GLAUCOMA

Comparison of measurement error of Cirrus HD-OCT and Heidelberg Retina Tomograph 3 in patients with early glaucomatous visual field defect

Alexander A. Shpak · Maria K. Sevostyanova · Svetlana N. Ogorodnikova · Irina N. Shormaz

Abstract

Background To compare the measurement error of Cirrus HD-OCT and Heidelberg Retina Tomograph 3 (HRT3) in patients with early glaucomatous visual field defect.

Methods Thirty-nine consecutive patients (39 eyes) with early glaucomatous visual field defect were included. One eye of a patient was examined with Cirrus HD-OCT and HRT3 in one session. Each instrument was used by two operators, each taking two measurements in turn. We performed measurements of average retinal nerve fiber layer (RNFL) thickness and RNFL thickness in four quadrants with Cirrus HD-OCT and 13 stereometric parameters of the optic nerve head and RNFL with HRT3.

Results Cirrus HD-OCT parameters performed much better than HRT3 stereometric parameters. Mean (for both operators) intraoperator within-subject coefficient of variation of the best Cirrus HD-OCT parameter (average RNFL thickness) was 2.6–2.7 times lower than the best HRT3 parameters [mean cup depth and rim area ($P<0.001$)]. Mean intraoperator variability of RNFL thickness in quadrants (except nasal quadrant) was also significantly lower with OCT than with HRT. The interoperator within-subject coefficients of variation for both average RNFL thickness and RNFL thickness in all quadrants were significantly lower than the interoperator variability of best HRT3 parameter [mean cup depth ($P<0.001$)]. The within-subject coefficient of variation of the average/mean RNFL thickness assessed by both instruments was 5.4–7.3 times

lower for Cirrus HD-OCT. Among HRT3 parameters, mean cup depth, rim area and linear cup/disk ratio were the least variable, while cup volume, cup area and cup/disc area ratio were the most variable parameters.

Conclusions Cirrus HD-OCT provides excellent intrasession intra- and interoperator repeatability of the RNFL measurements, especially of the average RNFL thickness in primary open-angle glaucoma (POAG) patients and POAG suspects with early glaucomatous visual field defect. The measurement error (expressed as within-subject coefficient of variation) of RNFL measurements by Cirrus HD-OCT is much lower than the error of HRT3 measurements of stereometric parameters of the optic nerve head and RNFL.

Keywords Early primary open-angle glaucoma · Spectral-domain optical coherence tomography · Cirrus HD-OCT · Heidelberg Retina Tomograph 3 · Repeatability · Within-subject coefficient of variation

A. A. Shpak (✉) · M. K. Sevostyanova · S. N. Ogorodnikova · I. N. Shormaz
The S. Fyodorov
Eye Microsurgery Complex — Federal State Institution,
Beskudnikovsky Blvd. 59-A,
Moscow 127486, Russian Federation
e-mail: a_shpak@inbox.ru

Introduction

Spectral-domain optical coherence tomography (SD-OCT) and confocal scanning laser ophthalmoscopy (CSLO) are among the best known imaging methods widely used for diagnosing and following of primary open-angle glaucoma (POAG). The repeatability and variability of measurements are important features of any diagnostic method. There are several studies of repeatability and variability of stereometric parameters of the optic disc and retinal nerve fiber layer (RNFL) measurements by CSLO with Heidelberg Retina Tomograph (HRT) [1–7], but their results are sometimes conflicting. A number of recent studies [8–14] have dealt with repeatability and variability of RNFL

measurements by SD-OCT with Cirrus HD-OCT [8–12] or other instruments [13, 14], but to our knowledge no comparison with HRT was performed. The purpose of the present study was to compare the intrasession intraoperator and interoperator repeatability and within-subject coefficients of variation of measurements of the RNFL by Cirrus HD-OCT and stereometric parameters of the optic nerve head (ONH) and RNFL by HRT3 in patients with early glaucomatous visual field defect.

Material and methods

Subjects

Patients were recruited consecutively between January and April 2010 at the out-patient department of the S. Fyodorov Eye Microsurgery Complex — Federal State Institution, Moscow, Russian Federation. The enrolled patients underwent recent complete ophthalmologic evaluation. The inclusion criteria were a best-corrected visual acuity of 20/40 or better, spherical equivalent refraction within ±5.0 D, astigmatism within ±3.0 D, clear media, open angles at gonioscopy, controlled intraocular pressure (IOP), reliable repeated standard automated perimetry (Humphrey Field Analyzer II, Carl Zeiss Meditec Inc., Dublin, CA, USA) using the Swedish Interactive Threshold Algorithm (SITA) Standard 30–2 program (perimetry was considered reliable if false positives, fixation losses and false negatives did not exceed 20% and no observable testing artefacts were present). The exclusion criteria were poor fixation, pronounced miosis, history of ocular surgery within 6 months of enrollment date, coexisting retinal disease, uveitis, optic nerve abnormalities or non-glaucomatous optic neuropathy. A total of 39 patients (39 eyes) were enrolled. There were 12 men and 27 women. Mean age was 62.9±8.7, range 35–76 years. One eye of each patient was included; one eye was randomly selected if both eyes were suitable for this study. Early POAG was diagnosed in 30 patients with glaucomatous ONH abnormalities (cup-to-disc ratio ≥0.5, cup-to disc ratio asymmetry ≥0.2, or focal thinning of the rim), a history of elevated IOP prior to study entry, and early glaucomatous visual field defect. Early glaucomatous visual field defect was established according to the staging system suggested by Mills et al. [15] by mean deviation≥−6.00 dB and at least one of the following: glaucoma hemifield test outside normal limits, pattern standard deviation significant at $P<0.05$ or a cluster of three or more points depressed below the 5% level in an expected location on pattern deviation plot, at least one of which depressed below the 1% level. Nine patients with suspicious but less pronounced ONH changes and early glaucomatous visual field defect were classified as POAG suspects. Best-

corrected visual acuity was not less than 20/25 in all eyes except one eye with mixed astigmatism and visual acuity 20/40. Spherical equivalent refraction ranged from −4.25 to +2.5 D. Astigmatism was not more than 2.0 D in all cases except one. This study adhered to the tenets of the Declaration of Helsinki and had local ethics committee approval, with informed consent obtained from all subjects.

Imaging methods

SD-OCT was performed with the Cirrus HD-OCT (Carl Zeiss Meditec Inc., Dublin, CA, USA), software version 4.5.1.11. The scan protocol was "Optic Disc Cube 200× 200" analyzed by "RNFL Thickness Analysis". Data acquisition and analysis software of Cirrus HD-OCT has been described in detail by many authors [8–12]. The analysis protocol is fully automatic, and does not need manual corrections. In this study, we analyzed overall average RNFL thickness and RNFL thickness in temporal, superior, nasal and inferior quadrants. Image quality (signal strength) on a 0–10 scale was 8 or more in 91.0% and 7 in 7.7% of cases; two images (1.3%) had a signal strength of 6.

CSLO was performed with HRT3 (Heidelberg Engineering, Heidelberg, Germany), software version 1.5.1.0. Maximal possible correction of astigmatism was performed. Each measurement consisted of three consecutive scans automatically averaged to a mean topographic image. Image quality was excellent or very good in 71.2%, and good in 26.9% of cases (mean standard deviation ≤30 μm); three images (1.9%) had mean standard deviations of 31, 33 and 38 μm respectively. A contour line was drawn manually to outline the optic disc margin on the mean topographic image. The contour line placement was thoroughly controlled using black-and-white and 3D modes of imaging. Once the contour line was drawn, the software automatically calculated all the ONH measurements. Thirteen stereometric parameters of the ONH and RNFL were analyzed in this study (see Tables 1 and 2).

Each patient was examined both by HRT3 and Cirrus HD-OCT without pupil dilation in one session on the same day; the order of the two methods was selected at random. SD-OCT was performed by two experienced operators (MKS and SNO), each taking two measurements in turn. After each measurement, the patient was asked to remove the head from the headrest and replace it again. CSLO was performed in the same fashion by two experienced operators (MKS and INS). HRT3 automatically transfers the contour line to all further images. When comparing operators to each other, we avoided this transfer by creating two separate entries for each patient in the patients' database.

Table 1 Intraoperator, interoperator repeatability and means of Cirrus HD-OCT and HRT3 measurements in patients with early glaucomatous visual field defect

Parameter	Repeatability			Mean
	Operator 1, mean (95% CI)	Operator 2, mean (95% CI)	Interoperator*, mean (95% CI)	
SD-OCT				
Average RNFL thickness (μm)	3.81 (2.96–4.65)	4.53 (3.52–5.53)	5.04 (3.92–6.16)	83.3
RNFL thickness in temporal quadrant (μm)	4.66 (3.62–5.69)	4.10 (3.19–5.01)	6.28 (4.89–7.68)	61.8
RNFL thickness in inferior quadrant (μm)	8.92 (6.94–10.89)	8.35 (6.49–10.20)	9.94 (7.74–12.16)[†]	104.3
RNFL thickness in superior quadrant (μm)	8.83 (6.87–10.76)	8.88 (6.91–10.85)	10.02 (7.80–12.25)	98.8
RNFL thickness in nasal quadrant (μm)	8.49 (6.61–10.38)	6.95 (5.41–8.49)	8.87 (6.90–10.83)	68.4
HRT				
Mean cup depth (mm)	0.019 (0.015–0.024)	0.032 (0.025–0.039)	0.045 (0.035–0.055)	0.195
Rim area (mm^2)	0.194 (0.151–0.237)[†]	0.170 (0.132–0.208)	0.423 (0.330–0.517)	1.365
Linear cup/disk ratio	0.078 (0.060–0.095)	0.089 (0.069–0.109)	0.147 (0.115–0.180)	0.479
Maximum cup depth (mm)	0.101 (0.079–0.124)	0.146 (0.114–0.179)	0.183 (0.143–0.224)	0.528
Cup shape measure	0.055 (0.043–0.068)	0.047 (0.037–0.058)	0.064 (0.050--0.079)	−0.169
Mean RNFL thickness (mm)	0.059 (0.046–0.072)	0.066 (0.051–0.081)	0.103 (0.080–0.126)	0.231
Rim volume (mm³)	0.081 (0.063–0.100)	0.096 (0.075–0.118)	0.168 (0.131–0.205)	0.349
Height variation contour (mm)	0.113 (0.088–0.138)	0.108 (0.084–0.132)	0.171 (0.133–0.209)[†]	0.373
RNFL cross sectional area (mm^2)	0.300 (0.233–0.367)	0.305 (0.237–0.373)	0.570 (0.444–0.697)	1.119
Cup/disc area ratio	0.073 (0.057–0.090)	0.093 (0.072–0.114)	0.159 (0.124–0.195)	0.259
Cup area (mm^2)	0.272 (0.211–0.332)[†]	0.170 (0.132–0.208)	0.357 (0.278–0.436)[†]	0.516
Cup volume (mm³)	0.073 (0.057–0.090)[†]	0.071 (0.055–0.087)[†]	0.104 (0.081–0.127)[†]	0.098
Disc area (mm^2)	0**	0**	0.239 (0.186–0.292)	1.881

* For HRT: each operator drew his own contour line (without automatic placement)

** Automatic placement of contour line provided for equality of the first and second measurements

[†] Individual subjects' standard deviations related to their means as shown by statistically significant Kendall's tau

Statistical Analysis

Statistical analysis was performed using the R software package version 2.10.1 (The R Foundation for Statistical Computing, http://www.r-project.org, accessed January, 28, 2010). The minimal required number of subjects was 31. This was calculated by setting the confidence interval as 25% on either side of the estimate of repeatability: $n = 1.96^2/(2*0.25^2*(m–1))$, where n is the number of subjects and m is the number of observations [16]. Repeatability of measurements (sometimes called the repeatability coefficient) was evaluated according to the Bland and Altman method [17] as $\sqrt{2}*1.96*S_w$, where S_w is the within-subject standard deviation. S_w was calculated as the square root of the average of within-subject variances (defined as half the square of the difference of two measurements). For the assessment of interoperator repeatability, we used only the first measurements of each operator. The standard error (se) of S_w was calculated as $se = S_w/\sqrt{2n}$ (where n is the number of subjects) and the 95% confidence interval (CI) was defined as $\pm 1.96*se$ [18]. We also calculated the within-subject coefficients of variation (CV_w) as $100*S_w/M$, where M is the mean of all paired measurements. Another method was used when individual subjects' standard deviations were related to their means, as confirmed by statistically significant Kendall's rank correlation tau (eight out of 52 cases). First, the within-subject coefficient of variation was defined as the square root of the mean of squared coefficients of variation for each subject (root mean square method) [19]. Then the S_w was calculated back from the coefficient of variation ($S_w = CV_w*M/100$), and repeatability was again defined as $\sqrt{2}*1.96*S_w$. The comparison of coefficients of variation was performed by the Z test [20]. Standard errors (se) of coefficients of variation were calculated as $se = CV_w/\sqrt{2n}$, where n is the number of subjects [21].

Results

Intraoperator and interoperator repeatability of all parameters studied are shown in Table 1. To simplify the

Table 2 Intraoperator and interoperator within-subject coefficients of variation of Cirrus HD-OCT and HRT3 measurements in patients with early glaucomatous visual field defect

Parameter	Operator 1, mean (95% CI)	Operator 2, mean (95% CI)	Interoperator*, mean (95% CI)
	SD-OCT		
Average RNFL thickness	1.65 (1.28–2.01)	1.96 (1.53–2.40)	2.18 (1.70–2.67)
RNFL thickness in temporal quadrant	2.71 (2.11–3.31)	2.40 (1.87–2.94)	3.67 (2.86–4.48)
RNFL thickness in inferior quadrant	3.09 (2.40–3.77)	2.88 (2.24–3.52)	3.44 (2.67–4.20)[†]
RNFL thickness in superior quadrant	3.23 (2.51–3.94)	3.24 (2.52–3.96)	3.66 (2.85–4.47)
RNFL thickness in nasal quadrant	4.47 (3.48–5.47)	3.67 (2.85–4.48)	4.68 (3.64–5.72)
	HRT		
Mean cup depth	3.61 (2.81–4.41)	5.88 (4.57–7.18)	8.38 (6.52–10.24)
Rim area	5.20 (4.05–6.36)[†]	4.44 (3.46–5.43)	11.14 (8.67–13.62)
Linear cup/disk ratio	5.79 (4.51–7.08)	6.75 (5.25–8.25)	11.17 (8.69–13.64)
Maximum cup depth	6.89 (5.36–8.42)	10.02 (7.80–12.25)	12.60 (9.80–15.39)
Cup shape measure	11.47 (8.93–14.02)	10.42 (8.11–12.74)	13.87 (10.79–16.95)
Mean RNFL thickness	9.24 (7.19–11.29)	10.20 (7.93–12.46)	15.94 (12.40–19.48)
Rim volume	8.52 (6.63–10.41)	9.82 (7.64–12.0)	17.13 (13.33–20.93)
Height variation contour	10.71 (8.33–13.09)	10.71 (8.33–13.08)	16.30 (12.68–19.92)[†]
RNFL cross sectional area	9.82 (7.64–12.0)	9.69 (7.54–11.84)	18.21 (14.17–22.25)
Cup/disc area ratio	10.10 (7.86–12.34)	13.17 (10.24–16.09)	22.51 (17.51–27.50)
Cup area	18.51 (14.40–22.61)[†]	12.22 (9.51–14.93)	25.27 (19.67–30.88)[†]
Cup volume	26.37 (20.52–32.22)[†]	26.68 (20.76–32.60)[†]	39.35 (30.61–48.08)[†]
Disc area	0**	0**	4.58 (3.56–5.59)

*, **, [†] See the footnotes to Table 1

comparison of different parameters, their means (taking into account all four measurements in each patient) are also shown in the right column.

Intraoperator and interoperator within-subject coefficients of variation of all parameters studied are shown in Table 2 and Fig. 1. In both Tables 1 and 2, parameters are listed in the order of increase of coefficients of variation (means of interoperator and averaged intraoperator data).

Disc area is included in both tables for information purposes only, because this parameter is not used directly in diagnosing of POAG or evaluation of POAG progression. Disc area as measured by HRT3 was from 1.6 to 2.5 mm^2 in most patients; in four patients it was less than 1.6 mm^2, and in one patient it was 3.05 mm^2.

Both operators demonstrated similar intraoperator repeatability data, with the exception of mean cup depth and

Fig. 1 Intraoperator and interoperator within-subject coefficients of variation (CV$_w$) of Cirrus HD-OCT and HRT3 measurements in patients with early glaucomatous visual field defect (the intraoperator data for both operators are averaged; disc area is not included). *, **, *** — SD-OCT data significantly lower than the best HRT parameter — mean cup depth (P<0.05, 0.01, and 0.001 respectively)
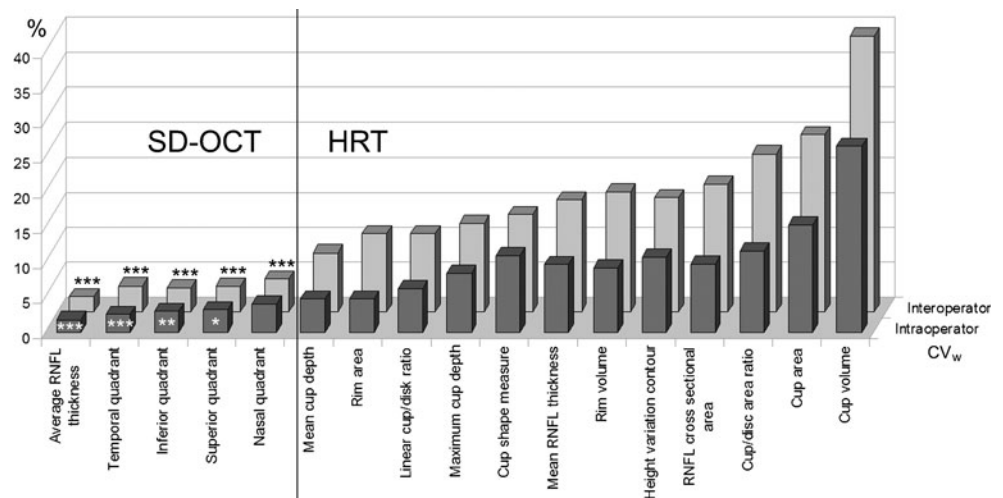
cup area (P<0.01) and maximum cup depth (P<0.05). The interoperator repeatability values of Cirrus HD-OCT parameters did not differ significantly from averaged intraoperator repeatability of both operators, with the exception of RNFL thickness in temporal quadrant (P<0.05). In contrast, the interoperator repeatability values for most HRT3 parameters except cup shape measure were substantially and significantly increased compared to averaged intraoperator repeatability of both operators (P<0.05-P<0.001). Intraoperator and interoperator within-subject coefficients of variation demonstrated the same relationships as repeatability values.

According to Table 2 and Fig. 1, the Cirrus HD-OCT parameters performed much better than HRT3 stereometric parameters. Mean (for both operators) intraoperator within-subject coefficient of variation of the best Cirrus HD-OCT parameter (average RNFL thickness) was 2.6–2.7 times lower than the best HRT3 parameters [mean cup depth and rim area (P<0.001)]. Mean intraoperator variability of RNFL thickness in quadrants (except nasal quadrant) was also significantly lower (P<0.05-P<0.001). The interoperator within-subject coefficients of variation both for average RNFL thickness and RNFL thickness in all quadrants were significantly lower than interoperator variability for the best HRT3 parameter [mean cup depth (P<0.001)].

The within-subject coefficients of variation of the average/mean RNFL thickness as assessed by both methods demonstrated a striking difference: for SD-OCT they were 5.4–7.3 times lower.

Among HRT3 parameters, mean cup depth, rim area, and linear cup/disk ratio were the least variable, while cup volume, cup area, and cup/disc area ratio were the most variable parameters.

Discussion

Quantitative evaluation of glaucoma progression by any method could not be performed without the knowledge of its measurement error. It is necessary to differentiate changes caused by the test variability from the real progression of the disease. Although other studies have been published on the repeatability and variability of SD-OCT and CSLO, the importance of the present study lies in the fact that it reports for the first time repeatability and variability data for two modern units — Cirrus HD-OCT and HRT3 — in a head-to-head comparison in eyes of the same patients with early glaucomatous visual field defect. This information is valuable for selecting the parameters which are potentially most useful for monitoring glaucoma progression. As such, we believe the results are worth reporting.

Repeatability (repeatability coefficient) is a very clear and practical characteristic of the measurement error. It shows the maximal possible difference between two measurements taken in the same subject (valid for 95% of pairs of observations) [17], and so permits differentiation of real change from measurement error. For example, according to our data the relatively small 5-μm change in average RNFL thickness (6 μm if we account for CI) could be a significant sign of early POAG progression when measured by the same operator on Cirrus HD-OCT. This is in close agreement with the results of other studies [8–11, 14]. Repeatability data for other Cirrus HD-OCT and HRT3 parameters shown in Table 1 could be used for evaluation of other changes seen in a patient with early POAG or suspected POAG.

The most important question raised in the present study was the comparison of measurement errors of Cirrus HD-OCT and HRT3. This question could not be answered by repeatability studies, because repeatability does not permit comparison of measurement errors of the parameters of different nature (scale) such as the Cirrus HD-OCT and HRT3 parameters. Even the repeatability of similar parameters assessed by both instruments — average / mean RNFL thickness — could not be compared directly due to the large difference in their scales (as shown by the mean values: 83.3 and 231 μm). The proper comparison of parameters of different nature (scale) could be performed by using coefficients of variation that are independent of the unit and scale of measurement [21]. So the within-subject coefficients of variation were used for comparison of measurement errors of Cirrus HD-OCT and HRT3.

The results of the present study demonstrate quite low variability of RNFL measurements by the new method of SD-OCT. Our data are in agreement with the majority of the recently published papers. The coefficient of variation of average RNFL thickness in healthy and glaucomatous persons varies mostly from 1.4 to 3.0% [8–11, 14]. Only a few authors have shown worse data (4.3–4.5% or more) [12, 13].

The studies on variability of CSLO (HRT) measurements are contradictory, though more recent studies demonstrate results similar to ours. Most authors consider rim area, mean and/or maximal cup depth among the least variable HRT parameters [1–3, 6, 7], while cup volume (and sometimes other cup measures [2]), mean RNFL thickness, and RNFL cross-sectional area are among the most variable parameters [1, 3, 6]. Some authors showed different results — low variability of cup volume [4, 5] and cup area [4], and high variability of rim area [5]. But they used the earliest model of HRT. The quantitative data of these studies are not consistent, and sometimes could not be compared to each other due to differences in statistical methods used. Nevertheless, some data are very similar to ours e.g.,

interoperator coefficients of variation in the study by Strouthides et al. [6].

It should be noted that variability of HRT3 parameters, much more than variability of Cirrus HD-OCT parameters, depended on several cases with outliers (data differing by more than 3 SD from the mean). Outliers in standard reference height (its values and difference between measurements) were the most important. Outliers in scan depth (difference between measurements), optic disc size, astigmatism, age and, for comparison of operators, in difference in disc area (contour line position) also played a role. In practice, a very experienced HRT operator could notice such outliers and perform additional control examinations. But even after exclusion of outliers, the variability of HRT3 parameters still could not approach that of the Cirrus HD-OCT parameters.

The most important result of the present study is the obvious superiority of Cirrus HD-OCT over HRT3 in terms of variability of measurements. That implies that Cirrus HD-OCT might be a better instrument for monitoring of glaucoma progression. But that would be a hasty conclusion, because a smaller measurement variability of any parameter is not equivalent to a higher sensitivity to detecting change: e.g., disk area has the smallest interoperator variability of HRT stereometric parameters, but it might not be an indicator of glaucoma progression at all. Still, if we compare parameters that show similar trends with glaucoma progression, the lower variability of measurements would indeed mean the higher sensitivity to detecting change. It could be said, for example, about a comparison of average RNFL thickness (Cirrus HD-OCT) and mean RNFL thickness (HRT3). Clinicians should be aware of these results when following glaucoma patients.

The results of this study must be interpreted by recognizing some limitations. First, we quantified only intrasession and not intersession (intervisit) measurement error. But we believe that intrasession variability better reflects the measurement error of the method itself, because it is not influenced by the changes of the patient's condition. Besides, there are conflicting data on the relationship between intrasession and intersession variability. Intrasession within-subject coefficients of variation have been found to be higher [6] or lower [10] than intervisit ones, or sometimes practically did not differ at all [8]. Secondly, we studied only patients with early POAG and POAG suspects, so the repeatability and variability data established in this study could not be extended to patients with other stages of the disease. Finally, we studied only stereometric HRT3 parameters and not the Moorfields Regression Analysis (MRA), Glaucoma Probability Score (GPS) or discriminant functions such as those of Frederick S. Mikelberg (FSM), Reinhard O.W. Burk (RB) etc. However, it should be noted that all these measures are primarily based on stereometric parameters, and the measurement error of such parameters plays a role in any subsequent analysis.

In conclusion, the present study showed that SD-OCT with Cirrus HD-OCT provides excellent intrasession repeatability (both intra- and interoperator) of the RNFL measurements, especially of the average RNFL thickness in patients with early glaucomatous visual field defect. The measurement error (expressed as within-subject coefficient of variation) of RNFL measurements by Cirrus HD-OCT is much lower than the error of HRT3 measurements of stereometric parameters of the ONH and RNFL.

References

1. Miglior S, Albé E, Guareschi M, Rossetti L, Orzalesi N (2002) Intraobserver and interobserver reproducibility in the evaluation of optic disc stereometric parameters by Heidelberg Retina Tomograph. Ophthalmology 109:1072–1077
2. Verdonck N, Zeyen T, Van Malderen L, Spileers W (2002) Short-term intra-individual variability in Heidelberg Retina Tomograph II. Bull Soc Belge Ophtalmol 286:51–57
3. Tan JC, Garway-Heath DF, Hitchings RA (2003) Variability across the optic nerve head in scanning laser tomography. Br J Ophthalmol 87:557–559
4. Funk J, Mueller H (2003) Comparison of long-term fluctuations: laser scanning tomography versus automated perimetry. Graefes Arch Clin Exp Ophthalmol 241:721–724
5. Watkins RJ, Broadway DC (2005) Intraobserver and interobserver reliability indices for drawing scanning laser ophthalmoscope optic disc contour lines with and without the aid of optic disc photographs. J Glaucoma 14:351–357
6. Strouthidis NG, White ET, Owen VM, Ho TA, Hammond CJ, Garway-Heath DF (2005) Factors affecting the test-retest variability of Heidelberg retina tomograph and Heidelberg retina tomograph II measurements. Br J Ophthalmol 89:1427–1432
7. Prata TS, Meira-Freitas D, Lima VC, Guedes LM, Magalhães FP, Paranhos Junior A (2010) Factors affecting the variability of the Heidelberg Retina Tomograph III measurements in newly diagnosed glaucoma patients. Arq Bras Oftalmol 73:354–357
8. Leung CK, Cheung CY, Weinreb RN, Qiu Q, Liu S, Li H, Xu G, Fan N, Huang L, Pang CP, Lam DS (2009) Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: a variability and diagnostic performance study. Ophthalmology 116:1257–1263
9. Vizzeri G, Weinreb RN, Gonzalez-Garcia AO, Bowd C, Medeiros FA, Sample PA, Zangwill LM (2009) Agreement between spectral-domain and time-domain OCT for measuring RNFL thickness. Br J Ophthalmol 93:775–781
10. Mwanza JC, Chang RT, Budenz DL, Durbin MK, Gendy MG, Shi W, Feuer WJ (2010) Reproducibility of peripapillary retinal nerve fiber layer thickness and optic nerve head parameters measured with Cirrus HD-OCT in glaucomatous eyes. Invest Ophthalmol Vis Sci 51:5724–5730
11. Savini G, Carbonelli M, Parisi V, Barboni P (2010) Effect of pupil dilation on retinal nerve fibre layer thickness measurements and their repeatability with Cirrus HD-OCT. Eye (Lond) 24:1503–1508
12. Cremasco F, Massa G, Gonçalves Vidotti V, de Carvalho P, Lupinacci A, Costa VP (2011) Intrasession, intersession, and

interexaminer variabilities of retinal nerve fiber layer measurements with spectral-domain OCT. Eur J Ophthalmol 21:264–270

13. Labiris G, Giarmoukakis A, Katsanos A, Gkika MG, Fanariotis M, Pavlidou E, Kozobolis VP (2010) Reproducibility of OCT/SLO measurements in healthy eyes. Eur J Ophthalmol 20:552–558

14. González-García AO, Vizzeri G, Bowd C, Medeiros FA, Zangwill LM, Weinreb RN (2009) Reproducibility of RTVue retinal nerve fiber layer thickness and optic disc measurements and agreement with Stratus optical coherence tomography measurements. Am J Ophthalmol 147:1067–1074

15. Mills RP, Budenz DL, Lee PP, Noecker RJ, Walt JG, Siegartel LR, Evans SJ, Doyle JJ (2006) Categorizing the stage of glaucoma from pre-diagnosis to end-stage disease. Am J Ophthalmol 141:24–30

16. Bland JM (2010) How can I decide the sample size for a repeatability study? Available at: http://www-users.york.ac.uk/~mb55/meas/sizerep.htm. Accessed January 24, 2011

17. Bland JM, Altman DG (1996) Statistics notes: measurement error. BMJ 313:744

18. Bland JM (2004) What is the standard error of the within-subject standard deviation, S_w? Available at: http://www-users.york.ac.uk/~mb55/meas/seofsw.htm. Accessed January 24, 2011

19. Bland JM (2006) How should I calculate a within-subject coefficient of variation? Available at: http://www-users.york.ac.uk/~mb55/meas/cv.htm. Accessed January 24, 2011

20. Zar JH (2010) Biostatistical analysis, 5th edn. Pearson Prentice-Hall, Upper Saddle River, NJ, pp 160–161

21. Abdi H (2010) Coefficient of variation. In: Salkind N, Dougherty DM, Frey B (eds) Encyclopedia of research design. Sage, Thousand Oaks, CA, pp 169–171