Factors affecting the variability of the Heidelberg Retina Tomograph III measurements in newly diagnosed glaucoma patients

Fatores associados com a variabilidade das medidas do Heidelberg Retina Tomograph III em pacientes com glaucoma recém-diagnosticados

TIAGO SANTOS PRATA1, DANIEL MEIRA-FREITAS1, VERÔNICA CASTRO LIMA1, LIA MANIS GUIDE2, FERNANDA PEDREIRA MAGALHÃES1, AUGUSTO PARANHOS JUNIOR1

ABSTRACT

Purpose: To determine factors associated with the test-retest variability of optic nerve head (ONH) topography measurements with confocal scanning laser ophthalmoscopy (CSLO) in newly diagnosed glaucomatous patients.

Methods: Consecutive patients with newly diagnosed primary open-angle glaucoma were prospectively enrolled. Patients presenting with any ocular disease other than glaucoma were excluded. All patients underwent CSLO using the Heidelberg Retina Tomograph III (HRT-III) in one randomly selected eye (three consecutive scans; performed by the same examiner). For each Heidelberg Retina Tomograph III parameter, repeatability was assessed using within subject standard deviation (Sv) and coefficient of variation (Cv), repeatability coefficient (RC) and intraclass correlation coefficient (ICC). Scatter plots and regression lines were constructed to identify which factors influenced test-retest measurement variability.

Results: A total of 32 patients were included (mean age, 65.4 ± 13.8 years). Most patients were female (65%) and white (50%). Among all Heidelberg Retina Tomograph III parameters evaluated, rim area and mean cup depth had the best measurement repeatability. Vertical cup-to-disc ratio (CDR, as determined by optic disc stereophotograph examination) was significantly associated (R²=0.21, p<0.01) with test-retest measurement variability. Eyes with larger CDR showed less variable measurements. Other factors, including age, disc area, central corneal thickness and intraocular pressure were not significant (p>0.14).

Conclusion: Heidelberg Retina Tomograph III showed good test-retest repeatability for all ONH topographic measurements, mainly for rim area and mean cup depth. Test-retest repeatability seemed to improve with increasing CDR. These findings suggest that HRT-III topographic measurements should be cautiously interpreted when evaluating longitudinally glaucoma patients with early structural damage (small CDR).

Keywords: Glaucoma/diagnosis; Glaucoma, open-angle/diagnosis; Diagnostic techniques, ophthalmological; Optic disk; Optic nerve diseases/diagnosis; Tomography/methods; Ophthalmoscopy; Image interpretation, computer-assisted; Visual acuity; Visual fields

RESUMO

Objetivo: Determinar os fatores associados à variabilidade (teste-reteste) das medidas topográficas da cabeça do nervo óptico (CNO) utilizando o oftalmoscopia confocal de varredura a laser (CSLO) em pacientes com glaucoma recém-diagnosticados.

Métodos: Neste estudo, pacientes com glaucoma primário de ângulo aberto recém-diagnosticados foram prospectivamente incluídos. Aqueles que apresentassem outras doenças oculares (exceto glaucoma) foram excluídos. Todos os pacientes incluídos no estudo foram submetidos à CSLO usando o aparelho Heidelberg Retina Tomograph III (HRT-III) em um olho aleatoriamente selecionado (três exames consecutivos realizados pelo mesmo examinador). Para cada parâmetro do Heidelberg Retina Tomograph III, a repetibilidade foi avaliada através dos seguintes indicadores: desvio padrão (DP) e coeficiente de variação (CV) individual, coeficiente de repetibilidade (CR) e coeficiente de correlação intraclasse (CCI). Diagramas de dispersão e linhas de regressão foram construídos para identificar quais fatores pudessem influenciar a variabilidade das medidas.

Resultados: Trinta e dois pacientes foram incluídos no estudo (idade média, 65,4 ± 13,8 anos). A maior parte era composta por mulheres (65%) e pacientes brancos (50%). Dentre os parâmetros de Heidelberg Retina Tomograph III avaliados, a área da rima e a profundidade média da escavação apresentaram os melhores valores de repetibilidade. A relação escavação/disco (E/D) vertical (baseada na análise de estereofotografia do disco óptico), foi significativamente associada (R²=0,21, p<0,01) com a variabilidade teste-reteste. Pacientes com relação E/D maiores apresentaram medidas menos variáveis. Outros fatores como idade, área do disco, espessura corneal central e pressão intraocular não foram significativos (p>0,14).


Descritores: Glaucoma/diagnóstico; Glaucoma de ângulo fechado/diagnóstico; Técnicas de diagnóstico oftalmológico; Disco óptico; Doenças do nervo óptico/diagnóstico; Tomografia/métodos; Oftalmoscopia; Interpretação de imagem assistida por computador; Acuidade visual; Campos visuais

INTRODUCTION

Glaucoma is a progressive optic neuropathy that causes morphologic changes of the optic nerve head (ONH) and retinal nerve fiber layer (RNFL). Detection and monitoring of glaucoma patients are based on identification of structural and functional changes. Therefore, clinical assessment of the ONH is of utmost importance in the management of the disease. However, it greatly depends on the physician’s clinical ability and is limited by subjective analysis and intra- and inter-observer variability[1]. In order to improve detection of early...
damage and its progression over time, different technologies for objective and quantitative measurement of structural changes have emerged\textsuperscript{6-8}. The Heidelberg retina tomograph (HRT, Heidelberg Engineering, Heidelberg, Germany) is a well-established confocal scanning laser ophthalmoscope (CSLO) that provides automated evaluation of the ONH topography and RNFL\textsuperscript{9,10}. The HRT has been widely used to objective diagnose and monitor glaucomatous optic neuropathy (GON). It assesses structural changes of the ONH by analysis of a series of images acquired over time. Although it provides useful information that helps the physician to distinguish between progressing and non-progressing patients, Heidelberg Retina Tomograph (HRT) images are not identically scaled versions of the true ONH and are affected by measurement variability during image acquisition (intra-scan noise) and between visits during follow-up (inter-scan noise)\textsuperscript{11,12}. Although previous studies have found good reproducibility for HRT topographic measurements, most of these studies used previous versions of the HRT and only a few assessed potential sources of measurement variability\textsuperscript{13,14}. It is well known that a device reproducibility is critical for following glaucoma patients over time. Therefore, the purpose of this study was to identify possible factors associated with the test-retest variability of ONH topography parameters with Heidelberg Retina Tomograph III (HRT-III) and to establish which parameters provide the most repeatable measurements, in patients with newly diagnosed primary open-angle glaucoma (POAG) with various degrees of glaucomatous damage.

METHODS

This cross-sectional study adhered to the tenets of the Declaration of Helsinki and was approved by the Ethics Committee of the Federal University of São Paulo. In addition, written informed consent was obtained from all subjects.

PATIENTS

Consecutive patients with newly diagnosed POAG from the general clinic of the Ophthalmology Department of the Federal University of São Paulo were prospectively enrolled. Patients who met our inclusion criteria were promptly directed to the glaucoma clinic, where a complete ophthalmological examination was performed. Exclusion criteria were previous ocular surgery or trauma, secondary glaucoma, spherical equivalent \(> \pm 4.0\) D, history of using oral or topical steroids, and any ocular disease other than glaucoma (including moderate or advanced cataract). POAG was defined as characteristic GON and visual field loss, with intraocular pressure (IOP) >21 mmHg on two separate occasions and a widely open-angle.

Typical GON was defined as a vertical cup-to-disc ratio (CDR) greater than 0.5, asymmetry of the cup-to-disc ratio \(\geq 0.2\) between eyes, presence of localized RNFL defects and/or neuroretinal rim defects in the absence of any other abnormalities that could explain such findings. Assessment of CDR values was based on optic disc stereophotographs evaluation by one single experienced examiner (TSP) masked to patient's clinical findings. A glaucomatous visual field defect in the standard automated perimetry (Humphrey SITA - Standard 24-2, Carl Zeiss Meditec, Dublin, CA) was defined as three or more points in clusters with a probability of \(< 5\%\) (excluding those on the edge of the field or directly above and below the blind spot) on the pattern deviation plot, a pattern standard deviation index with a probability of \(< 5\%\), or a glaucoma hemifield test with results outside the normal limits.

PROCEDURES

Baseline data assessed were age, gender, self-described race, central cornea thickness (CCT) and intraocular pressure (IOP) (Goldmann applanation tonometry). All patients underwent confocal scanning laser ophthalmoscopy using the Heidelberg Retina Tomograph III (HRT; Heidelberg Engineering, Heidelberg, Germany). A series of three scans was performed by the same examiner (LMD) in one randomly selected eye. Patients were refocused before the first HRT exam, and whenever the cylinder was \(\geq 1.00\) diopter, correction for corneal astigmatism was performed. Quality control of the scans was set to less than 30 \(\mu m\) of standard deviation. Two patients were excluded from the analysis due to poor quality scans. The examiner was masked for patient's clinical data. After the final exam, glaucoma treatment was initiated, ancillary exams were scheduled and patients were followed regularly in the glaucoma clinic.

STATISTICAL ANALYSIS

Descriptive analysis was used to present demographic and clinical data. The following HRT parameters were chosen for analysis: rim area and volume, cup area and volume, and mean and maximum cup depth. Within subject standard deviation (Sw) and coefficient of variation (CVw= Sw/mean of all repeated measurements) were used to examine the repeatability of each HRT parameter. Repeatability was also assessed by estimating the repeatability coefficient (RC= sqrt(2)\times 1.96\times Sw\textsuperscript{2}). Intraclass correlation coefficient (ICC; testing for absolute agreement) was used to estimate the reliability of the parameters generated. Scatter plots and regression lines were constructed to identify which factors influenced test-retest measurement variability with significant associations assumed at \(p<0.05\). The factors evaluated were age, intraocular pressure (IOP), CCT, disc area and vertical cup-to-disc ratio. As rim area represents a more clinically meaningful measure and was used in previous studies to assess factors associated with test-retest measurement variability\textsuperscript{9}, we used this HRT parameter (expressed by the rim area CW) as the dependent variable in our analysis. All statistical analyses were performed using Medcalc version 7.4.2.0 (Medcalc Software, Mariakerke, Belgium).

RESULTS

A total of 32 patients were included (mean age, 65.4 ± 13.8 years). Most patients were female (65%) and white (50%). Baseline characteristics of study patients are summarized in table 1. All HRT parameters evaluated showed good repeatability results as shown in table 2 (Sw range, 0.01 - 0.03; CVw range, 0.02 - 0.07; RC range, 0.08 - 0.22; ICC range, 0.94 - 0.99). Judged by the CVw, rim area (0.02) and mean cup depth (0.02) had the best measurement repeatability. Judged by the ICC, rim area (0.99), cup area (0.99) and mean cup depth (0.99) were the most reliable parameters.

Vertical CDR (as determined by stereophotograph) was the only factor significantly associated with test-retest variability (expressed by the rim area CW; R\textsuperscript{2}=0.21, p<0.01). Figure 1 show a scatter plot of rim area CVw against vertical CDR. Table 3 summarizes the relationship between all sources of variability evaluated and rim area CVw. No significant associations were observed for age, disc area, CCT and IOP (p\textsuperscript{2}>0.31).

DISCUSSION

The high-technology devices introduced into glaucoma practice over the last decade are intended to detect the earliest structural and functional changes and to support ophthalmologists in clinical decision-making. One of the most important requisites of such devices is the reproducibility of their findings, as this can affect their diagnostic ability. Besi-
des, it is essential to know whether changes observed during follow-up examination are related to the test variability or indicate a real progression of the disease. To the best of our knowledge, this is the first study to report a significant association between the relative size of the cup and test-retest repeatability of HRT-III topographic measurements.

Based on different indices, good reproducibility results were found for all HRT-III topographic parameters, with the mean ICC ranging from 0.94 to 0.99 for cup- and rim-related parameters. Considering published studies with previous HRT versions, our results are comparable to most of the data reported on intraobserver-intrasession repeatability (13-18). Although the reproducibility of stereometric parameters from previous HRT versions has been evaluated in several studies, only a few of them assessed possible sources of variability (9,19).

Considering all stereometric parameters evaluated in our study, rim area and cup depth were the most repeatable parameters. These findings are in agreement with previous publications (8,19). A study that evaluated the reproducibility of different stereometric parameters using HRT-II reported that mean cup depth and cup area (which is highly correlated with rim area measurements) were the least variable parameters (8). Other studies using previous versions of the HRT also reported these parameters as the most repeatable (9,19). Rim area, which was found to be a highly repeatable measurement in our study, is an important parameter used by the Moorfields Regression Analysis (in a cross-sectional manner) for glaucoma classification. It has also been previously chosen as a key parameter (dependent variable) to investigate possible sources of measurement variability for the HRT-II (9). As rim area comprises the retinal ganglion cell axons, it is directly related to structural damage in glaucoma and therefore has been discussed as a more clinically meaningful parameter for physicians (9,19). Finally, it has been effectively used to discriminate between normal, glaucomatous, and ocular hypertensive subjects (20-22). Therefore, based on its high reproducibility, rim area is a potential parameter for assessment of disease progression.

Among different sources of variability evaluated in our study, we found that a larger CDR was significantly associated with a decreased HRT measurement variability, as determined by the rim area coefficient of variation. Other factors were not significant in this analysis, including age, disc area, CCT and IOP. Considering the relative size of the cup as a measurement of disease severity, our findings suggest that patients with advanced structural glaucomatous damage (for example, a CDR of 0.95) would have more repeatable measurements than those with mild damage (for example, a CDR of 0.5). These findings are relevant regarding the management of glaucomatous patients overtime, as we use HRT to follow glaucoma patients with early to moderate structural damage more often than we use for those with very advanced damage (total cupping for instance). There is scant and controversial information in the literature about the influence of glaucomatous damage on HRT measurements variability. For instance, some authors (23), evaluating glaucomatous patients and healthy subjects, found no significant association between linear cup-to-disc ratio (as determined by the HRT) and test-retest variability. Strouthidis et al. (9), found no association between baseline rim area (also considered a measurement of structural damage) and test-retest variability. In that study, cup-to-disc ratio was not evaluated directly as a possible source of variability. Finally, DeLeon et al. (24) found an increased test-retest variability associated with advanced glaucomatous damage (based on visual field loss) for some HRT parameters, but no association for others, even after taking into account the effects from age, visual acuity and lens status.

### Table 1. Baseline characteristics of study patients

<table>
<thead>
<tr>
<th>Variable</th>
<th>Patients (n=32)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>65 ± 13.8</td>
</tr>
<tr>
<td>Gender (female/male)</td>
<td>21/11</td>
</tr>
<tr>
<td>Race (white/african descent/mixed)</td>
<td>16/9/7</td>
</tr>
<tr>
<td>Baseline intraocular pressure (mmHg)</td>
<td>27.4 ± 7.4</td>
</tr>
<tr>
<td>Central cornea thickness (μm)</td>
<td>542.1 ± 37.5</td>
</tr>
<tr>
<td>Disc area (mm²)</td>
<td>2.4 ± 0.4</td>
</tr>
<tr>
<td>Linear cup-to-disc ratio (HRT)</td>
<td>0.64 ± 0.15</td>
</tr>
<tr>
<td>Vertical cup-to-disc ratio (based on stereophotographs)</td>
<td>0.72 ± 0.12</td>
</tr>
<tr>
<td>Mean cup depth (mm)</td>
<td>0.32 ± 0.11</td>
</tr>
<tr>
<td>Visual field mean deviation (dB)</td>
<td>-5.3 ± 5.5</td>
</tr>
</tbody>
</table>

HRT = Heidelberg Retina Tomograph

* = data are given as mean ± standard deviation whenever indicated

### Table 2. Reproducibility indices for each topographic parameter

<table>
<thead>
<tr>
<th>HRT parameters</th>
<th>Sw (mean)</th>
<th>CVw (mean)</th>
<th>RC (mean)</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rim area (mm²)</td>
<td>0.03</td>
<td>0.02</td>
<td>0.22</td>
<td>0.99 (0.98-0.99)</td>
</tr>
<tr>
<td>Rim volume (mm³)</td>
<td>0.02</td>
<td>0.07</td>
<td>0.17</td>
<td>0.96 (0.93-0.98)</td>
</tr>
<tr>
<td>Cup area (mm²)</td>
<td>0.03</td>
<td>0.04</td>
<td>0.21</td>
<td>0.99 (0.98-0.99)</td>
</tr>
<tr>
<td>Cup volume (mm³)</td>
<td>0.02</td>
<td>0.06</td>
<td>0.15</td>
<td>0.98 (0.97-0.99)</td>
</tr>
<tr>
<td>Mean CD (mm)</td>
<td>0.01</td>
<td>0.02</td>
<td>0.08</td>
<td>0.99 (0.98-0.99)</td>
</tr>
<tr>
<td>Maximum CD (mm)</td>
<td>0.03</td>
<td>0.04</td>
<td>0.19</td>
<td>0.94 (0.90-0.97)</td>
</tr>
</tbody>
</table>

HRT = Heidelberg Retina Tomograph; CD = cup depth; Sw = within subject standard deviation; CVw = within subject coefficient of variation; RC = repeatability coefficient; ICC = intraclass correlation coefficient (95% confidence interval)

### Figure 1. Scatter plot of vertical cup-to-disc ratio (as determined by funduscopy) against rim area coefficient of variation. The regression line is also shown (R²=0.21, p<0.01).
Different factors have been previously identified as potential sources of HRT measurement variability. These include patient/scanner misalignment\textsuperscript{1,2}, inter-test differences in disc area contour line drawing and reference height\textsuperscript{3,5,21}, image quality (influenced by lens opacity and cylindrical refractive error)\textsuperscript{19}, age and degree of astigmatism\textsuperscript{16}. In the present study, the disc area contour was drawn by one single observer (LMG) during the first scan and then imported to the subsequent images. By keeping a constant disc area contour, this possible source of variability was minimized and therefore not explored. Since we did not include patients with significant media opacity or high refractive error (spherical equivalent $< +4.0 \, \text{D}$), and scan quality control was set to $< 30 \, \mu \text{m}$ of standard deviation, the presence of cataract, astigmatism and poor image quality were also mitigated and not investigated as possible sources of measurement variability. Ultimately, the fact that our key factor associated with test-retest variability (vertical CDR) was assessed by one single examiner instead of two (plus a third one for adjudication) should be considered while interpreting our results.

**CONCLUSION**

In summary, HRT-III showed good test-retest repeatability for all ONH topographic measurements. Rim area and mean cup depth were the least variable parameters. Test-retest repeatability seemed to improve with increasing CDR. These findings suggest that HRT-III topographic measurements should be cautiously interpreted when evaluating longitudinal glaucoma patients with early structural damage (small CDR).

**REFERENCES**