Published in final edited form as:

Ophthalmology. 2008 April ; 115(4): 673–677. doi:10.1016/j.ophtha.2007.05.045.

Comparison of Parameters from Heidelberg Retina Tomographs 2 and 3

Michelle L. Gabriele, BS, Gadi Wollstein, MD, Richard A. Bilonick, PhD, Zvia Burgansky-Eliash, MD, Hiroshi Ishikawa, MD, Larry E. Kagemann, MS, and Joel S. Schuman, MD

From University of Pittsburgh Medical Center Eye Center, Eye and Ear Institute, Ophthalmology and Visual Science Research Center, Department of Ophthalmology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.

Abstract

Purpose—To compare stereometric parameters and classification results from the Heidelberg Retina Tomograph version 2 (HRT2); HRT3; and HRT3 Glaucoma Probability Score (GPS), an automated method of obtaining optic nerve head analysis without the need for manual definition of disc margin.

Design—Retrospective cross-sectional study.

Participants—Five hundred four eyes from 281 consecutive subjects (glaucoma, glaucoma suspect, and healthy) evaluated in a glaucoma clinic.

Methods—All participants had HRT2 scanning of the optic nerve head. Inclusion criteria were scans with good centration and focus, even illumination, an overall quality score by HRT3 of acceptable or better, and standard deviation < 50 μ m. A Bland–Altman analysis was used for the comparison of HRT2 and HRT3. From these results, calibration equations were determined to permit conversion of the measurements between devices. The agreement between HRT2 and HRT3 Moorfields regression analysis (MRA) and HRT3 GPS classification methods was measured using κ statistics.

Main Outcome Measures—Heidelberg Retina Tomograph version 2 and HRT3 stereometric parameters, MRA, and global GPS.

Results—There was a statistically significant difference between HRT2 and HRT3 global disc area, rim area, cup area, rim volume, cup volume, height variation contour, and retinal nerve fiber layer cross-sectional area stereometric parameters. All of those parameters were smaller using HRT3, due to a manufacturer-reported horizontal scaling error of 4% in HRT2 that was corrected in HRT3. κ s for agreement were 0.60 between classifications (within normal limits, borderline, and outside normal limits) of MRA by HRT2 and HRT3 and 0.47 between HRT3 MRA and GPS.

Conclusions—The HRT3 generally provided smaller stereometric disc measurements than HRT2. There was no clear conversion between HRT3 and GPS parameters, as the 2 methods for measuring the stereometric parameters differ.

The confocal scanning laser ophthalmoscope uses a rapid-scanning diode laser light source to obtain quantitative measurements of the optic nerve head (ONH) region. Structural

changes in the appearance of the ONH and surrounding tissues are indicative of the presence and progression of diseases such as glaucoma.^{1–3} The commercially available Heidelberg Retina Tomograph (HRT; Heidelberg Engineering GmbH, Heidelberg, Germany) is a confocal scanning laser ophthalmoscope that creates 3-dimensional topography images and then derives several structural measurement parameters representing the ONH globally and in predefined segments. The recently released HRT version (HRT3) incorporates a larger normative database and thus permits an analysis that is more specific to ethnic background, age, and disc size than the commonly used HRT2. Both versions of HRT also include Moorfields regression analysis (MRA), which has been shown to be highly capable of discriminating between healthy and glaucomatous eyes.^{4,5}

One limitation of HRT technology has been the requirement that the operator manually define a contour line marking the inner border of the ONH margin as defined by the scleral ring. Many of the quantitative measurements are derived from the contour line placement, thereby inducing measurement variability.^{6,7} The HRT3 provides an alternative option of using an automated ONH analysis that does not require operator placement of a contour line. Stereometric parameters are instead derived by applying parametric mathematical modeling of the ONH shape, as suggested by Swindale et al.⁸ A 3-dimensional model is fit to the ONH, and the probability that an eye is glaucomatous is determined using relevance vector machine learning classifiers (Glaucoma Probability Score [GPS]).

In this study, 3 different analyses were applied to a single dataset: (1) HRT2 with operator-defined contour, (2) HRT3 with operator-defined contour, and (3) HRT3 with GPS automated disc margin identification. The purpose of the study was to compare corresponding HRT2, HRT3, and HRT3 GPS values.

Materials and Methods

Subjects

The study was approved by the University of Pittsburgh Institutional Review Board/Ethics Committee and adhered to the Declaration of Helsinki and Health Insurance Portability and Accountability Act. The data for this study were collected retrospectively during the Pittsburgh Imaging and Novel Diagnostics for Glaucoma Study, a prospective longitudinal study carried out in the glaucoma service at the University of Pittsburgh Medical Center Eye Center. All subjects attending the service who were willing to participate in the study and qualified according to the criteria below were included. Subjects (healthy, glaucoma suspect, and glaucoma) were at least 18 years old with best-corrected visual acuity of 20/60 or better, refractive error between +3.00 and –6.00 diopters. Exclusion criteria were a history of diabetes and eyes with signs of posterior pole pathology other than those attributed to glaucoma or significant media opacity in which the fundus was not visible. Informed consent was obtained from all subjects.

Heidelberg Retina Tomograph

All participants in this study had HRT2 scanning of the ONH. Three images were acquired automatically after initial positioning by an experienced operator. Qualified HRT scans were well centered and well focused, with even illumination; had an overall quality score as defined by HRT3 of acceptable or better; and had an interscan standard deviation (SD) of <50 μm . The ONH contour line was then drawn by the same operator, with the ONH margin defined as the inner border of Elschning's ring. The contour line was drawn in HRT2 software, and the data were exported to HRT3 and processed without altering the location of the contour line. The HRT3 software also performed an automated GPS analysis on all data. Thus, global stereometric parameters for HRT2, HRT3, and HRT3 with GPS were acquired

from the HRT2 scans. Moorfields regression analysis categorical classification (within normal limits, borderline, or outside normal limits) was recoded from HRT2 and HRT3. Glaucoma Probability Score analysis provided by HRT3 includes the following parameters: horizontal and vertical retinal nerve fiber layer (RNFL) curvature, cup size, rim steepness, and global GPS. We compared GPS cup size with 2 HRT3 conventional stereometric parameters (disc area and cup area).

Analysis

The R language and environment for statistical computing and graphics was used in this analysis.⁹ Bland–Altman analyses were used to assess agreement between HRT2 and HRT3 as well as GPS and HRT3 global stereometric parameters.¹⁰ To account for clustering within subjects, a mixed-effects model was used. To evaluate the agreement using the Bland and Altman method, the 2 methods should have similar imprecision. The Grubbs method was used to estimate each device's imprecision.¹¹ To assess the nature of the bias, the difference between the 2 methods was plotted and regressed against the average of the 2 methods across subjects for each parameter. For each parameter, the bias was shown to be either nonconstant but linear or roughly constant with the level of the parameter. Calibration equations relating the devices were generated from the regression equations relating the differences to the averages, permitting the conversion between devices. Constant bias could then be assessed for parameters that did not demonstrate a statistically significant nonconstant bias using a paired *t* test.

κ statistics were used to analyze the agreement between HRT2 and HRT3 MRA classification as well as HRT3 MRA and GPS classification. Glaucoma Probability Score data were converted to ordinal classification using the manufacturer-defined cutoffs: GPS score of 0% to 27%, within normal limits; 28% to 64%, borderline; and 65% to 100%, outside normal limits.

Results

A total of 281 subjects were recruited for this study (504 eyes [252 left and 252 right, 301 female and 203 male]). The subjects ranged in age from 18 to 93 years (mean, 55.3 ± 16.1). Eighty-two percent of the subjects (231; 418 eyes) were Caucasian, 16% (44; 75 eyes) were African American, and the remaining subjects were Indian.

The mean stereometric results for HRT3 appear in Table 1. The difference between HRT2 and HRT3 measurements of global stereometric parameters as a function of the corresponding average (i.e., a nonconstant bias) was statistically significant for disc area, rim area, cup area, rim volume, cup volume, height variation contour, and RNFL cross-sectional area (Table 1). This table also shows calibration equations for converting between HRT2 and HRT3 global stereometric parameters. Most of the parameters that showed a statistically significant difference were 3% to 4% larger in HRT2 than in HRT3, as indicated by the slope of calibration line equation (Fig 1). Exceptions were height variation contour and RNFL cross-sectional area. Table 1 also shows the estimated bias for parameters that had a constant bias (parameters with $P > 0.05$ for difference as a function of average). Mean RNFL thickness and height variation contour showed a statistically significant constant bias between HRT2 and HRT3 measurements.

For the HRT3 stereometric parameters and GPS comparison, the imprecisions in the 2 methods were not similar, and thus, a direct comparison between methods was not possible.

Good agreement was observed between HRT2 MRA and HRT3 MRA ($\kappa = 0.60$; Table 2), whereas moderate agreement was observed between HRT3 MRA and GPS classification (κ

= 0.47; Table 3). Raw agreement, indicating the percentage of classifications that showed exact agreement between HRT2 MRA and HRT3 MRA, was 74%. However, HRT3 MRA tended to report more advanced structural damage than HRT2 MRA. Seventy-nine eyes (29%) labeled as within normal limits by HRT2 MRA were classified as borderline by HRT3 MRA, and 58% of eyes classified as borderline by HRT2 MRA were classified as outside normal limits by HRT3 MRA. The raw agreement between HRT3 MRA and GPS classification was 64%, with no overt pattern in the difference between the methods' classifications.

Discussion

This study evaluated the differences between measurements obtained by HRT2 and HRT3. Due to a manufacturer-reported 4% horizontal scaling error in HRT2 that was corrected in HRT3, we expected to see area and volume stereometric parameter differences between the 2 methods. The magnification error was not expected to affect any parameters that depended only on z-direction measurements.

An approximate 4% difference was evident in all area and volumetric measurements, as expected, with HRT2 larger than HRT3 (Fig 1). In addition, height variation contour and RNFL cross-sectional area were also affected by the scaling correction. Differences may have been induced by a rounding error that resulted in slight alterations in the precise location of the contour line after correcting for the HRT2 magnification error. This in turn might affect the height variation contour, which is defined as the height difference between the most elevated and most depressed points of the contour line, and a shift of even one pixel may alter this discrete measurement. Retinal nerve fiber layer cross-sectional area is dependent on the length of the contour line and, thus, is affected by the horizontal scaling correction. As expected, stereometric parameters that represent ratios (cup to disc) were not affected by the scaling. Cup shape measure was also unaltered, as it represents the cup slope in the z-axis direction. Among the parameters with a constant bias, a significant difference was noted for mean RNFL thickness and global reference height. The source for this is unknown but does not hold any clinical implication, as the difference was <1% of the mean value.

Due to the large difference in the imprecisions for HRT3 stereometric parameters and GPS, it was impossible to establish reliable conversion formulas. This reflects a fundamental difference in the methodology for obtaining these measurements, which precludes a direct comparison.

The agreement between HRT2 and HRT3 MRA classifications was better than the agreement between HRT3 and GPS. This is likely because HRT2 and HRT3 use the same classification system, whereas GPS classification is derived from a different method. Although the comparison between HRT2 and HRT3 MRA showed a consistent trend of more advanced damage with HRT3, the comparison between HRT3 and GPS did not show a consistent trend. This is likely because the methods are sensitive to different ONH features. The clinical implication of this result is not yet known.

In conclusion, calibration equations have been presented to allow for the conversion between HRT2 and HRT3 parameters. Although stereometric parameters from HRT2 can be converted into HRT3 and vice versa, it was not possible to present a simple calibration for GPS and HRT3 parameters.

Acknowledgments

Supported in part by the National Institutes of Health, Bethesda, Maryland (grant nos. RO1-EY013178-6, P30-EY008098); Eye and Ear Foundation, Pittsburgh, Pennsylvania; and Research to Prevent Blindness, Inc., New York, New York (unrestricted grant).

References

1. Iester M, Mikelberg FS, Courtright P, Drance SM. Correlation between the visual field indices and Heidelberg retina tomograph parameters. *J Glaucoma* 1997;6:78–82. [PubMed: 9098814]
2. Eid TM, Spaeth GL, Katz LJ, et al. Quantitative estimation of retinal nerve fiber layer height in glaucoma and the relationship with optic nerve head topography and visual field. *J Glaucoma* 1997;6:221–230. [PubMed: 9264301]
3. Chauhan BC, McCormick TA, Nicolela MT, LeBlanc RP. Optic disc and visual field changes in a prospective longitudinal study of patients with glaucoma: comparison of scanning laser tomography with conventional perimetry and optic disc photography. *Arch Ophthalmol* 2001;119:1492–1499. [PubMed: 11594950]
4. Wollstein G, Garway-Heath DF, Hitchings RA. Identification of early glaucoma cases with the scanning laser ophthalmoscope. *Ophthalmology* 1998;105:1557–1563. [PubMed: 9709774]
5. Burgansky-Eliash Z, Wollstein G, Bilonick RA, et al. Glaucoma detection with the Heidelberg retina tomograph 3. *Ophthalmology* 2007;114:466–471. [PubMed: 17141321]
6. Garway-Heath DF, Poinoosawmy D, Wollstein G, et al. Inter- and intraobserver variation in the analysis of optic disc images: comparison of the Heidelberg retina tomograph and computer assisted planimetry. *Br J Ophthalmol* 1999;83:664–669. [PubMed: 10340972]
7. Miglior S, Albe E, Guareschi M, et al. Intraobserver and interobserver reproducibility in the evaluation of optic disc stereometric parameters by Heidelberg retina tomograph. *Ophthalmology* 2002;109:1072–1077. [PubMed: 12045046]
8. Swindale NV, Stjepanovic G, Chin A, Mikelberg FS. Automated analysis of normal and glaucomatous optic nerve head topography images. *Invest Ophthalmol Vis Sci* 2000;41:1730–1742. [PubMed: 10845593]
9. R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing; 2006.
10. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–310. [PubMed: 2868172]
11. Jaech, JL. *Statistical Analysis of Measurement Errors*. New York: Wiley; 1985. p. 49-103.

Figure 1. Bland–Altman plot of rim area measurements obtained with the Heidelberg Retinal Tomograph Version 2 (HRT2) and HRT3. A nonconstant bias is evident as the difference between measurements increases as a function of the average measurement (solid line). Dashed lines, 95% confidence intervals.

Mean Stereometric Results for the Heidelberg Retina Tomograph Version 3 (HRT3) and Equations to Convert Global Stereometric Parameters from HRT3 to HRT2

Table 1

Parameter	HRT3 Mean Value	Conversion	P Value	Mean Difference	P Value*
Disc area (mm ²)	1.903±0.528	HRT2 = 1.039 × HRT3	<0.0001		
Rim area (mm ²)	1.220±0.375	HRT2 = 0.004 + 1.038 × HRT3	<0.0001		
Cup area (mm ²)	0.683±0.508	HRT2 = -0.001 + 1.039 × HRT3	<0.0001		
Rim volume (mm ³)	0.298±0.146	HRT2 = 0.002 + 1.034 × HRT3	<0.0001		
Cup volume (mm ³)	0.168±0.191	HRT2 = -0.001 + 1.040 × HRT3	<0.0001		
Cup-to-disc area ratio	0.336±0.187	HRT2 = 0.999 × HRT3	0.46	-0.0005	0.13
Linear cup-to-disc area ratio	0.552±0.177	HRT2 = 0.998 × HRT3	0.39	-0.0005	0.14
Cup shape measure	-0.150±0.097	HRT2 = 0.999 × HRT3	0.66	-0.0002	0.50
Mean cup depth (mm)	0.238±0.100	HRT2 = 0.998 × HRT3	0.23	<0.0001	0.99
Maximum cup depth (mm)	0.609±0.198	HRT2 = 0.001 + 0.999 × HRT3	0.70	0.0001	0.84
Height variation contour (mm)	0.368±0.112	HRT2 = 0.004 + 0.990 × HRT3	0.04		
Mean RNFL thickness (mm)	0.218±0.079	HRT2 = 0.001 + 1.003 × HRT3	0.46	0.0014	<0.0001
RNFL cross-sectional area (mm ²)	1.051±0.392	HRT2 = 0.007 + 1.019 × HRT3	<0.0001		
Global reference height (mm)	0.326±0.118	HRT2 = 0.001 + 0.996 × HRT3	0.09	-0.0009	0.02

RNFL = retinal nerve fiber layer.

The mean difference between HRT2 and HRT3 is given for parameters that did not show a statistically significant difference between HRT2 and HRT3 as a function of the mean value.

* For paired *t* test assessing constant bias.

Table 2

Agreement between the Heidelberg Retina Tomograph Version 2 (HRT2) and HRT3 Moorfields Regression Analysis (MRA) Classification (Raw Agreement = 74%, $\kappa = 0.60$)

	HRT3 MRA			
	Within Normal Limits	Borderline	Outside Normal Limits	Total
HRT2 MRA				
Within normal limits				
Count	196	79	1	276
Total %	38.89	15.67	0.20	54.76
Column %	100.00	68.10	0.52	
Row %	71.01	28.62	0.36	
Borderline				
Count	0	36	49	85
Total %	0.00	7.14	9.72	16.87
Column %	0.00	31.03	25.52	
Row %	0.00	42.35	57.65	
Outside normal limits				
Count	0	1	142	143
Total %	0.00	0.20	28.17	28.37
Column %	0.00	0.86	73.96	
Row %	0.00	0.70	99.30	
Total	196	116	192	504
	38.89	23.02	38.10	

Table 3

Agreement between the Heidelberg Retina Tomograph Version 3 (HRT3) Moorfields Regression Analysis (MRA) and Glaucoma Probability Score (GPS) (Raw Agreement = 64%, $\kappa = 0.47$)

	HRT3 MRA			Total
	<i>Within Normal Limits</i>	<i>Borderline</i>	<i>Outside Normal Limits</i>	
GPS				
Within normal limits				
Count	154	54	26	234
Total %	30.56	10.71	5.16	46.43
Column %	78.57	46.55	13.54	
Row %	65.81	23.08	11.11	
Borderline				
Count	27	33	31	91
Total %	5.36	6.55	6.15	18.06
Column %	13.78	28.45	16.15	
Row %	29.67	36.26	34.07	
Outside normal limits				
Count	15	29	135	179
Total %	2.98	5.75	26.79	35.52
Column %	7.65	25.00	70.31	
Row %	8.38	16.20	75.42	
Total	196	116	192	504
	38.89	23.02	38.10	