Title
Local Variability of Macular Thickness Measurements With SD-OCT and Influencing Factors.

Journal
Translational vision science & technology, 5(4)

ISSN
2164-2591

Authors
Miraftabi, Arezoo
Amini, Navid
Gornbein, Jeff
et al.

Publication Date
2016-07-01

DOI
10.1167/tvst.5.4.5

Peer reviewed

tvst

DOI: 10.1167/tvst.5.4.5

Article

Local Variability of Macular Thickness Measurements With SD-OCT and Influencing Factors

Arezoo Miraftabi[1,2], Navid Amini[1], Jeff Gornbein[1], Sharon Henry[1], Pablo Romero[1], Anne L. Coleman[1], Joseph Caprioli[1], and Kouros Nouri-Mahdavi[1]

[1] Glaucoma Division, Stein Eye Institute, David Geffen School of Medicine, University of California-Los Angeles, Los Angeles, CA, USA
[2] Eye Research Center, Rasoul Akram Hospital, Iran University of Medical Sciences, Tehran, Iran

Correspondence: Kouros Nouri-Mahdavi, MD MSc, 100 Stein Plaza, Los Angeles, CA 90095, USA.
e-mail: nouri-mahdavi@jsei.ucla.edu

Received: 25 September 2015
Accepted: 9 May 2016
Published: 19 July 2016

Keywords: SD-OCT; macular imaging; glaucoma; variability

Citation: Miraftabi A, Amini N, Gornbein J, et al. Local variability of macular thickness measurements with SD-OCT and influencing factors. Trans Vis Sci Tech. 2016;5(4):5, doi:10.1167/tvst.5.4.5

Purpose: To compare the intrasession variability of spectral-domain optical coherence tomography (SD-OCT)-derived local macular thickness measures and explore influencing factors.

Methods: One hundred two glaucomatous eyes (102 patients) and 21 healthy eyes (21 subjects) with three good quality macular images during the same session were enrolled. Thickness measurements were calculated for 3° superpixels for the inner plexiform (IPL), ganglion cell (GCL), or retinal nerve fiber layers (mRNFL), GC/IPL, ganglion cell complex, and full macular thickness. Spatial distribution and magnitude of measurement errors (ME; differences between the 3 individual superpixel values and their mean) and association between MEs and thickness, age, axial length, and image quality were explored.

Results: MEs had a normal distribution with mostly random noise along with a small fraction of outliers (1.2%–6.6%; highest variability in mRNFL and on the nasal border) based on M-estimation. Boundaries of 95% prediction intervals for variability reached a maximum of 3 μm for all layers and diagnostic groups after exclusion of outliers. Correlation between proportion of outliers and thickness measures varied among various parameters. Age, axial length, or image quality did not influence MEs ($P > 0.05$ for both groups).

Conclusions: Local variability of macular SD-OCT measurements is low and uniform across the macula. The relationship between superpixel thickness and outlier proportion varied as a function of the parameter of interest.

Translational Relevance: Given the low and uniform variability within and across eyes, definition of an individualized 'variability space' seems unnecessary. The variability measurements from this study could be used for designing algorithms for detection of glaucoma progression.

Introduction

Glaucoma is characterized by progressive loss of retinal ganglion cells (RGCs). With the advent of spectral-domain optical coherence tomography (SD-OCT), there has been increased interest in measuring RGC loss or thinning of the inner retina in the macula.[1–5] Approximately 50% of the RGCs are located within 4.5 mm (16°) of the foveal center.[6] The central macula is the only part of the retina where the RGC layer is more than one cell thick; the peak RGC density occurs approximately 750 to 1100 μm from the center of the fovea.[6] Regional inner retinal thickness measurements such as ganglion cell/inner plexiform layer (GC/IPL) thickness perform as well as regional retinal nerve fiber layer (RNFL) thickness measures for detection of early glaucoma.[3,4,7] Hood et al.[8] demonstrated that GC/IPL thinning could be demonstrated in a group of patients with preperimetric glaucoma, although other studies have not found macular thickness to perform as well as RNFL thickness for detection of glaucoma in such patients.[5,9]

The role of macular SD-OCT imaging for detection of glaucoma progression is yet to be fully elucidated. Early results have indicated its potential use.[10–12] Reproducibility of any structural outcome measure is one of the main determinants of how

useful it would be for detection of change. Factors that affect the variability of macular SD-OCT images have not been fully explored; existing studies have shown very good to excellent reproducibility for various macular thickness measurements with SD-OCT.[2,13–19] However, it is not yet clear which macular parameter would perform best for detection of glaucoma progression. One caveat of the aforementioned studies is that they have mostly reported thickness variability either globally or in larger predetermined areas of the macula such as hemiretinal regions or sectors. Estimation of measurement variability in smaller regions of the macula would potentially allow us to define patterns of change and use this knowledge to better detect glaucoma deterioration.

Thickness measurement of individual macular retinal layers has not been widely available due to the suboptimal quality of the SD-OCT images or the segmentation algorithms. New software implemented on the Spectralis SD-OCT (Heidelberg Engineering, Heidelberg, Germany) called Glaucoma Module Premium Edition or GMPE is now able to segment individual macular retinal layers, including the macular RNFL (mRNFL), ganglion cell layer (GCL), and IPL. Thickness data are provided in an 8×8 grid with each square cell or 'superpixel' measuring 3° in width.

The aim of the current study was to measure and compare the local intrasession variability for individual inner retinal layers or combination of layers in the macula at the superpixel level in a group of glaucoma and healthy subjects. The main layers of interest include the IPL, GCL, mRNFL, GC/IPL, ganglion cell complex (GCC; combination of GC/IPL and mRNFL), and full macular thickness. A secondary aim of this study is to explore factors that influence such measurement variability.

Methods

This study was approved by the institutional review board at the University of California Los Angeles (UCLA) before the original study started and all procedures followed the tenets of the Declaration of Helsinki. Patients were consented and enrolled from The Advanced Glaucoma Progression Study (AGPS), which is a longitudinal prospective study underway in the Glaucoma Division of Stein Eye Institute at UCLA. The goal of AGPS is to define the role of macular SD-OCT imaging to detect deterioration in patients with advanced glaucoma. One hundred two glaucoma patients (102 eyes) and 21 healthy subjects (21 eyes) with three good-quality macular SD-OCT images at baseline exam were enrolled. Only patients with a diagnosis of primary open-angle glaucoma, pseudoexfoliative, pigmentary, or primary chronic angle-closure glaucoma were eligible for recruitment. Study eyes underwent a thorough eye exam at baseline including best-corrected visual acuity, automated refraction, corneal pachymetry, slit-lamp exam, intraocular pressure measurement with applanation tonometry, gonioscopy, dilated fundus exam, biometry with IOLMaster, and achromatic visual field testing (SITA standard 24-2 and 10-2 fields with Humphrey Field Analyzer). Glaucoma was defined as the presence of glaucomatous optic nerve damage (i.e., vertical cup-to-disc ratio > 0.6, cup to disc asymmetry > 0.2, or presence of focal rim thinning or notching) along with an associated visual field defect on standard achromatic perimetry (SAP). A visual field defect was considered to be present on 24-2 fields if both of the following criteria were met: (1) Glaucoma Hemifield Test outside normal limits, and (2) four abnormal points with P less than 5% on the pattern deviation plot, both confirmed at least once.[20] Patients were also required to meet the following criteria: less than 3 diopters (D) of astigmatism, mean deviation of −6.0 dB or worse or evidence of central visual field involvement (i.e., at least two points within the central 10° with $P < 0.05$ on pattern deviation plot on 24-2 test), and no significant retinal or neurological disease. The heathy subjects had a normal eye exam, open angles, normal appearing optic discs, RNFL, and SAP visual fields.

Imaging Protocol

The Posterior Pole algorithm of the Spectralis SD-OCT was used to obtain a $30° \times 25°$ volume scan of the macula centered on the fovea. The algorithm performs 61 horizontal B-scans tilted parallel to the fovea-disc axis approximately 120 μm apart. The central $24° \times 24°$ of the measurement cube is segmented by the software and data are presented in an 8×8 array with each cell or superpixel $3° \times 3°$ in width. Each B-scan is repeated between 9 and 11 times to improve image quality. As per study protocol, recruited patients had three consecutive macular images taken during a single session by the same operator at the baseline visit. After taking each image, the patient was instructed to lean back and his/her head was repositioned to mimic as far as possible repeat imaging at separate sessions. Only images with

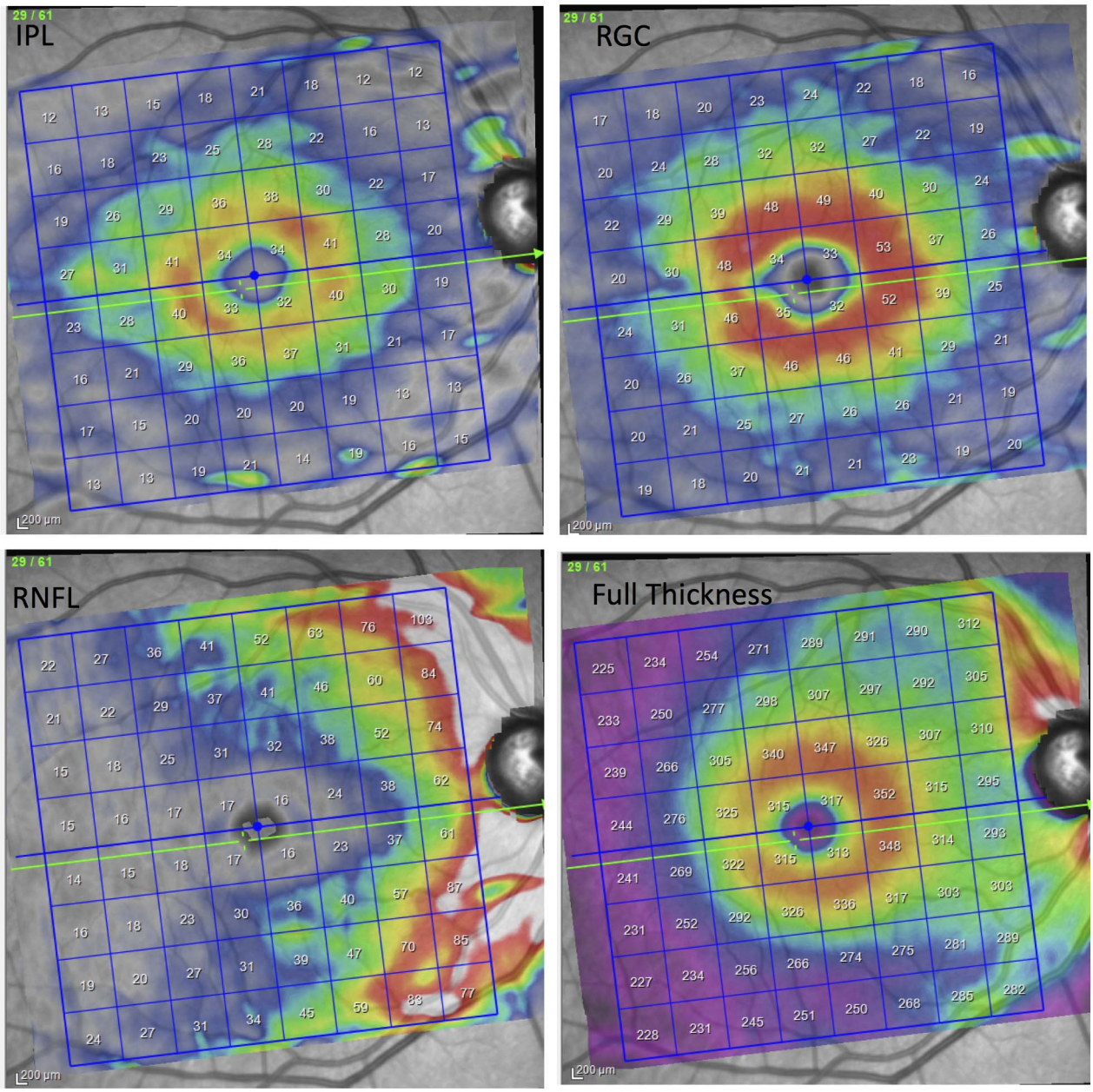

Figure 1. An example of a macular OCT image after segmentation demonstrating distribution of the IPL, GCL, mRNFL, and full retinal thickness in a healthy eye.

a quality factor of 15 or higher were included. One of the investigators (AM) reviewed all the B-scans and measurement grid positions to make sure the images were centered on the fovea and to check for image artifacts. If more than two B-scans in any individual volume scan were of inadequate quality, that eye was excluded from analyses. A poor quality B-scan image was defined as the presence of more than 10% missing data or inadequate segmentation, or any artifacts such as mirror artifacts. The Spectralis GMPE software performs segmentation of individual retinal layers, and the data are exported as extensible markup language files. The segmented images from one eye of each patient was also reviewed by one of two observers (PR and SH) to verify the accuracy of the layer segmentations and manually corrected as necessary. The thickness in combinations of layers such as GC/IPL and GCC) can be calculated by simply adding the thickness of individual layers (Fig. 1). The macular layers (or combination of layers) of

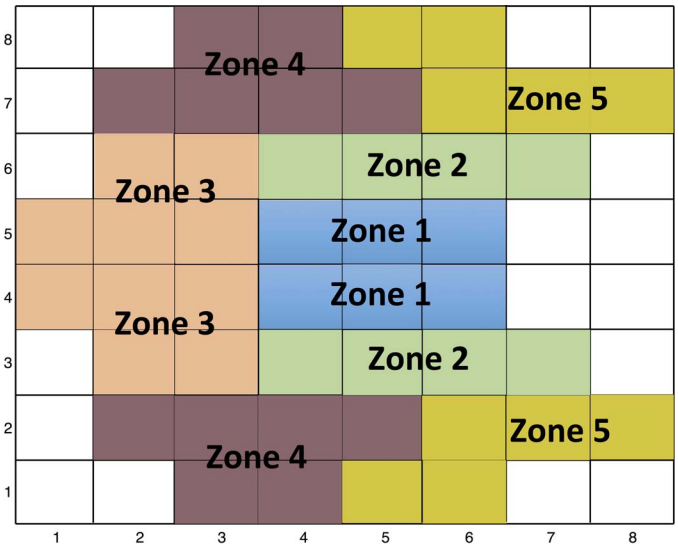

Figure 2. Representation of 'Glaucoma Hemifield' sectors as described originally by Um and colleagues.[22] The individual superpixels are numbered starting temporally and inferiorly.

interest in this study were: IPL, GCL, and mRNFL layers, GC/IPL, GCC, and full macular thickness. All the data are presented in right eye format (Fig. 1).

Statistical Methods

For each eye, three thickness measurements were available in each of the $8 \times 8 = 64$ superpixels. Mean, SD, and measurement errors (MEs), defined as the difference between each of the three values and the mean of three measurements at each superpixel, were calculated for all layers for a total of 64×3 measurements per eye per parameter. ME histograms and normal quantile plots were reviewed.

Robust M estimation with a bisquare weight function was used to estimate the ME mean and SD assuming a normal distribution after pooling data for each group (i.e., glaucoma versus healthy) and layer. We thereby detected outliers by identifying the outlier threshold (largest value that is not an outlier) beyond the normal distribution (defined as 0.5% and 99.5% percentiles). The M estimate SD is the SD of MEs with outliers omitted. For a given group and layer, the model assumes that the ME distribution is normal with a constant mean and SD for most of the errors but that there is a small percentage of outliers (or contaminants). By estimating the mean and SD with robust methods, the percentage (probability) of outliers is also estimated. Normal quantile plots of the MEs with outliers removed were subsequently examined to confirm that the ME values with outliers omitted had a normal distribution. To compare the magnitude of ME variability across layers and diagnostic groups, SDs were compared with the Fligner modification of the Brown-Forsythe method, which is based on the corresponding median absolute deviations (MADs).[21] The latter takes into account the systematic or random effects of eye, and superpixel location. After controlling for these systematic and random effects, the residual replicate MEs are independent or uncorrelated. Linear-regression analysis was used to explore the potential factors affecting superpixel variability.

We used a Glaucoma Hemifield Test pattern as suggested by Um et al.[22] (Fig. 2) to define five sectors in the superior and inferior hemiretinas where the average thickness was calculated for all the parameters. Afterwards, the intraclass correlation coefficient (ICC) was calculated in all the sectors for all layers so as to be able to compare results to the current literature. Analyses were done with the R software (version 3.2.4, http://www.R-project.org) or Stata software version 12.0 (StataCorp LP, College Station, TX).

Results

The mean (±SD) age was 66.7 (±9.4) years in the glaucoma group and 61.1 (±8.8) years in the healthy subjects ($P = 0.014$). Visual field mean deviation was −7.3 (±6.3) dB and −0.3 (±0.9) dB in the glaucoma and healthy subjects, respectively (Table 1). No eye was excluded because of poor image quality and all images had a quality factor greater than 15. Four eyes

Table 1. Demographic and Clinical Characteristics of the Study Sample

Number (Patients/Eyes)	Glaucoma (102 Eyes)	Healthy (21 Eyes)	*P* Value
Age, y (mean ± SD)	66.7 (±9.4)	61.1 (±8.8)	0.014
Sex, female/male	42/60	13/8	0.858
Visual acuity, logMar (mean ± SD)	0.07 (±0.09)	0.02 (±0.04)	<0.001
Mean deviation, dB (mean ± SD)	−7.25 (±6.3)	−0.30 (±0.9)	<0.001
Axial length, mm (mean ± SD)	24.75 (±1.53)	23.67 (±0.97)	0.002

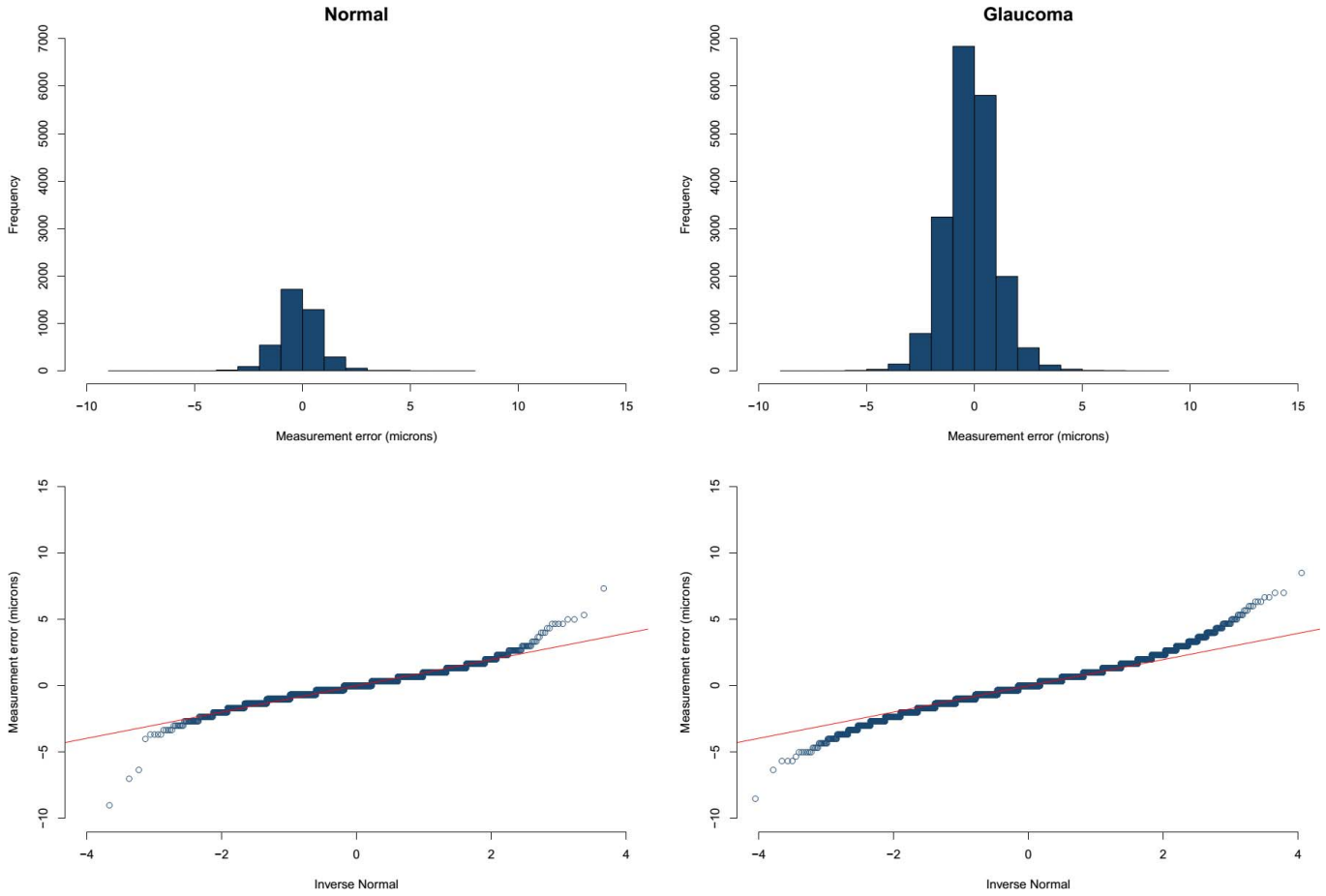

Figure 3. *Top*: histogram demonstrating the distribution of measurement errors for the IPL in the healthy (A), and glaucoma (B) groups. The distributions are quite normal except for a small number of outliers toward both tails of the distribution. *Bottom*: the normal quantile plots for the same layer and groups demonstrating the small tails of outliers in superpixel thickness data starting at approximately a z score of 3.

were excluded from the original sample because of suspected macular pathology. Examination of normal quartile plots indicated that the ME distributions mostly followed a Gaussian distribution except for a small number of outliers. The distribution of the IPL MEs in healthy and glaucoma groups is shown in Figure 3 as a representative example. Between 1.2% and 6.6% of the superpixel measurements were detected as outliers as identified by the M estimation approach in the two groups (Table 2). Figure 4 demonstrates the proportion of superpixels considered as outliers by our approach as a function of location within the macular grid. It can be seen that the mRNFL and full thickness parameters had the highest proportion of outliers among all layers in both healthy and glaucoma groups (6.1% and 5.1% vs. 3.3% and 6.6% in the healthy and glaucoma groups, respectively). Also, for most parameters, the outliers were concentrated at the nasal and superior borders

Table 2. Proportion of Outliers (i.e., Measurement Errors Considered Not to Represent Random Noise According to Macular Outcome Measure and Diagnosis [Glaucoma vs. Normal])

Layer	Group	Number of Superpixels	Proportion of Outliers
IPL	Glaucoma	19,519	1.5%
IPL	Healthy	4,032	1.2%
GCL	Glaucoma	19,521	1.8%
GCL	Healthy	4,032	1.5%
mRNFL	Glaucoma	19,525	5.1%
mRNFL	Healthy	4,032	6.1%
GCL+IPL	Glaucoma	19,584	1.7%
GCL+IPL	Healthy	4,032	3.4%
GCC	Glaucoma	19,584	2.3%
GCC	Healthy	4,032	1.8%
Full	Glaucoma	19,534	6.6%
Full	Healthy	4,032	3.3%

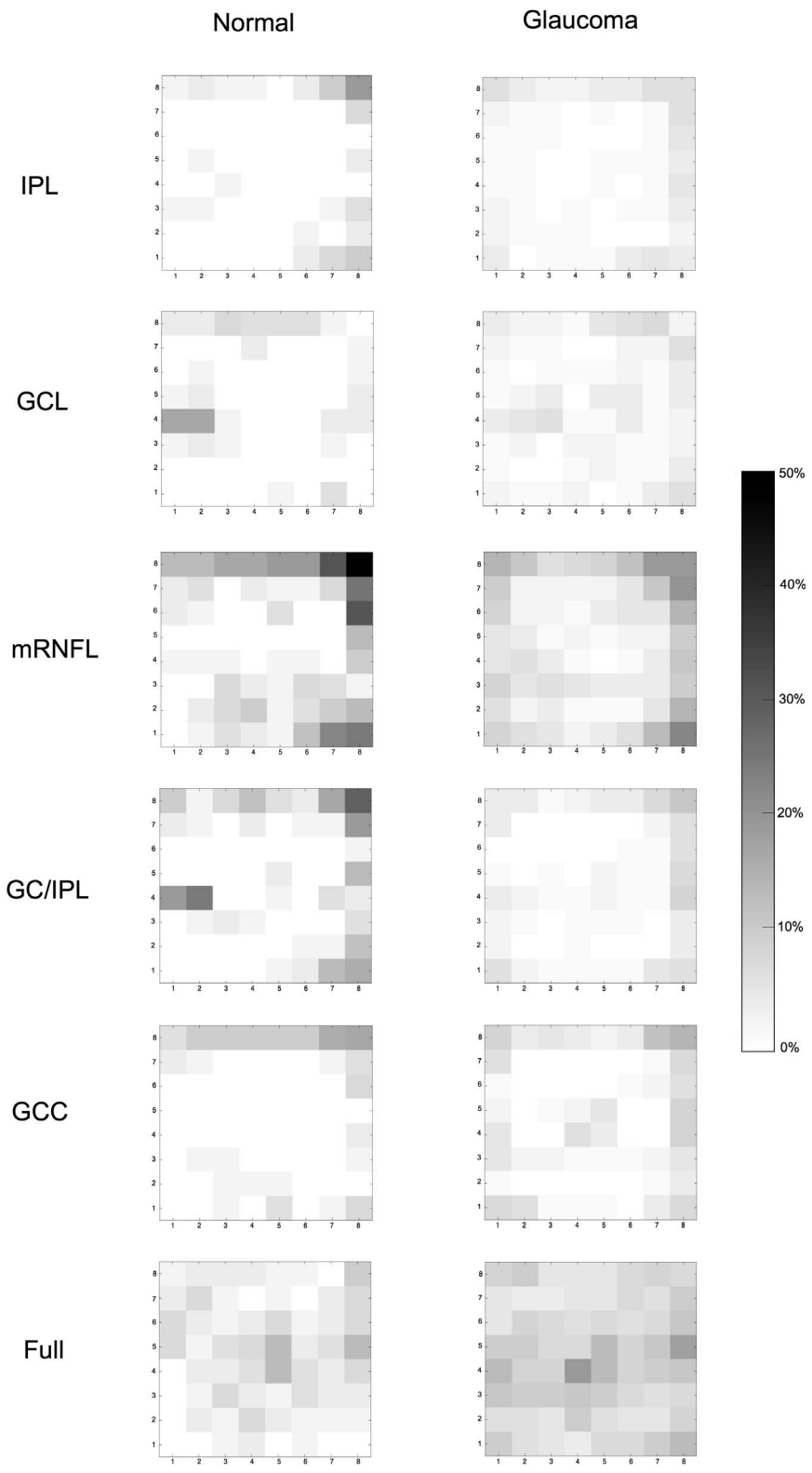

Figure 4. Gray scale images demonstrate the topographic distribution of outliers for various macular outcome measures. IPL: inner plexiform layer; GCL; ganglion cell layer; mRNFL: macular retinal nerve fiber layer; GC/IPL: ganglion cell/inner plexiform layer; GCC: ganglion cell complex; Full: full macular thickness.

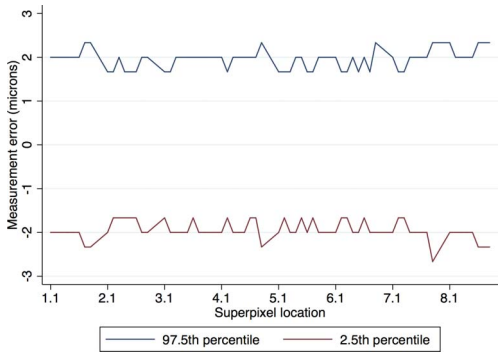

Figure 5. An example of the 95% prediction interval for variability for the most variable outcome measure (full macular thickness in glaucoma eyes) demonstrates that after removing of outliers (nonrandom noise), the intrasession variability reaches a maximum of 3 μm.

of the measurement grid. For all 12 layers and groups, the M estimation method determined that absolute MEs greater than approximately 3 μm were outliers; in other words, the 2.5 percentile ME was always larger than −3 μm and the 97.5 ME was always smaller than 3 μm. Therefore, a conservative rule would be that the 95% prediction interval bounds for thickness ME are between −3 and +3 μm for all layers in glaucoma and healthy eyes after exclusion of outliers. As an example, the 95% prediction boundaries for full macular thickness in glaucoma patients are shown Figure 5. Given the high variability in the nasal most column and the most superior row of the 8 × 8 array of superpixels, we recalculated the prevalence of superpixels demonstrating a decrease in thickness greater than 3 μm in the remaining 7 × 7 grid after excluding the nasal most column and the most superior row. Table 3 demonstrates the prevalence of such superpixels for various macular parameters based on diagnostic group after exclusion of the most superior row and the most nasal column. The prevalence of outliers decreased to less than 2% for all parameters, which is approximately 1 in 49 superpixels except for full thickness and mRNFL. Simply put and as a rule of thumb, if more than one superpixel is showing a decrease of more than 3 μm on a follow-up Posterior Pole image for RGC, GC/IPL, or GCC in the central 7 × 7 array, it would be likely that the change is real.

We also compared SDs of ME in the glaucoma and healthy patients after removing outliers (Table 4). The SDs of ME were significantly higher in the glaucoma group for the IPL and GCIPL layers ($P = 0.002$ and $= 0.014$, respectively); whereas the SD for

Table 3. Proportion of Superpixels on the Posterior Pole Algorithm of the Spectralis SD-OCT That Could Demonstrate a Random Decrease in Thickness Greater Than 3 μm After Excluding the Most Nasal and Superior Rows of the 8 × 8 Array of Superpixels, Where the Variability is High, Based on Diagnostic Group

Layer	Group	Number of Superpixels	Proportion of Outliers
IPL	Glaucoma	14,973	0.9%
IPL	Healthy	3,087	0.4%
GCL	Glaucoma	14,974	1.4%
GCL	Healthy	3,087	1.2%
RNFL	Glaucoma	14,975	3.3%
RNFL	Healthy	3,087	2.7%
GC+IPL	Glaucoma	14,994	1.0%
GC+IPL	Healthy	3,087	1.7%
GCC	Glaucoma	14,994	1.4%
GCC	Healthy	3,087	0.5%
Full	Glaucoma	14,978	6.4%
Full	Healthy	3,087	3.1%

ME of mRNFL was higher in the healthy subjects ($P = 0.003$); however, the magnitude of the differences was at most 0.2 μm and would not be considered clinically important. Comparisons between left versus right eyes after controlling for diagnosis and location showed that left versus right eye differences were never larger than 22% (translating into 0.2 μm) and none were significant. Scatter plots showing the relationship between the proportions of outliers at individual superpixels as a function of mean superpixel thickness showed no linear relationship between the mean superpixel thickness and proportion of outliers for IPL, GCL, GC/IPL, GCC, and full thickness layers. Nevertheless, there was an apparent increase at very low thickness measurements in both healthy and glaucoma groups for IPL, GCL, and GC/IPL layers (Fig. 6). The mRNFL measurements tended to have increased numbers of outliers with increasing thickness measurements at individual superpixels (Fig. 6); this finding was consistent among glaucoma and healthy subjects. Linear regression showed no relationship between age or axial length with ME ($P > 0.999$ for both, Fig. 7 top left and right). There was no correlation between ME and image quality factor regardless of diagnostic group ($P > 0.999$, Fig. 7 bottom).

Table 4. Comparison of Mean Absolute Deviation for Measurement Errors across Various Layers

	Glaucoma		Normal			
Outcome Measure	MAD, μm	SE	MAD, μm	SE	Percent Difference	*P* Value
IPL	0.846	0.010	0.694	0.048	−19.7%	0.0020
GCL	0.727	0.013	0.656	0.059	−10.3%	0.2356
mRNFL	0.778	0.015	0.997	0.071	24.6%	0.0027
GCL/IPL	1.061	0.018	0.855	0.082	−21.5%	0.0140
GCC	1.082	0.020	1.147	0.094	5.9%	0.4953
FULL	0.938	0.027	0.995	0.125	5.9%	0.6543

Mean absolute deviation is directly proportional to standard deviation (SD ≈ 1.2 MAD) and is used here to compare SDs across the various layers and groups with the Fligner modification of the Brown-Forsythe method.

Tables 5 and 6 describe the ICCs and their 95% confidence intervals for the SD-OCT Glaucoma Hemifield sectors in the healthy and glaucoma groups, respectively. The ICCs were very good to excellent across the macular region. Full macular thickness measurements had the best reproducibility across all sectors in the heathy subjects (ICC: 0.981–0.998), whereas it had the lowest ICCs in glaucoma patients (0.879–0.927). The mRNFL measurements had overall similar reproducibility compared with other layers (ICC: 0.916–0.994 in the healthy group and 0.925–0.996 in the glaucoma group).

Discussion

One of the main goals of glaucoma management is detection of disease deterioration as early as possible so that appropriate treatment alterations can be made. Spectral-domain OCT is rapidly becoming the automated device of choice for detecting presence of glaucoma or disease worsening through structural measures, although solid data are scarce with regard to the latter task. Most available studies have relied on measurement of the circumpapillary RNFL thickness for detecting structural change in glaucoma.[23–28] More recently, the use of macular thickness

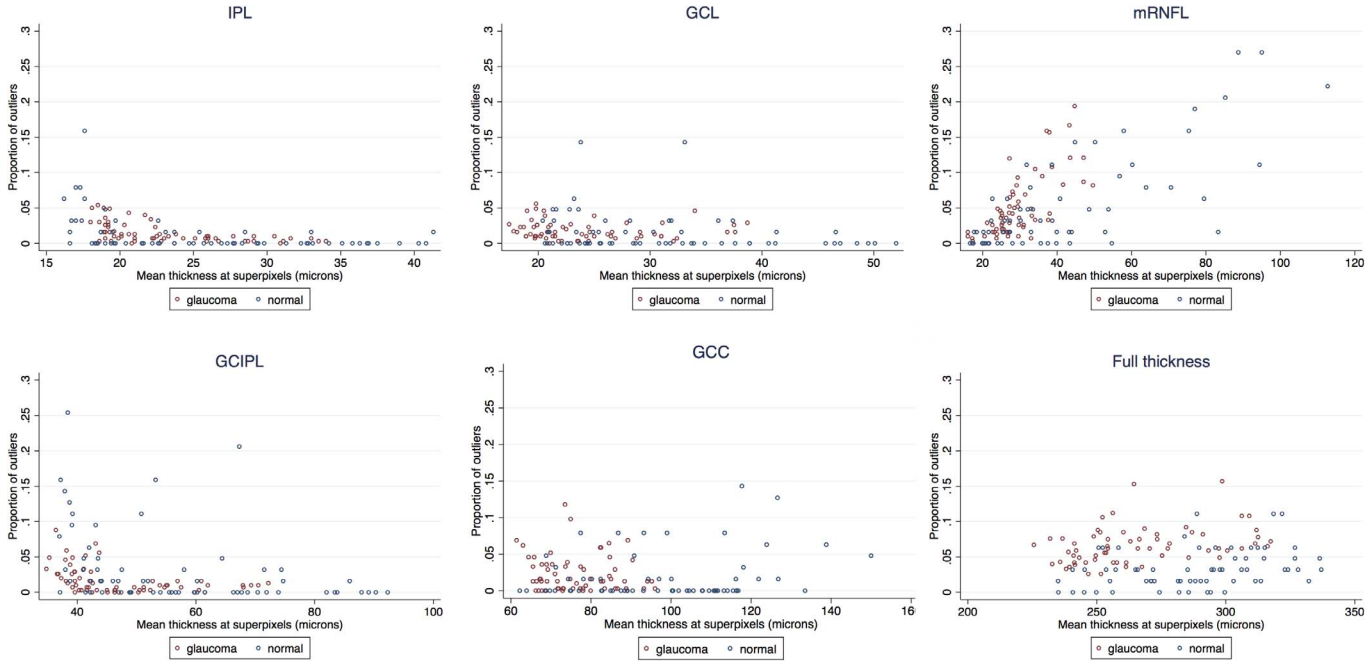

Figure 6. Scatter plots demonstrate the proportion of outliers across superpixels as a function of mean superpixel thickness for the six macular outcome measures according to diagnosis: (*top left*) IPL; (*top middle*) GCL; (*top right*) RNFL, (*bottom left*) GC/IPL, (*bottom middle*) GCC layer, and (*bottom right*) full macular thickness.

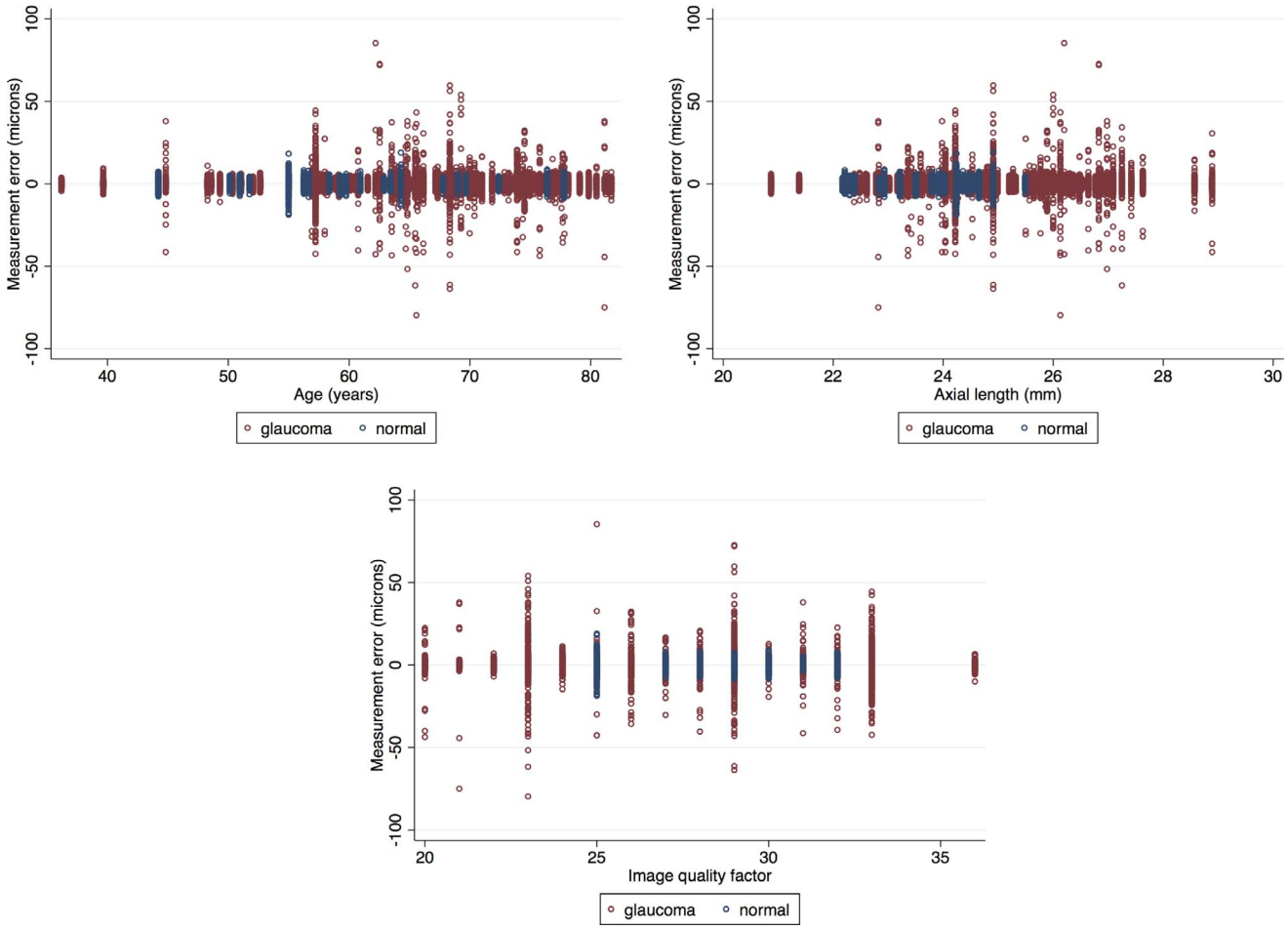

Figure 7. Association of variability of macular full thickness measurements (measurement errors at 3° × 3° superpixels) with age (*top left*), axial length (*top right*), and image quality factor (*bottom*). None of the corresponding regression coefficients were different from zero.

measurements has been explored for this purpose because such measurements have demonstrated very good reproducibility with current SD-OCT devices.[2,13-19,29] However, all these studies reported reproducibility of macular thickness parameters for global measures or larger hemiretinal areas or sectors and only addressed full macular thickness, GCC, or GC/IPL measurements. Reproducibility data have not been previously reported for individual retinal layers of interest in glaucoma, especially for smaller areas of the macula. We explored and compared the intrasession variability of various macular layers or combinations of layers in glaucoma and healthy subjects. We also evaluated factors that may influence the variability of such measurements.

Our findings can be summarized as follows: (1) variability of macular measurements were uniform when such measurements were performed at the level of superpixels 3° × 3° in size, (2) measurement variability could be divided into a Gaussian component (random noise) and outliers that were found to be nonrandom and were mostly explained by topography and possible imaging-related technical issues (see below), (3) after removal of the outliers, variability at the level of superpixels was very similar regardless of the layer(s) being measured (at most 3 μm for all layers), (4) while the variability was overall similar between glaucoma and healthy subjects, in some layers (IPL, GCIPL) the variability was slightly higher in glaucoma subjects whereas the mRNFL variability was higher in healthy subjects; such differences were not clinically significant, (5) the proportion of outliers was least for the IPL and GCL parameters (1.2%–1.5% for IPL and 1.5%–1.8% for GCL), and (6) age, axial length, and image quality did not have a significant influence on the magnitude of measurement error; superpixels located on the

Table 5. Intraclass Correlation Coefficients for Macular OCT Sector Thicknesses as Defined by Um et al.[22] for Various Outcome Measures in 21 Eyes of 21 Healthy Subjects

Cluster	IPL	GCL	mRNFL	GC/IPL
Superior zone 1	0.925 (0.855–0.966)	0.967 (0.935–0.986)	0.916 (0.838–0.962)	0.970 (0.940–0.987)
Inferior zone 1	0.925 (0.855–0.966)	0.967 (0.935–0.986)	0.916 (0.838–0.962)	0.970 (0.940–0.987)
Superior zone 2	0.941 (0.883–0.973)	0.912 (0.830–0.960)	0.981 (0.962–0.992)	0.972 (0.944–0.988)
Inferior zone 2	0.950 (0.902–0.978)	0.984 (0.967–0.993)	0.946 (0.894–0.976)	0.984 (0.968–0.993)
Superior zone 3	0.941 (0.883–0.973)	0.912 (0.830–0.960)	0.981 (0.962–0.992)	0.972 (0.944–0.988)
Inferior zone 3	0.923 (0.850–0.965)	0.746 (0.558–0.877)	0.988 (0.976–0.995)	0.925 (0.854–0.966)
Superior zone 4	0.943 (0.887–0.974)	0.988 (0.976–0.995)	0.924 (0.878–0.956)	0.964 (0.928–0.984)
Inferior zone 4	0.961 (0.922–0.983)	0.968 (0.936–0.986)	0.931 (0.864–0.969)	0.981 (0.961–0.992)
Superior zone 5	0.950 (0.901–0.978)	0.988 (0.975–0.995)	0.991 (0.984–0.995)	0.972 (0.943–0.987)
Inferior zone 5	0.926 (0.857–0.967)	0.960 (0.919–0.982)	0.994 (0.988–0.997)	0.965 (0.930–0.985)

Numbers in parentheses represent 95% confidence intervals.

nasal and superior borders of the measurement grid had the highest variability especially for mRNFL.

Patterns of glaucoma progression have not been well defined for macular thickness measures. The definition of such patterns would first require measurement of variability in smaller regions within the macula. Knowledge of the magnitude of variability and its determinant is necessary if macular measures from SD-OCT are to be used for the detection of progression. Specifically, given the fairly low frequency of OCT tests in the clinical setting (every 6 or 12 months), event analyses may be better suited for detection of disease deterioration at least during the first few years of follow-up because trend analyses require a large number of data points to achieve an adequate power.[30] Although the 3° × 3° superpixels do not have anatomical or physiological correlates, given the fairly small size of such superpixels, patterns of progression can be better studied with this approach as compared with larger anatomical divisions frequently used by other devices such as the wedge-shaped sectors with Cirrus HD-OCT.

Based on exploratory histograms and normal quantile plots, we posited that the variability within macular images could be divided into a random component (real measurement error), which had Gaussian distribution, and a small nonrandom component (called outliers by the M estimation method). We observed that after removal of outliers, the variability was constant across locations, eyes, and diagnostic groups, and that the upper and lower bounds of the 95% prediction limits for ME did not exceed 3 μm for any layer. This strongly suggests that macular thickness measurements can serve as valid outcomes for detection of glaucoma deterioration. The outliers had a nonrandom topographic distribution across all layers and were more frequent on the most nasal edge of the macular volume scan near the optic disc and superiorly (Fig. 4). The proportion of outliers was related to mean superpixel thickness only for mRNFL (Fig. 6); there was a direct relationship between the mRNFL thickness and proportion of outliers across the 64 superpixels. A topographic analysis of the frequency of outliers showed that the mRNFL had the highest proportion of such measurements among all the macular layers. This pattern is consistent with the location of major retinal vessels because most of the larger retinal vessels are located in the RNFL layer. Approximately 13% of the average RNFL thickness and as

Table 5. Extended

Cluster	GCC	Full
Superior zone 1	0.978 (0.956–0.990)	0.981 (0.962–0.992)
Inferior zone 1	0.978 (0.956–0.990)	0.981 (0.962–0.992)
Superior zone 2	0.983 (0.966–0.993)	0.992 (0.984–0.996)
Inferior zone 2	0.980 (0.956–0.991)	0.988 (0.976–0.995)
Superior zone 3	0.983 (0.966–0.993)	0.992 (0.984–0.996)
Inferior zone 3	0.983 (0.966–0.993)	0.991 (0.982–0.996)
Superior zone 4	0.983 (0.966–0.993)	0.995 (0.989–0.998)
Inferior zone 4	0.969 (0.938–0.986)	0.990 (0.981–0.996)
Superior zone 5	0.985 (0.970–0.993)	0.996 (0.992–0.998)
Inferior zone 5	0.995 (0.989–0.998)	0.998 (0.995–0.999)

much as 25% of the RNFL thickness in the temporal arcuate sectors is contributed by blood vessels depending on the number of blood vessels within these sectors.[31,32] Given the topography of outliers discussed above, it seems that the presence of blood vessels is a major source of variability for mRNFL measurements while other sources such as peripapillary atrophic changes could be contributory as well. A review of the segmented B-scans also demonstrated that the segmentation algorithm failed mostly on the very nasal edge of the OCT images. We detected no increased variability around the fovea as expected based on prior work by Knighton and colleagues.[33] One reason for this finding could be the fairly large size of superpixels so that the higher variability of thickness measurements for superpixels around the fovea is actually averaged with the lower variability of thickness measurements in the surrounding areas. Interestingly, for IPL, GCL, and GC/IPL layers, although there was no relationship between mean superpixel thickness and proportion of outliers, there was a significant spike in the latter at the very low end of thickness measurements implying that segmentation failure could become common with thinning of the aforementioned layers. Although the quality of macular SD-OCT images is very good with current technology, the segmentation of individual retinal layers is still not flawless and segmentation failure can understandably become more frequent with progressive thinning of inner retinal layers. Because of significant thinning, even manual correction of segmentation would be suboptimal under such circumstances.

Our findings are promising with regard to detection of glaucoma progression in more advanced stages of glaucoma. One could argue that until the late stages of glaucoma when the individual retinal layers suddenly show increased variability, various macular thickness parameters could be used for this purpose and that parameters based on combination of layers, especially GCC, might be especially well suited to this aim because the proportion of outliers is essentially stable over the entire range of thickness measurements. The latter finding is likely a result of averaging of the noise. Also, segmentation of the entire GCC thickness without measuring individual layers is a technically less demanding imaging task.

To compare the results of our study with prior investigations, we also calculated the ICCs for macular sectors as defined by Um and colleagues.[22] Our results were consistent with the previous literature demonstrating very good to excellent reproducibility of measurements across such fairly large areas of the macula. This is partly a result of averaging of the noise across larger sectors or hemiretinal regions of the macula. The ICCs were highest for full macular thickness measurements in healthy subjects; interestingly, full macular thickness had the lowest ICCs in glaucoma patients. From an imaging point of view, automated detection of the vitreous-internal limiting membrane interface and retinal pigment epithelium is a fairly straightforward task and therefore, it is less prone to failure. One has to keep in mind that the variability of combination of layers in our study would have been expected to be higher than previous reports merely as a result of our approach. Because the combination of layers were calculated by adding individual layers rather than primarily segmenting the outer and inner borders, the noise in our data would be expected to be higher. Although our findings are promising, the high reproducibility of local OCT measurements needs to be established between sessions as well. Available studies suggest that the magnitude of intersession variability is very close to that of intrasession variability.[13,19]

Table 6. Intraclass Correlation Coefficients for Macular OCT Sectors as Defined by Um et al.[22] as a Function of Outcome Measures in 102 eyes of 102 Glaucoma Patients

Cluster	IPL	GCL	mRNFL	GC/IPL
Superior zone 1	0.978 (0.970–0.985)	0.990 (0.987–0.993)	0.925 (0.898–0.946)	0.994 (0.991–0.996)
Inferior zone 1	0.978 (0.970–0.985)	0.990 (0.987–0.993)	0.925 (0.898–0.946)	0.994 (0.991–0.996)
Superior zone 2	0.968 (0.956–0.977)	0.989 (0.984–0.992)	0.990 (0.986–0.993)	0.993 (0.990–0.995)
Inferior zone 2	0.974 (0.964–0.982)	0.990 (0.986–0.993)	0.938 (0.916–0.956)	0.992 (0.989–0.994)
Superior zone 3	0.968 (0.956–0.977)	0.989 (0.984–0.992)	0.990 (0.986–0.993)	0.993 (0.990–0.995)
Inferior zone 3	0.960 (0.944–0.971)	0.980 (0.972–0.986)	0.983 (0.976–0.988)	0.983 (0.976–0.988)
Superior zone 4	0.898 (0.861–0.926)	0.952 (0.933–0.966)	0.994 (0.991–0.995)	0.945 (0.924–0.961)
Inferior zone 4	0.938 (0.916–0.956)	0.974 (0.965–0.982)	0.993 (0.990–0.995)	0.979 (0.971–0.985)
Superior zone 5	0.808 (0.746–0.859)	0.910 (0.878–0.936)	0.996 (0.994–0.997)	0.921 (0.893–0.944)
Inferior zone 5	0.886 (0.846–0.918)	0.955 (0.934–0.968)	0.996 (0.994–0.997)	0.945 (0.924–0.961)

Numbers in parentheses represent 95% confidence intervals. Please refer to Figure 2 for definition of zones.

Our results indicate that measurement variability is uniform within and across eyes beyond the nasal most column and the most superior row of superpixels and as a result, it may not be necessary to define an individual variability space for each patient. We demonstrated that after removing the most nasal column and the most superior row of the 8×8 array of superpixels, the prevalence of outliers was less than 2% for most parameters (except mRNFL and full macular thickness) in the remaining 49 superpixels. That means that if more than one superpixel demonstrates a decrease of more than 3 μm for any of the macular parameters, that amount of change would likely be real. Some imaging devices such as the Heidelberg Retina Tomograph obtain three images at each session to minimize overall variability. Given the excellent local measurement reproducibility shown in this study, we advocate that this may not be necessary with SD-OCTs. Variability had essentially the same pattern and magnitude in the glaucoma and healthy eyes, although for IPL and GCIPL layers, the healthy eyes tended to have slightly lower variability (at most 20%); in contrast, mRNFL had higher variability in healthy eyes (25%). The magnitude of the difference seems very small and is likely of no clinical significance.

At present, it is not entirely clear which macular parameter would have the best performance for detection of glaucoma deterioration, given the very good reproducibility of all such measurements. It has been reported that the dynamic range, defined as the number of progression 'steps' exceeding test-retest variability, may be smaller for structural testing than for perimetry.[34] The most appropriate approach short of a longitudinal comparison would be to compare the measurement variability to the dynamic range for each of the macular layers. Although the performance of GC/IPL and GCC parameters has been reported to be better than that of full thickness macular parameters for detection of early glaucoma, this does not necessarily mean that the same macular layers would perform better for detection of glaucoma deterioration. We are currently carrying out the next phase of this study to address this issue.

We also examined the association of age, axial length, and quality of the macular images with measurement variability. Neither age nor axial length had an influence on the measurement variability regardless of diagnostic group. The same was also

Table 6. Extended

Cluster	GCC	Full
Superior zone 1	0.993 (0.991–0.995)	0.879 (0.837–0.912)
Inferior zone 1	0.993 (0.991–0.995)	0.879 (0.837–0.912)
Superior zone 2	0.997 (0.996–0.997)	0.888 (0.848–0.919)
Inferior zone 2	0.983 (0.977–0.988)	0.883 (0.842–0.915)
Superior zone 3	0.997 (0.996–0.998)	0.888 (0.848–0.919)
Inferior zone 3	0.986 (0.980–0.990)	0.927 (0.900–0.948)
Superior zone 4	0.992 (0.989–0.994)	0.871 (0.827–0.907)
Inferior zone 4	0.994 (0.992–0.996)	0.855 (0.806–0.895)
Superior zone 5	0.996 (0.994–0.997)	0.920 (0.891–0.942)
Inferior zone 5	0.994 (0.992–0.996)	0.891 (0.852–0.921)

true with respect to image quality. However, the quality factor reported by SD-OCT devices is a global measure, and therefore, it may not reflect local variations in image quality that could potentially affect variability at the superpixel level. Also, we should note that low-quality images, defined as those with quality factor less than 15, were already excluded from all analyses and therefore, our findings are not generalizable to such images.

The results presented here represent an optimal-case scenario where most of the B-scans that belong to the same imaging session have good quality and are free of significant artifacts. Therefore, one would expect that the variability of macular images might not be as good under typical clinical situations. Another caveat is that manual correction of image segmentation could have led to decreased variability; however, analyses of uncorrected data showed the results to be very similar, except for higher variability of the mRNFL layer thickness (data not shown). The average axial length was 24.8 mm in the glaucoma group; hence, our results may not be generalizable to eyes with axial high myopia, in which the incidence of image artifacts is higher. One could argue that the intersession variability is a more important parameter to consider when detection of glaucoma deterioration is desired. The published literature on global and sectoral intersession variability suggests that most of the intersession variability is explained by the intrasession variability of macular SD-OCT and that the magnitude of intersession variability is very similar to that of intrasession variability in larger areas of the macula.[13,19] Our team is carrying out a prospective study to compare intrasession and intersession variability of the macular thickness measures explored in this study to better clarify this issue.

Our findings have significant clinical implications with regard to detection of glaucoma progression. There is no well-established algorithm for this task based on macular OCT images. According to the 95% prediction intervals of variability defined in this study, patterns of progression can be identified in glaucoma patients and functional correlations sought; also, rules for establishing disease deterioration based on the magnitude and topography of change and other criteria such as clustering of changing superpixels can be created. The sensitivity and specificity of such criteria needs to be subsequently measured in cohorts of glaucoma and healthy subjects.

In summary, we measured the local intrasession variability for various macular parameters derived from SD-OCT images of the central macula including individual layers of the inner retina. The variability was very low and uniform across eyes and layers, with some nonrandom variability that could be explained by the topography of the macular images and the thickness of the layers of interest. These findings are most relevant for designing algorithms for defining glaucoma progression with structural means.

Acknowledgments

Supported by a Mid-Career Clinician-Scientist Grant from the American Glaucoma Society (KNM) and a National Institutes of Health Mentored Patient-oriented Research Career Development Award (5K23EY022659; KNM) and an unrestricted Departmental Grant from Research to Prevent Blindness.

Presented as a poster at the annual meeting of the Association for Research in Vision and Ophthalmology, May 3–7, 2015, Denver, CO, USA.

Disclosure: **A. Miraftabi,** None; **N. Amini,** None; **J. Gornbein,** None; **S. Henry,** None; **P. Romero,**

None; **A.L. Coleman,** None; **J. Caprioli,** None; **K. Nouri-Mahdavi,** None

References

1. Guedes V, Schuman JS, Hertzmark E, et al. Optical coherence tomography measurement of macular and nerve fiber layer thickness in normal and glaucomatous human eyes. *Ophthalmology*. 2003;110:177–189.
2. Tan O, Chopra V, Lu AT, et al. Detection of macular ganglion cell loss in glaucoma by Fourier-domain optical coherence tomography. *Ophthalmology*. 2009;116:2305–2314, e2301–e2302.
3. Mwanza JC, Durbin MK, Budenz DL, et al. Glaucoma diagnostic accuracy of ganglion cell-inner plexiform layer thickness: comparison with nerve fiber layer and optic nerve head. *Ophthalmology*. 2012;119:1151–1158.
4. Nouri-Mahdavi K, Nowroozizadeh S, Nassiri N, et al. Macular ganglion cell/inner plexiform layer measurements by spectral domain optical coherence tomography for detection of early glaucoma and comparison to retinal nerve fiber layer measurements. *Am J Ophthalmol*. 2013;156:1297–1307, e1292.
5. Begum VU, Addepalli UK, Yadav RK, et al. Ganglion cell-inner plexiform layer thickness of high definition optical coherence tomography in perimetric and preperimetric glaucoma. *Invest Ophthalmol Vis Sci*. 2014;55:4768–4775.
6. Curcio CA, Allen KA. Topography of ganglion cells in human retina. *J Comp Neurol*. 1990;300:5–25.
7. Sung MS, Yoon JH, Park SW. Diagnostic validity of macular ganglion cell-inner plexiform layer thickness deviation map algorithm using cirrus HD-OCT in preperimetric and early glaucoma. *J Glaucoma*. 2014;23:e144–e151.
8. Hood DC, Raza AS, de Moraes CG, Johnson CA, Liebmann JM, Ritch R. The nature of macular damage in glaucoma as revealed by averaging optical coherence tomography data. *Transl Vis Sci Technol*. 2012;1(1):3.
9. Lisboa R, Paranhos A, Weinreb RN, Zangwill LM, Leite MT, Medeiros FA. Comparison of different spectral domain OCT scanning protocols for diagnosing preperimetric glaucoma. *Invest Ophthalmol Vis Sci*. 2013;54:3417–3425.
10. Na JH, Sung KR, Baek S, et al. Detection of glaucoma progression by assessment of segmented macular thickness data obtained using spectral domain optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2012;53:3817–3826.
11. Sung KR, Sun JH, Na JH, Lee JY, Lee Y. Progression detection capability of macular thickness in advanced glaucomatous eyes. *Ophthalmology*. 2012;119:308–313.
12. Lee KS, Lee JR, Na JH, Kook MS. Usefulness of macular thickness derived from spectral-domain optical coherence tomography in the detection of glaucoma progression. *Invest Ophthalmol Vis Sci*. 2013;54:1941–1949.
13. Garas A, Vargha P, Holló G. Reproducibility of retinal nerve fiber layer and macular thickness measurement with the RTVue-100 optical coherence tomograph. *Ophthalmology*. 2010;117:738–746.
14. Mwanza JC, Oakley JD, Budenz DL, Chang RT, Knight OJ, Feuer WJ. Macular ganglion cell-inner plexiform layer: automated detection and thickness reproducibility with spectral domain-optical coherence tomography in glaucoma. *Invest Ophthalmol Vis Sci*. 2011;52:8323–8329.
15. Matlach J, Wagner M, Malzahn U, Göbel W. Repeatability of peripapillary retinal nerve fiber layer and inner retinal thickness among two spectral domain optical coherence tomography devices. *Invest Ophthalmol Vis Sci*. 2014;55:6536–6546.
16. Wadhwani M, Bali SJ, Satyapal R, et al. Test-retest variability of retinal nerve fiber layer thickness and macular ganglion cell-inner plexiform layer thickness measurements using spectral-domain optical coherence tomography. *J Glaucoma*. 2015;24;e109–e115.
17. Carpineto P, Aharrh-Gnama A, Ciciarelli V, Mastropasqua A, Di Antonio L, Toto L. Reproducibility and repeatability of ganglion cell-inner plexiform layer thickness measurements in healthy subjects. *Ophthalmologica*. 2014;232:163–169.
18. Kim KE, Yoo BW, Jeoung JW, Park KH. Long-term reproducibility of macular ganglion cell analysis in clinically stable glaucoma patients. *Invest Ophthalmol Vis Sci*. 2015;56:4857–4864.
19. Ghasia FF, El-Dairi M, Freedman SF, Rajani A, Asrani S. Reproducibility of spectral-domain optical coherence tomography measurements in adult and pediatric glaucoma. *J Glaucoma*. 2015;24:55–63.
20. Johnson CA, Sample PA, Cioffi GA, Liebmann JR, Weinreb RN. Structure and function evaluation (SAFE): I. criteria for glaucomatous visual field loss using standard automated perimetry

(SAP) and short wavelength automated perimetry (SWAP). *Am J Ophthalmol*. 2002;134:177–185.
21. Morton B, Brown ABF. Robust test for the equality of variances. *American Statistical Association*. 1974;69:364–367.
22. Um TW, Sung KR, Wollstein G, Yun SC, Na JH, Schuman JS. Asymmetry in hemifield macular thickness as an early indicator of glaucomatous change. *Invest Ophthalmol Vis Sci*. 2012;53:1139–1144.
23. Wollstein G, Schuman JS, Price LL, et al. Optical coherence tomography longitudinal evaluation of retinal nerve fiber layer thickness in glaucoma. *Arch Ophthalmol*. 2005;123:464–470.
24. Leung CK, Cheung CY, Weinreb RN, et al. Evaluation of retinal nerve fiber layer progression in glaucoma with optical coherence tomography guided progression analysis. *Invest Ophthalmol Vis Sci*. 2010;51:217–222.
25. Park SB, Sung KR, Kang SY, Kim KR, Kook MS. Comparison of glaucoma diagnostic capabilities of Cirrus HD and Stratus optical coherence tomography. *Arch Ophthalmol*. 2009;127:1603–1609.
26. Leung CK, Cheung CY, Weinreb RN, et al. Evaluation of retinal nerve fiber layer progression in glaucoma: a comparison between the fast and the regular retinal nerve fiber layer scans. *Ophthalmology*. 2011;118:763–767.
27. Leung CK, Chiu V, Weinreb RN, et al. Evaluation of retinal nerve fiber layer progression in glaucoma: a comparison between spectral-domain and time-domain optical coherence tomography. *Ophthalmology*. 2011;118:1558–1562.
28. Leung CK, Liu S, Weinreb RN, et al. Evaluation of retinal nerve fiber layer progression in glaucoma a prospective analysis with neuroretinal rim and visual field progression. *Ophthalmology*. 2011;118:1551–1557.
29. Francoz M, Fenolland JR, Giraud JM, et al. Reproducibility of macular ganglion cell-inner plexiform layer thickness measurement with cirrus HD-OCT in normal, hypertensive and glaucomatous eyes. *Br J Ophthalmol*. 2014;98:322–328.
30. Chauhan BC, Garway-Heath DF, Goni FJ, et al. Practical recommendations for measuring rates of visual field change in glaucoma. *Br J Ophthalmol*. 2008;92:569–573.
31. Hood DC, Fortune B, Arthur SN, et al. Blood vessel contributions to retinal nerve fiber layer thickness profiles measured with optical coherence tomography. *J Glaucoma*. 2008;17:519–528.
32. Hood DC, Kardon RH. A framework for comparing structural and functional measures of glaucomatous damage. *Prog Retin Eye Res*. 2007;26:688–710.
33. Knighton RW, Gregori G. The shape of the ganglion cell plus inner plexiform layers of the normal human macula. *Invest Ophthalmol Vis Sci*. 2012;53:7412–7420.
34. Gardiner SK, Ren R, Yang H, Fortune B, Burgoyne CF, Demirel S. A method to estimate the amount of neuroretinal rim tissue in glaucoma: comparison with current methods for measuring rim area. *Am J Ophthalmol*. 2014;157:540–549, e541–e542.