Quantification and Repeatability of Vessel Density and Flux as Assessed by Optical Coherence Tomography Angiography

Konstantinos Pappelis1,2, and Nomdo M. Jansonius1,2

1 Department of Ophthalmology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
2 Graduate School of Medical Sciences (Research School of Behavioural and Cognitive Neurosciences), University of Groningen, Groningen, The Netherlands

Correspondence: Konstantinos Pappelis, Department of Ophthalmology, University Medical Center Groningen, PO Box 30.001, 9700 RB Groningen, Netherlands. e-mail: k.pappelis@rug.nl
Nomdo M. Jansonius. e-mail: n.m.jansonius@umcg.nl

Keywords: OCT-A; repeatability; test-retest; perfusion; vessel density

Citation: Pappelis K, Jansonius NM. Quantification and repeatability of vessel density and flux as assessed by optical coherence tomography angiography. Trans Vis Sci Tech. 2019;8(3):3, https://doi.org/10.1167/tvst.8.3.3

Copyright 2019 The Authors

Purpose: To determine the intrasession repeatability (test-retest variability) of parafoveal and peripapillary perfused capillary density (PCD) and normalized flux index (NFI) as assessed with Canon OCT-HS100 angiography.

Methods: Pairs of optical coherence tomography angiography (OCT-A) images were obtained from the parafoveal and peripapillary region of 30 eyes of 30 healthy subjects. PCD and NFI were calculated using generic image-processing software. Macular ganglion-cell complex thickness (GCC) and peripapillary retinal nerve fiber layer thickness (RNFLT) were also recorded. Bland-Altman analysis was performed and the coefficient of repeatability (CoR) and intraclass correlation coefficient (ICC) were calculated. Correlations of parafoveal PCD/NFI with GCC and of peripapillary PCD/NFI with RNFLT were also computed. Correlations of parafoveal PCD/NFI with GCC and of peripapillary PCD/NFI with RNFLT were also computed.

Results: Mean (standard deviation) parafoveal and peripapillary PCD were 40.0% (1.8%) and 44.5% (1.3%), respectively. Corresponding values for NFI were 151.2 (6.8) and 164.2 (3.9). For PCD, ICC was 0.76 for parafoveal and 0.79 for peripapillary measurements; corresponding CoRs were 2.7% and 1.8%. Corresponding values for NFI were 0.62 and 0.67 for ICC and 13.3 and 7.0 for CoR. Average measures ICC was 0.87/0.88 and 0.76/0.80 for the parafoveal/peripapillary PCD and NFI, respectively. PCD and NFI were weakly correlated with GCC (r = 0.39, P = 0.035; r = 0.33, P = 0.077) and moderately correlated with RNFLT (r = 0.43, P = 0.017; r = 0.55, P = 0.002).

Conclusions: Repeatability of a commercially available OCT-A with generic image-processing software was good (NFI) to excellent (PCD). Our results indicate that changes surpassing the variability in healthy subjects should be easily detectable in a clinical setting.

Translational Relevance: Repeatability estimates provide information regarding the relevance of changes in retinal perfusion.

Introduction

Optical coherence tomography angiography (OCT-A) is a novel imaging modality enabling the fast, noninvasive, and depth-resolved visualization of the retinal and optic nerve head (ONH) microcirculation, thus potentially assisting in the diagnosis and follow-up of numerous ocular pathologies.1,2 However, in order to interpret findings in an objective rather than qualitative manner, standardized quantitative metrics must be established, and their repeatability must enable the detection of clinically relevant changes.

The main principle behind OCT-A is the detection of erythrocyte movement by comparing the static and dynamic signal properties between consecutive B-scans. Different image acquisition algorithms have been implemented in current commercial OCT-A systems. Moreover, image post-processing and depiction as well as the quantitative measurements offered by some OCT manufacturers introduce discrepancies
in the technique’s reproducibility. Thus, OCT-A currently struggles with the proprietary nature of the metrics and therefore requires each instrument to be evaluated separately.

Perfused capillary density (PCD) is one of the most frequently used metrics within the context of OCT-A: a reduction in PCD may be present in various ocular disorders, but particular interest has been taken in its potential use for evaluating glaucoma. Even though the relationship between OCT-A signal and blood flow is still unclear, PCD can provide information not only on the structure but also on the function of the microvasculature, since OCT-A allows visualization of only the flowing capillaries. Another metric, namely Normalized Flux Index (NFI), has been regarded as even more informative by certain investigators.

PCD is available as a parameter on several commercial OCT-A devices and its repeatability has been reported as well. However, other systems—such as the angiographic module of Canon OCT-HS100—currently lack this quantitative analysis and therefore the (intra-device) repeatability has not been evaluated yet, at least not outside the foveal avascular zone. NFI on the contrary is not available at the moment in any manufacturer’s quantification software. Related to that, information regarding its repeatability is scarce.

The aim of this study was to determine the intrasession repeatability (test-retest variability) of parafoveal and peripapillary PCD and NFI as assessed with Canon OCT-HS100. For this purpose, we obtained OCT-A images from a well-defined healthy sample and analyzed the data using generic image processing software in order to compensate for the system’s absence of quantitative measurements and allow for harmonization of data across different OCT-A brands.

**Methods**

**Study Population**

This is a prospective, cross-sectional study. All subjects between 50 and 65 years of age who responded to our advertisement received an information letter. We imaged one random eye of the first 34 participants who satisfied the following inclusion criteria after initial screening: best-corrected visual acuity ≥ 0.8; spherical equivalent refraction between −3 and +3 D and astigmatism not exceeding 2 D; intraocular pressure (IOP) ≤ 21 mm Hg as assessed by a noncontact tonometer (Tonoref II, Nidek, Aichi, Japan); no reproducibly abnormal visual field test locations in Frequency Doubling Technology C20-1 screening mode (Carl Zeiss, Jena, Germany); no ophthalmic, vascular, cardiac, or blood disease (except for hypertension) as assessed by fundus imaging (TRC-NW400, Topcon Corporation, Tokyo, Japan) and a medical history questionnaire; no family history of glaucoma.

The ethics board of the University Medical Center Groningen (UMCG) approved the study protocol (#NL61508.042.17). All participants provided written informed consent. The study followed the tenets of the Declaration of Helsinki.

**Data Collection**

Prior to the imaging session, the pupil of the chosen eye was dilated with tropicamide 0.5%. In addition, blood pressure (BP) was recorded twice in sitting position with an automatic BP monitor (Omron M6 Comfort, Omron Healthcare, Kyoto, Japan) from the brachial artery.

Subsequently, three 6 x 6 mm scans centered at the fovea and three 6 x 6 mm scans centered at the ONH were obtained in succession with the angiographic module of the Canon OCT-HS100 (Angio eXpert, OCTA version 2.0, Tokyo, Japan). In order to optimize image quality, while keeping at the same time scanning duration to tolerable levels (less than 10 seconds per scan), the resolution was set to “medium” (928 x 928 pixels) after repeats within each scan were set to “2.” Between the scans participants were asked to remove and reposition their head on the chinrest. The interval between the scans was typically less than 1 minute, being the time needed for removing and repositioning the participant’s chin on the chinrest and refocusing. All scans were obtained in the late afternoon (5:00 PM–6:00 PM).

Out of the six angiographic scans obtained, the first two parafoveal and the first two peripapillary scans with sufficient quality (≥7 as assessed by the device itself) and free of motion or blinking artifacts were included in the repeatability analysis. Participants whose images did not satisfy the aforementioned criteria were excluded from the study. Of the included scans, the median image quality was 8 (range, 7–9) for the parafoveal scans and 7 (7–8) for the peripapillary scans. To elucidate whether the angiographic imaging signal and its subsequent analysis are truly informative, we also recorded as a secondary outcome the parafoveal ganglion-cell complex thickness (GCC) and the peripapillary retinal...
nerve fiber layer thickness (RNFLT), which have been previously shown to positively correlate with PCD even in healthy subjects.14

Image Segmentation and Analysis

The angiographic images containing information from the maximum intensity projection signal of the superficial capillary plexus (top: inner limiting membrane; bottom: ganglion cell/inner plexiform border; offset: +50 μm) as determined by the manufacturer’s segmentation were stored in an uncompressed format (bitmap). Even though the device can visualize the deep capillary plexus with a projection-resolved algorithm, we limited the analysis to the inner retinal layers to ensure that the signal is not a result of projection artifacts.15

Pairs of images from the same subject corresponding to the same region (parafoveal or peripapillary) were registered by means of a rigid body transformation matrix (rotation and translation) so that their structural features coincide (ImageJ; public domain software, National Institutes of Health, Bethesda, MD16). Local Otsu’s thresholding algorithm was then applied in 14 × 14 pixel blocks to binarize each image in signal of flow and nonflow. Additionally, the big blood vessels were masked out of the images centered at the ONH through a combination of Hessian-based Frangi vesselness filtering algorithm17 (made available to the public by Dirk-Jan Kroon, 2009) and global thresholding. Our analysis was confined to a well-defined ring around the ONH with inner and outer radii of 1.03 and 1.84 mm, respectively18 and a 3-mm-diameter disc around the fovea (Fig. 1).19 An experienced grader (KP) determined the center of the ONH and of the foveal avascular zone.

We calculated PCD as the percentage of pixels occupied by capillaries inside the total measurement area. NFI was calculated as the average signal

Figure 1. Regions of interest defined. (A) Peripapillary OCT-A image; (B) Peripapillary region of interest after local thresholding; (C) Peripapillary large vessel mask; (D) Parafoveal OCT-A image; (E) Parafoveal region of interest after local thresholding (no large vessel mask applied parafoveally).
strength (grayscale intensity) of the pixels associated with perfused capillaries and is a unitless number between 0 and 255 (being the 8-bit intensity coding in the raw image, before binarization). For the peripapillary scan, the area occupied by large vessels was not included in these calculations.

Image processing was performed in MATLAB R2014a (MathWorks, Natick, MA). The scripts are available on request.

Data Analysis

Means and standard deviations (SDs) were used to describe normally distributed variables. Variables with a skewed distribution were described by median and interquartile range (IQR). We generated scatterplots of the second measurement as a function of the first as well as the corresponding Bland-Altman plots for the parafoveal and peripapillary PCD and NFI.

We computed the mean difference in the two repeated scans for the parafoveal and peripapillary PCD and NFI. For these variables we also calculated (1) the coefficient of repeatability (CoR) as twice the SD of the difference in two repeated scans and (2) the two-way mixed intraclass correlation coefficient (ICC), both for single and average measures (ICC[2,1]/ICC[2,2]). For the interpretation of ICC values, we used the guidelines of Cicchetti et al.

Additionally, we computed Pearson’s correlation coefficient for parafoveal PCD/NFI versus GCC as well as peripapillary PCD/NFI versus RNFLT. All analysis was performed using R (version 3.4.2; R Foundation for Statistical Computing, Vienna, Austria) and WINPEPI (PAIRSetc version 3.59; Abramson JH, 2004). A P-value of 0.05 or less was considered statistically significant.

Results

Of the 34 participants satisfying the screening criteria, four were excluded due to significant image artifacts. Therefore, a total of 30 participants were included in the analysis. Table 1 shows the demographics and characteristics of the study population.

Table 2 summarizes mean values for PCD and NFI in both regions of interest, mean differences between the two consecutive scans, and the repeatability estimates. Based on the ICC, repeatability was excellent for PCD and good for NFI. A one-sample t-test ensured that the difference between the two consecutive scans was not significantly different from 0 (all P-values ≥ 0.4). Moreover, the absolute interscan difference for the various outcomes did not depend on age, BP, or body mass index, as assessed by the significance of Pearson’s correlation coefficient corrected for multiple testing with the Bonferroni method (all P-values > 0.2). Figures 2 and 3 present the scatterplots of the second measurement metrics plotted as a function of the first measurement metrics and the corresponding Bland-Altman plots, for the PCD and NFI, respectively.

Regarding our secondary outcome, only PCD but not NFI was significantly correlated with GCC in the parafoveal region ($r = 0.39$ [P = 0.035] and $r = 0.33$ [P = 0.077], respectively). In the peripapillary region, both PCD and NFI were significantly correlated with RNFLT ($r = 0.43$ [P = 0.017] and $r = 0.55$ [P = 0.002], respectively).

Table 1. Characteristics of Study Population

<table>
<thead>
<tr>
<th></th>
<th>Mean (IQR)</th>
<th>Gender, % female</th>
<th>IOP, mean (SD), mm Hg</th>
<th>SBP, mean (SD), mm Hg</th>
<th>DBP, mean (SD), mm Hg</th>
<th>GCC, mean (SD), µm</th>
<th>RNFLT, mean (SD), µm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, median (IQR), y</td>
<td>58 (53–61)</td>
<td>53</td>
<td>13 (3)</td>
<td>130 (14)</td>
<td>85 (8)</td>
<td>92 (5)</td>
<td>100 (10)</td>
</tr>
<tr>
<td>Gender</td>
<td>53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOP</td>
<td>13 (3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SBP</td>
<td>130 (14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DBP</td>
<td>85 (8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCC</td>
<td>92 (5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNFLT</td>
<td>100 (10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SBP, systolic blood pressure; DBP, diastolic blood pressure.

Table 2. Repeatability Estimates for Parafoveal and Peripapillary PCD and NFI

<table>
<thead>
<tr>
<th></th>
<th>Mean (SD)</th>
<th>Mean (SD) Difference</th>
<th>CoR (95% CI)</th>
<th>ICC[2,1] (95% CI)</th>
<th>ICC[2,2] (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parafoveal PCD (%)</td>
<td>40.0 (1.8)</td>
<td>0.1 (1.4) [P = 0.8]</td>
<td>2.7 (1.8–3.6)</td>
<td>0.76* (0.56–0.88)</td>
<td>0.87* (0.72–0.94)</td>
</tr>
<tr>
<td>Peripapillary PCD (%)</td>
<td>44.5 (1.3)</td>
<td>0.1 (0.9) [P = 0.4]</td>
<td>1.8 (1.2–2.4)</td>
<td>0.79* (0.60–0.89)</td>
<td>0.88* (0.75–0.94)</td>
</tr>
<tr>
<td>Parafoveal NFI</td>
<td>151.2 (6.8)</td>
<td>1.0 (6.7) [P = 0.4]</td>
<td>13.3 (9.1–17.5)</td>
<td>0.62* (0.34–0.80)</td>
<td>0.76* (0.51–0.89)</td>
</tr>
<tr>
<td>Peripapillary NFI</td>
<td>164.2 (3.9)</td>
<td>0.4 (3.5) [P = 0.6]</td>
<td>7.0 (4.8–9.2)</td>
<td>0.67* (0.41–0.83)</td>
<td>0.80* (0.58–0.91)</td>
</tr>
</tbody>
</table>

CoR, coefficient of repeatability; ICC, intraclass correlation coefficient.

CI, confidence interval.

* Significant at P = 0.001.
Discussion

Parafoveal and peripapillary PCD measurements obtained by the Canon OCT-HS100 within the same session and from the same subject are in excellent agreement when considering the ICC values; upon retest their absolute values may differ up to 2.7% and 1.7%, respectively. NFI measurements are in good agreement; upon retest they may differ up to 13.1 and 6.9, respectively. Reliability from average measures is higher than reliability from single measures among all metrics in both regions. The angiographic metrics are weakly to moderately correlated with the corresponding retinal layer thicknesses.

A number of other studies have evaluated the intrasession repeatability of OCT-A metric PCD, albeit with different devices. Regarding the parafoveal region, Alnawaiseh et al.\textsuperscript{11} found the CoR and ICC (2,1) for the PCD to be 3.4% and 0.72, respectively, to be compared with 2.7% (95% confidence interval [CI]: 1.8%–3.6%) and 0.76 (0.56–0.88) in our study. Fang et al.\textsuperscript{23} reported a CoR of 3.2% and an ICC of 0.86, while the same repeatability variables reported by Coscas et al.\textsuperscript{24} were 3.3% and 0.78 and by Al-Sheikh et al.\textsuperscript{9} 3.4% and 0.90. The latter three studies did not clarify which ICC was presented. An ICC of 0.89 without further specification was also reported by Lei et al.\textsuperscript{10} A larger CoR of 4.9% was reported by Chen et al.\textsuperscript{25} because they included multiplication by the square root of 2 in the calculation; the corrected value is 3.4%. Venugopal et al.\textsuperscript{26} presented a CoR of 4.4% together with an ICC of 0.87 for the parafoveal region, as well as the corresponding values of 4.1% and 0.86 for the peripapillary region. Hence, the CoR for PCD that we report here as assessed by the angiographic module of Canon OCT HS-100 with customized software is at least as good as the ones
reported in most of the aforementioned studies for different devices. Venugopal et al.\textsuperscript{26} using a different device in their study, found the measurements to be significantly less repeatable than ours both in the parafoveal and the peripapillary region. This could be attributed to improved results of our customized software, to different inclusion criteria, or to differences between devices in terms of scanning and segmentation.

Chen et al.\textsuperscript{8} evaluated the repeatability of NFI but with a different statistical approach. They found a coefficient of variation (CV) of 3.3\% and 4.2\% for the parafoveal and peripapillary region, respectively. For the sake of comparison, we also calculated the CV (SD of differences divided by the mean), which was 2.3\% (CI: 1.7\%–2.9\%) for the parafoveal and 1.2\% (CI: 0.9\%–1.5\%) for the peripapillary region. This might suggest improved repeatability of our methods; however, using exclusively the CV might spuriously suggest that repeatability is worse (higher CV) in ocular diseases where the mean angiographic metrics decrease (e.g., in glaucoma) or in devices with lower average signal intensity. Indeed, a recent study found some differences between glaucoma and healthy in the CV of the inferotemporal region of the peripapillary scan, but not in the CoR.\textsuperscript{26}

Variation among consecutive scans could in general be a consequence of signal strength (even within the high quality scans), floaters, or measurement noise. Importantly, since OCT-A is only able to image capillaries that are perfused and visible given the resolution of the system, it is possible that a portion of the variation between consecutive scans could be attributed to physiological reasons such as small changes in perfusion pressure due to, for example, cardiac cycle related variability in IOP.\textsuperscript{27}

To our knowledge, this is the first study that addressed the PCD and NFI intrasession repeatability.
with the Canon OCT-HS100, and the first report of CoR and ICC for OCT-A metric NFI. Importantly, the Canon OCT-HS100 is currently the only device using a full-spectrum amplitude decorrelation algorithm; therefore, a separate evaluation was also deemed necessary. A strength of this study is the fact that the calculation of the quantification parameters and reasoning are described in detail. This generic approach avoids obscurities involved in metrics belonging to proprietary algorithms and thus allows for harmonization. For example, it is unclear if other algorithms include the larger vessels originating from the ONH in their calculations. A limitation of this study is the restriction to the superficial capillary plexus. Additionally, other metrics such as fractal dimension and foveal avascular zone were not considered in this analysis. Lastly, it is possible that the small sample size of the study affects the weak correlations of PCD and NFI with the GCC. A larger sample size could, for example, result in measured differences regressing toward the mean. However, our results are in agreement with Yu et al. who reported stronger correlations in the peripapillary than the parafoveal sector.

The results suggest that it is possible to quantify the retinal microvasculature through OCT-A with a satisfactory degree of accuracy. The additional information provided by OCT-A metrics can be helpful in differentiating between healthy and diseased eyes within the clinical setting, should the effect size be sufficiently large. However, this study shows that one single OCT-A image, which is most frequently obtained within the clinical setting, is not enough to guarantee a reliable absolute value estimation. Consequently, the use of these measurements for the evaluation and follow-up on an individual basis is not recommended. Instead, averaging of consecutive images or measurements might be a more informative approach, and this is also suggested in our study when comparing average measures versus single measures ICC values.

In conclusion, by applying a generic quantification algorithm to the images obtained with a commercially available OCT-A, we were able to quantify perfusion and estimate its intrasession repeatability. Small changes in perfusion fall within the test-retest variability; changes surpassing the variability in healthy subjects should be easily detectable in a clinical setting. This is important, since it provides insight on how the output of the specific device can be interpreted and handled in the clinic. Metrics with improved test-retest variability and diagnostic accuracy together with quantified blood velocity could potentially not only serve as additional clinical markers but also help unravel underlying pathophysiological mechanisms.

Acknowledgements

The authors thank K. Westra and W. Nijboer for their contribution in data collection, as well as R. Müskens and T. Heikka for their consulting role on technical aspects.

Supported by European Union’s Horizon 2020 Innovative Training Networks Program, under the Marie Sklodowska – Curie grant, Project ID 675033. The funding organization had no role in the design, conduct, analysis, or publication of this research.


Disclosure: K. Pappelis, None; N.M. Jansonius, None

References

16. Thévenaz P, Ruttimann UE, Unser M.
19. Ho J, Dans K, You Q, Nudleman ED, Freeman WR.
20. Bland JM, Altman DG.
21. Shrout PE, Fleiss JL.
22. Cicchetti DV.
23. Fang D, Tang FY, Huang H, Cheung CY, Chen H.


