. 3C$_1$) showed the worst agreement (ICC = 0.136 and ICC = 0.142, respectively). The complete list of other

Table 1. General Overview and Pixel Specifications of the Images

	Subject	Eye	Location	Mean Gray Value	Gray Value SD	Median Gray Value	Gray Value Skewness	Gray Value Range
Image 1	A	Right	Nasal	72.9	49.9	66.0	0.6	0-255
Image 2	A	Right	Temporal	69.5	50.6	59.0	0.8	0-255
Image 3	B	Left	Nasal	53.3	49.9	43.0	0.9	0-255
Image 4	C	Right	Temporal	49.1	49.3	38.0	1.0	0-255
Image 5	D	Right	Nasal	52.1	63.8	27.0	1.3	0-255
Image 6	D	Right	Temporal	55.5	55.8	41.0	1.1	0-255
Image 7	D	Left	Nasal	53.0	64.4	27.0	1.3	0-255
Image 8	D	Left	Temporal	60.6	60.5	44.0	1.2	0-255
Image 9	E	Right	Nasal	66.9	50.5	57.0	0.8	0-255
Image 10	F	Left	Nasal	49.5	51.4	33.0	1.2	0-255
Image 11	G	Right	Nasal	42.0	45.3	28.0	1.2	0-255
Image 12	G	Left	Nasal	38.6	51.5	18.0	1.7	0-255
Image 13	G	Left	Temporal	41.3	46.8	26.0	1.0	0-255
Image 14	H	Right	Nasal	50.7	47.6	40.0	1.0	0-255

Standard deviation, SD.
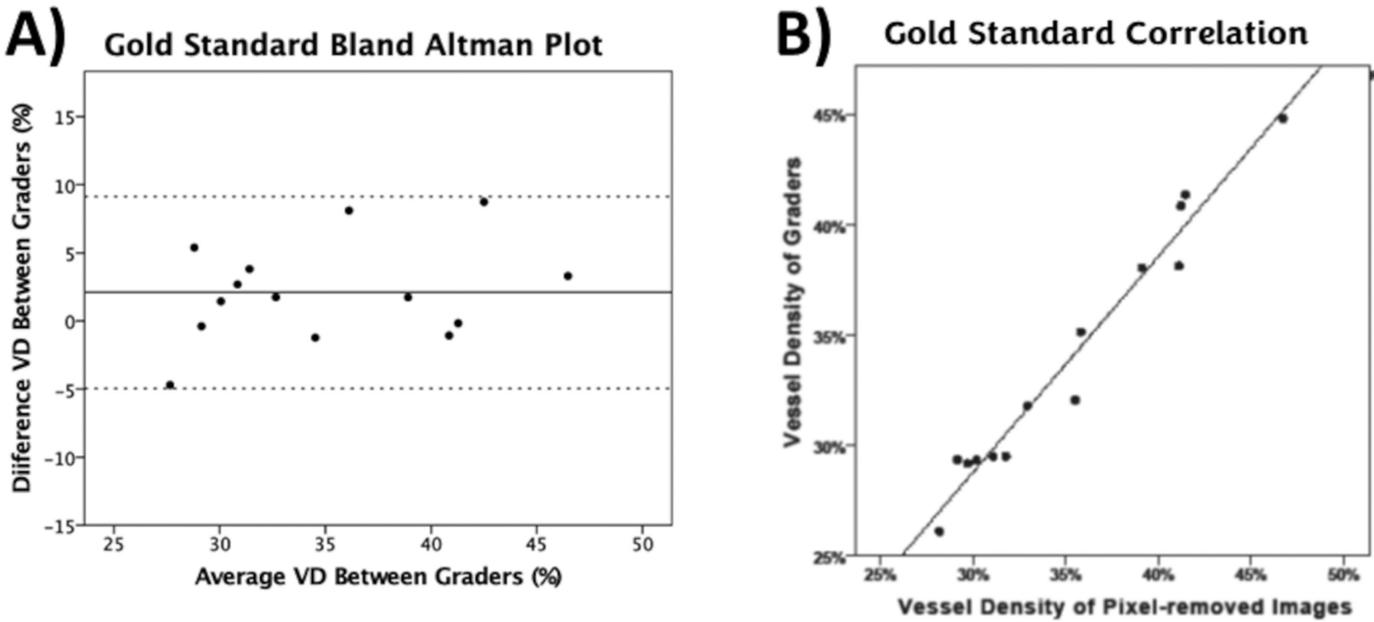
Figure 4. (A) Bland-Altman plots with the 95% upper and lower limits of the difference between the vessel density (VD) measurements between the graders. (B) The goodness-of-fit correlation between the VD measures of the three graders and the pixel-removed images for the gold-standard establishment.

methods that were tested but did not show good reliability are demonstrated as Supplemental Table S1, In general, the mean local threshold methods outperformed the Otsu and median local threshold methods. The latter methods, noticeably removed excessive amount of pixels overall.

The performance details of the high reliability threshold methods to assess the pixel-location of the

Table 2. Reliability of the Threshold Methodologies

Threshold Methodologies	Mean Gray Value VD (%)	±SD Gray Value VD (%)	ICC	95% CI Lower Bound	Upper Bound
1-Step Threshold Methods					
Manual Threshold	29.1	±4.9	0.525	−0.090	0.839
Mean Global Threshold	40.3	±2.7	0.245	−0.117	0.636
Otsu Global Threshold	29.7	±4.7	0.463	−0.056	0.791
Mean Local Threshold	34.3	±5.0	**0.836**	0.567	0.944
Otsu Local Threshold	28.8	±5.2	0.516	−0.099	0.836
Phansalkar Local Threshold	34.0	±6.1	**0.831**	0.547	0.943
2-Step Threshold Methods					
Close Morphology (1pxl) + Mean Local Threshold	34.1	±4.9	**0.838**	0.567	0.945
Close Morphology (1pxl) + Phansalkar Local Threshold	34.6	±6.2	**0.822**	0.541	0.939
Close Morphology (2pxl) + Mean Local Threshold	34.2	±4.8	**0.841**	0.574	0.946
Close Morphology (2pxl) + Phansalkar Local Threshold	36.3	±6.6	**0.763**	0.352	0.920
Bandpass Filter + Mean Local Threshold	34.7	±5.0	**0.835**	0.574	0.943
Bandpass Filter + Phansalkar Local Threshold	37.9	±6.0	**0.832**	0.556	0.943
Frangi Filter + Mean Local Threshold	49.1	±7.2	0.679	0.253	0.884
3-Step Threshold Methods					
Top hat Filter + Mean Local Threshold	29.7	±4.2	0.604	−0.086	0.882
Bandpass Filter + Gaussian Blur + Otsu Local Threshold	36.5	±4.9	**0.762**	0.213	0.927
Cornea Filter + Close Morphology + Mean Local Threshold	24.7	±5.3	0.316	−0.074	0.727
Gold Standard Images	33.9	±5.7	—	—	—

SD, standard deviation; pxl, pixel.
Methodologies with good reliability, ICC ≥0.750 (bold), that were selected for pxl location accuracy assessment.

Table 3. Performance of the High Reliability Threshold Methodologies

Threshold Methodologies	AUC	Sensitivity (%)	Specificity (%)	Accuracy (%)	Lower 95% CI	Upper 95% CI
1-Step Threshold Methods						
Mean Local Threshold	0.799	72.6	87.2	82.4	77.9	86.7
Phansalkar Local Threshold	0.808	73.5	88.1	82.9	78.7	87.3
2-Step Threshold Methods						
Close Morphology (1pxl) + Mean Local Threshold	0.804	72.8	88.0	82.7	78.4	87.0
Close Morphology (1pxl) + Phansalkar Local Threshold	0.728	57.9	87.8	73.8	68.8	78.8
Close Morphology (2pxl) + Mean Local Threshold	0.776	69.1	86.1	80.1	74.2	86.1
Close Morphology (2pxl) + Phansalkar Local Threshold	0.723	58.2	86.4	74.3	69.7	78.9
Bandpass Filter + Mean Local Threshold	0.809	74.2	87.6	83.5	79.2	87.8
Bandpass Filter + Phansalkar Local Threshold	0.826	79.0	86.1	86.9	83.0	91.0
3-Step Threshold Methods						
Bandpass Filter + Gaussian Blur + Otsu Local Threshold	0.723	64.6	80.0	76.5	70.7	82.4

Detailed performance of the pixel-location assessment of each methodology with averaged values from all images (e.g., AUC, sensitivity, specificity, accuracy, and 95% CI).
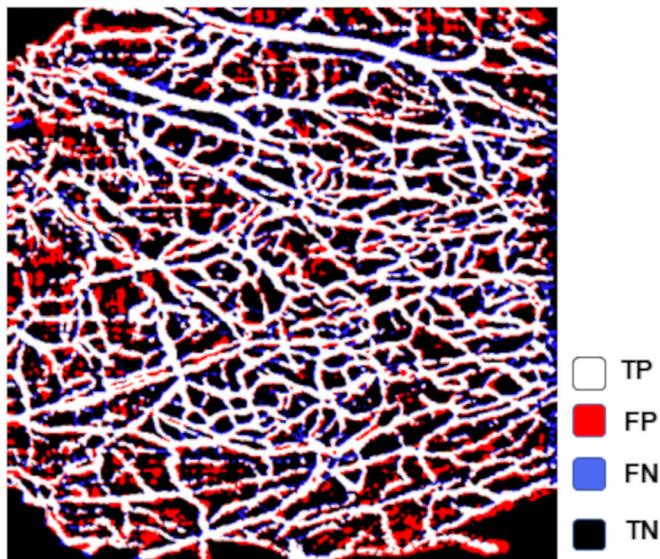
Figure 5. Representative OCTA image showing the overlapping pixels of the Bandpass filter + Phansalkar local threshold with the pixels from its gold-standard image. The *white pixels* represent true-positives (*TP*), the *red* false-positives (*FP*), the *blue* false-negatives (*FN*), and the *black* true-negatives (*TN*).

14 sample images are shown in Table 3. The bandpass filter with Phansalkar local threshold method showed the highest accuracy in identifying the correct location of pixels that represented flow in the gold standard images (86.9% ± 6.8%, range 71.3%–94.1%), as shown in Figure 5. Interestingly, although the combined method using a 2-pixel radius close morphology with mean local threshold showed a higher ICC value for VD measurement compared to its 1-pixel radius counterpart (ICC = 0.841 and ICC = 0.838, respectively), the latter was more accurate in identifying the correct blood flow pixels; mean accuracy of 82.9%

± 7.5%, compared to 81.2% ± 7.5% from its 2-pixel counterpart.

Discussion

This study highlights that only select image threshold methods commonly used and published in the literature show a good agreement in measuring VD. Moreover, the combined threshold method using a bandpass filter with a Phansalkar local threshold showed the highest accuracy in identifying pixels representing blood vessels. As OCTA continues to grow in research and clinical trials, and more recently for anterior segment, accurate and reproducible vessel measurement quantification is of paramount importance. This study demonstrates how different image thresholding techniques can significantly alter the quantification of VD, which further highlights the need for a standardization.

There is a wide range of binarization methods in the literature and to date no consensus for image processing and vessel quantification for both posterior and anterior segment. The goal of this study was not to exhaust every available method, rather, to compare commonly reported methods and their combination using a common, widely available, and open-source software. As such, we undertook a stepwise approach to compare the different image processing techniques. Some studies utilize manual thresholding on anterior segment OCTA images.[14,41,42] We showed that the agreement of manual thresholding was unsatisfactory for VD analysis, even though this approach was better than many automated threshold methods (global or local). Moreover, we showed agreement

between graders to be poor with manual thresholding, highlighting the high variability of measures by using 3 graders. We also assessed the agreement between one-step global and local thresholds commonly used in the literature,[6,17,23,32,43] and determined that local threshold outperformed global threshold methods. Accordingly, the local thresholding was used for consecutive combined thresholding methods, in accordance to previous reports.[6,32] Overall, the local threshold generated a more homogenous image by accounting for the local gray value differences in each region. The most commonly used 1-step threshold method in the literature is Otsu (global and local).[15,17,23,43] However, initial studies have shown a poor performance overall of this method, whereas recent approaches suggest Otsu threshold in combination with other image filters to improve its performance.[26,44,45] We found a similar trend in our data, where the 3-step method with Otsu local threshold outperformed its 1-step method.

Next, commonly used combined methods were tested and relevant adaptations were assessed. The most common combined methods utilized in the literature are the Gaussian blur with Bandpass filter and Otsu local threshold (Fig. 3A$_2$)[15,17,33,38] and the top hat filter with a bandpass filter and an Otsu local threshold (Fig. 3C$_1$).[23,43,46] We found that the top hat filter with a bandpass filter followed by an Otsu local threshold showed very poor agreement (ICC = 0.142). Although Gaussian blur with Bandpass filter followed by an Otsu local threshold met the criteria for accuracy assessment (ICC = 0.762) it was significantly outperformed by bandpass filter with Phansalkar local threshold (AUC = 0.732 and AUC = 0.826, respectively).

A top hat filter has been frequently proposed as a method to reduce noise and remove small particles in the image without distortion.[32,43] Although the combination of a top hat filter with a bandpass filter is the most commonly performed,[33,43,46,47] we further explored the impact each filter on the VD agreement. In this study, an isolated top hat filter followed by a local threshold method showed worse agreement compared to the combined top hat and bandpass filters (Table 2). Conversely, the isolated bandpass filter followed by a local threshold method showed better agreement than the combined filters (top hat and bandpass), showing that the latter excessively removed pixels representing vessels. Of note, the combined method using the cornea gray value as a global suppression threshold, simulating the avascular signal-to-noise calculation from the foveal avascular zone in retina OCTA,[39] showed a poor agreement. It is important to highlight that although this approach might be useful for corneal neovascularization thresholding,[16] where the cornea presents a higher signal-to-noise ratio, we found it was not adequate for conjunctival vessel thresholding. A hessian-based

Frangi vesselness filter is commonly used to identify vessel-like structures on a variety of images, including fundus photographs. However, the latter was designed to highlight vessels in noncontrasted images.[35,36] The Frangi filter showed suboptimal reliability given the blood flow in OCTA has already high contrast, increasing the gray value cutoff of the image and thus thresholding many fine capillaries as background pixel.

Image binarization is an important step when quantifying VD in research studies and clinical trials using OCTA. Mehta and associates[30] highlighted that different thresholding methods can significantly alter the absolute values as well as the directionality of trends, which could impact studies assessing quantitative vessel measurements. To date, there is no established parameter to compare and determine the quality of the thresholding methods. An ideal thresholding method is one that is widely available and accessible for research that can accurately binarize all pixels representing blood flow as foreground and all pixels that do not as background. Studies assessing the capability and sensitivity of OCTA in detecting true blood flow in vivo are lacking. One approach, widely accepted in the literature, is the manual removal of pixels on an OCTA image to produce a "gold standard" reference, then determining the parameters of a custom-built combined thresholding process (top hat filter with global threshold and local mean threshold) that more accurately reproduced the vessel measurements.[32] Herein, although we took a similar approach to determine the "gold standard" for comparison, we used several images rather than a single image and performed the binarization on an open-source software widely used in research. Furthermore, we were able to determine the accuracy of each method in binarizing pixels correctly after selecting methods that met the criteria for good agreement. The combined bandpass filter followed by Phansalkar local threshold method showed the highest accuracy in the large subset of images (Fig. 5). Although not every thresholding method was tested in this study, it remains the largest study comparing a wide variety of thresholding methods on conjunctival OCTA images. However, the inherit underlying limitations from the manual adjustments to determine the gold-standard images, which currently lack validation, should be considered. Additionally, not all the available methods were tested. Nevertheless, this study is hopefully a step forward toward a more accurate and accessible thresholding method, highlighting the importance of establishing a standardized image binarization method to allow comparison between future studies and clinical trials.

Many studies within the published literature, interestingly, either forgo mention or do not clearly

report the details of the thresholding method, resulting in difficulty in the interpretation and reproduction of the results. Also, while the use of the OCT device's built-in software is frequently used and advocated,[12,48,49] the inherent variability between devices that can result from distinct algorithm processing steps of each platform cannot guarantee accuracy,[30] especially in anterior segment where there are currently no specifically designed software. Another important point to highlight is that in our study, the cornea and eyelids were digitally removed from all OCTA images to focus on conjunctival vessels. This introduces determining conjunctival boundaries identification and manual selection, which is time consuming. Furthermore, not all individuals have the same interpalpebral exposure for conjunctival imaging, which may be a limitation of studies measuring conjunctival vessel parameters. Future trained algorithms designed to remove these features before image binarization are necessary for fully automated thresholding.

Studies have highlighted OCTA adapted for anterior segment capable of quantifying VD of small capillaries on the cornea and conjunctival better than slit-lamp photography and indocyanine green angiography.[46,50] Conversely, another study demonstrated that OCTA was less precise in detecting small capillaries of corneal neovascularization compared to indocyanine green angiography.[51] It is of note that these imaging techniques are not directly comparable when thresholding and quantifying vessels, given their differences in how flow is detected, the image contrast and resolution. Put differently, the same thresholding method for vessel quantification used in these images will not be optimal for all imaging techniques (i.e., OCTA, fluorescent angiography, color photography) and anatomical structures (i.e., macula, optic nerve, conjunctiva, cornea), as highlighted in our study. Accordingly, we believe that a consensus in OCTA thresholding and analysis for each anatomical structure is necessary for accurate vessel assessment.

There were several inherit limitations to our study. First, as previously discussed, is the lack of a gold standard. The manually created "gold standard" reference images as a means of comparison is suboptimal but necessarily commonly used, considering the available resources and knowledge from previous reports in the literature.[44] However, by removing pixels that represented noise or did not resemble vessels on the OCTA image, in some images, a large number of pixels were excluded because of higher image noise. This further highlights the discrepancy of thresholding methods that generated low VD values (i.e., Otsu

and manual threshold) and explains the high range of the 95% confidence intervals of ICC values; however, it simulates a more realistic assessment of conjunctival OCTA with variable noise. Furthermore, a relatively small number of images was analyzed, and not all threshold possibilities were tested in the current study. Manually removing pixels to create reference images is very time consuming and therefore challenging to create a large dataset of reference images for comparison. Nevertheless, our study power was 100%, assuming alpha of 0.05, for the gold standard validation. Finally, VD was the only vessel parameter studied herein. Although Reif et al.[32] proposed the same combined threshold method for VD and fractal dimension, larger studies are warranted to address the role of thresholding for each vessel parameter.

Collectively, our results show the impact of image thresholding on VD and further highlight the need for a standardization in image thresholding, particularly in anterior segment OCTA because it is a novel field with distinct features that may affect binarization, such as higher signal-to-noise ratio and artifacts compared to better established OCTA imaging for posterior segment.

In conclusion, a combined thresholding method, particularly the bandpass filter with a local Phansalkar threshold, showed a good reliability and higher accuracy for pixel location and overall performance when compared to other combined methods or 1-step threshold methods. Many of the published methods in the literature did not show a satisfactory agreement and did not meet the criteria for accuracy assessment. Thus, the effects of image thresholding should be considered in vessel quantification and interpretation as it can impact future research studies and clinical trials using OCTA.

Acknowledgments

P. Hamrah, "Systems and Methods for Determining Tissue Inflammation Levels" U.S. Patent Application Serial No. 62/913,307 (I); **D. Saukkonen,** None; **Y. Seyed-Razavi,** None; **A. Jamali,** None

References

1. Ohtani N. Laser Doppler flowmetry of the bulbar conjunctiva as a monitor of the cerebral blood flow. *Nihon Kyobu Geka Gakkai Zasshi*. 1996;44:1721–1728.

2. Wong TY, Klein R, Sharrett AR, et al. Retinal arteriolar narrowing and risk of coronary heart disease in men and women. The Atherosclerosis Risk in Communities Study. *JAMA*. 2002;287:1153–1159.

3. Cheung AT, Tomic MM, Chen PC, Miguelino E, Li CS, Devaraj S. Correlation of microvascular abnormalities and endothelial dysfunction in Type-1 Diabetes Mellitus (T1DM): a real-time intravital microscopy study. *Clin Hemorheol Microcirc*. 2009;42:285–295.

4. Yarmohammadi A, Zangwill LM, Diniz-Filho A, et al. Optical coherence tomography angiography vessel density in healthy, glaucoma suspect, and glaucoma eyes. *Invest Ophthalmol Vis Sci*. 2016;57(9):OCT451–OCT459.

5. de Carlo TE, Romano A, Waheed NK, Duker JS. A review of optical coherence tomography angiography (OCTA). *Int J Retina Vitreous*. 2015;1:5.

6. 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.

7. Van Melkebeke L, Barbosa-Breda J, Huygens M, Stalmans I. Optical coherence tomography angiography in glaucoma: a review. *Ophthalmic Res*. 2018;60:139–151.

8. Alagorie AR, Nittala MG, Velaga S, et al. Association of intravitreal aflibercept with optical coherence tomography angiography vessel density in patients with proliferative diabetic retinopathy: a secondary analysis of a randomized clinical trial. *JAMA Ophthalmol*. 2020;138:851–857.

9. Cennamo G, Montorio D, D'Alessandro A, Napolitano P, D'Andrea L, Tranfa F. Prospective study of vessel density by optical coherence tomography angiography after intravitreal bevacizumab in exudative age-related macular degeneration. *Ophthalmol Ther*. 2020;9:77–85.

10. Spaide RF, Fujimoto JG, Waheed NK, Sadda SR, Staurenghi G. Optical coherence tomography angiography. *Prog Retin Eye Res*. 2018;64:1–55.

11. Izatt JA, Hee MR, Swanson EA, et al. Micrometer-scale resolution imaging of the anterior eye in vivo with optical coherence tomography. *Arch Ophthalmol*. 1994;112:1584–1589.

12. Ang M, Cai Y, Shahipasand S, et al. En face optical coherence tomography angiography for corneal neovascularisation. *Br J Ophthalmol*. 2016;100:616–621.

13. Ang M, Sim DA, Keane PA, et al. Optical coherence tomography angiography for anterior segment vasculature imaging. *Ophthalmology*. 2015;122:1740–1747.

14. Cai Y, Alio Del Barrio JL, Wilkins MR, Ang M. Serial optical coherence tomography angiography for corneal vascularization. *Graefes Arch Clin Exp Ophthalmol*. 2017;255:135–139.

15. Liu YC, Devarajan K, Tan TE, Ang M, Mehta JS. Optical coherence tomography angiography for evaluation of reperfusion after pterygium surgery. *Am J Ophthalmol*. 2019;207:151–158.

16. Nanji A, Redd T, Chamberlain W, et al. Application of corneal optical coherence tomography angiography for assessment of vessel depth in corneal neovascularization. *Cornea*. 2020;39:598–604.

17. Liu Z, Karp CL, Galor A, Ai Bayyat GJ, Jiang H, Wang J. Role of optical coherence tomography angiography in the characterization of vascular network patterns of ocular surface squamous neoplasia. *Ocul Surf*. 2020;18:926–935.

18. Binotti WW, Koseoglu ND, Nose RM, Kenyon KR, Hamrah P. Novel parameters to assess the severity of corneal neovascularization using anterior segment optical coherence tomography angiography. *Am J Ophthalmol*. 2021;222:206–217.

19. Binotti WW, Mills H, Nose RM, Wu HK, Duker JS, Hamrah P. Anterior segment optical coherence tomography angiography in the assessment of ocular surface lesions. *Ocul Surf*. 2021;22:86–93.

20. Binotti WW, Nose RM, Koseoglu ND, Dieckmann GM, Kenyon K, Hamrah P. The utility of anterior segment optical coherence tomography angiography for the assessment of limbal stem cell deficiency. *Ocul Surf*. 2021;19:94–103.

21. Cheung N, Donaghue KC, Liew G, et al. Quantitative assessment of early diabetic retinopathy using fractal analysis. *Diabetes Care*. 2009;32:106–110.

22. Schreiber S, Bueche CZ, Garz C, Braun H. Blood brain barrier breakdown as the starting point of

cerebral small vessel disease? New insights from a rat model. *Exp Transl Stroke Med.* 2013;5(1):4.

23. Jiang H, Zhong J, DeBuc DC, et al. Functional slit lamp biomicroscopy for imaging bulbar conjunctival microvasculature in contact lens wearers. *Microvasc Res.* 2014;92:62–71.

24. Chen W, Batawi HI, Alava JR, et al. Bulbar conjunctival microvascular responses in dry eye. *Ocul Surf.* 2017;15:193–201.

25. Haralick RM, Sternberg SR, Zhuang X. Image analysis using mathematical morphology. *IEEE Trans Pattern Anal Mach Intell.* 1987;9:532–550.

26. Otsu N. A threshold selection method from gray-level histograms. *IEEE Trans Sys Man Cyber.* 1979;9:62–66.

27. Phansalkar N, More S., Sabale A., Joshi M. Adaptive local thresholding for detection of nuclei in diversity stained cytology images. *International Conference on Communications and Signal Processing.* 2011:218–220.

28. Hassan G, El-Bendary N, Hassanien AE, Fahmy A, Shoeb AM, Snasel V. Retinal blood vessel segmentation approach based on mathematical morphology. *Procedia Comput Sci.* 2015;65:612–622.

29. Cole ED, Moult EM, Dang S, et al. The definition, rationale, and effects of thresholding in OCT angiography. *Ophthalmol Retina.* 2017;1:435–447.

30. Mehta N, Liu K, Alibhai AY, et al. Impact of binarization thresholding and brightness/contrast adjustment methodology on optical coherence tomography angiography image quantification. *Am J Ophthalmol.* 2019;205:54–65.

31. William B, Nose RM, Dieckmann G, Seyed-Razavi Y, Hamrah P. Comparative assessment of semi-automated conjunctival vessel quantification methodologies for spectral domain anterior segment optical coherence tomography angiography. *Invest Ophthalmol Vis Sci.* 2018;59:1670.

32. 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;2012:509783.

33. Devarajan K, Di Lee W, Ong HS, et al. Vessel density and En-face segmentation of optical coherence tomography angiography to analyse corneal vascularisation in an animal model. *Eye Vis.* 2019;6:2.

34. Devarajan K, Sim R, Chua J, et al. Optical coherence tomography angiography for the assessment of choroidal vasculature in high myopia. *Br J Ophthalmol.* 2020;104:917–923.

35. Longo A, Morscher S, Najafababdi JM, Justel D, Zakian C, Ntziachristos V. Assessment of hessian-based Frangi vesselness filter in optoacoustic imaging. *Photoacoustics.* 2020;20:100200.

36. Shah SAA, Tang TB, Faye I, Laude A. Blood vessel segmentation in color fundus images based on regional and Hessian features. *Graefes Arch Clin Exp Ophthalmol.* 2017;255:1525–1533.

37. Haddad RA, Akansu A.N. A class of fast gaussian binomial filters for speech and image processing. *IEEE Trans Acoustics Speech Signal Processing.* 1991;39:723–727.

38. Anijeet DR, Zheng Y, Tey A, Hodson M, Sueke H, Kaye SB. Imaging and evaluation of corneal vascularization using fluorescein and indocyanine green angiography. *Invest Ophthalmol Vis Sci.* 2012;53:650–658.

39. 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.

40. Portney L, Watkins M. *Foundations of Clinical Research: Applications to Practice.* Hoboken, NJ: Prentice Hall; 2000:892.

41. Ang M, Cai Y, MacPhee B, et al. Optical coherence tomography angiography and indocyanine green angiography for corneal vascularisation. *Br J Ophthalmol.* 2016;100:1557–1563.

42. Patel CN, Antony AK, Kommula H, Shah S, Singh V, Basu S. Optical coherence tomography angiography of perilimbal vasculature: validation of a standardised imaging algorithm. *Br J Ophthalmol.* 2020;104:404–409.

43. Akagi T, Uji A, Huang AS, et al. Conjunctival and intrascleral vasculatures assessed using anterior segment optical coherence tomography angiography in normal eyes. *Am J Ophthalmol.* 2018;196:1–9.

44. BahadarKhan K, Khaliq AA, Shahid M. A morphological hessian based approach for retinal blood vessels segmentation and denoising using region based Otsu thresholding. *PLoS One.* 2016;11(7):e0158996.

45. Kloppel S, Abdulkadir A, Hadjidemetriou S, et al. A comparison of different automated methods for the detection of white matter lesions in MRI data. *Neuroimage.* 2011;57:416–422.

46. Stanzel TP, Devarajan K, Lwin NC, et al. Comparison of optical coherence tomography angiography to indocyanine green angiography and slit lamp photography for corneal vascularization in an animal model. *Sci Rep.* 2018;8(1):11493.

47. Binotti WW, Koseoglu ND, Nose RM, Kenyon KR, Hamrah P. Novel parameters to assess

the severity of corneal neovascularization using anterior segment optical coherence tomography angiography. *Am J Ophthalmol*. 2021;222:206–217.

48. Cai S, Zhao F, Du C. Repeatability of ocular surface vessel density measurements with optical coherence tomography angiography. *BMC Ophthalmol*. 2019;19:248.

49. Zhao F, Cai S, Huang Z, Ding P, Du C. Optical coherence tomography angiography in pinguecula and pterygium. *Cornea*. 2020;39:99–103.

50. Aicher NT, Nagahori K, Inoue M, Itoh Y, Hirakata A. Vascular density of anterior segment of eye determined by optical coherence tomography angiography and slit-lamp photography. *Ophthalmic Res*. 2020;63:572–579.

51. Brunner M, Romano V, Steger B, et al. Imaging of corneal neovascularization: optical coherence tomography angiography and fluorescence angiography. *Invest Ophthalmol Vis Sci*. 2018;59:1263–1269.