OPEN

Comparison of Neuroretinal Rim Area Measurements Made by the Heidelberg Retina Tomograph I and the Heidelberg Retina Tomograph II

Ya Xing Wang, MD,*† Neil O'Leary, MSc,‡§ Nicholas G. Strouthidis, MD, FRCOphth,*
Edward T. White, BA, Hons,* Tuan A. Ho, PhD,* and David F. Garway-Heath, MD, FRCOphth*

Purpose: To investigate the agreement between neuroretinal rim area (RA) measurements using the Heidelberg Retina Tomograph I (HRT Classic) and Heidelberg Retina Tomograph II (HRT II). To compare apparent RA changes in follow-up series of HRT II topographies when using either an HRT Classic or HRT II mean topography as baseline.

Design: Cross-sectional study and "no-change," short time series study.

Participants: Forty-three ocular hypertensive and 31 primary open angle glaucoma subjects.

Methods: Five HRT Classic and 5 HRT II examinations were acquired from 1 eye of each subject, across 2 visits within 6 weeks. For the cross-sectional study, follow-up RA measurements from HRT Classic and HRT II were compared, using the same HRT Classic mean topography as the baseline. The linear rates of RA change were compared in 2 short time series with either an HRT Classic or an HRT II mean topography as baseline, and 4 follow-up HRT II mean topographies. Intervals between topographies were arbitrarily set at 1 year for meaningful comparisons of rates. Rates of RA change over time were calculated by linear regression. Separate analyses were performed using 3 available reference planes (RP).

Main Outcome Measures: Global and sectoral RA measurements in HRT Classic and HRT II mean topographies; linear rates of RA change.

Results: HRT Classic minus HRT II mean differences (95% limits of agreement) were 0.09 (−0.17, 0.35) mm^2, 0.09 (−0.13, 0.32) mm^2, and 0.11 (−0.24, 0.46) mm^2 for the Moorfields, 320 μm, and standard RPs, respectively ($P < 0.001$ for all RPs, Wilcoxon rank sum test). In the time series, the mean differences (95% limits of agreement) of RA rates of change (HRT Classic baseline minus HRT II baseline) were −0.01 (−0.06, 0.03) mm^2/y, −0.01 (−0.06, 0.04) mm^2/y, and −0.0^2 (−0.09, 0.05) mm^2/y using the Moorfields, 320 μm, and standard RPs, respectively.

Conclusion: Although HRT software is backward-compatible, follow-up RA measurements made in the same eye using HRT Classic and HRT II devices display statistically and clinically meaningful systematic differences when HRT Classic topographies are used as a baseline.

Key Words: neuroretinal rim area, Heidelberg retina tomography, comparison, time series, reference plane

(J Glaucoma 2013;22:652–658)

Glaucoma is a chronic progressive optic neuropathy, and the examination of the optic nerve head (ONH) is a central component in the diagnosis and follow-up of subjects with, or at risk of, glaucoma.[1–3] Clinical examination of the ONH has long been enhanced by the evaluation of ONH and retinal nerve fiber layer images acquired using traditional photographic techniques. However, the introduction of semiautomated ONH imaging techniques has opened up the possibility of quantifying ONH structural parameters in a more objective and repeatable manner.[4]

Confocal scanning laser ophthalmoscopy with the Heidelberg Retina Tomograph (HRT) has been available in both the clinical and research environments for 2 decades. The original HRT, or "HRT Classic," has now been replaced by the HRT II, a later version of the device. The most recent innovation, the HRT III, is an upgrade of HRT II in terms of operating software although the acquisition hardware remains unchanged from the HRT II.[5]

Both the HRT Classic and HRT II generate 3-dimensional topography images of the ONH from which a range of stereometric parameters can be calculated. Although the 2 versions of the HRT differ in their field of view (10 × 10 degrees for the HRT Classic and 15 × 15 degrees for the HRT II) and their axial sampling, they share the same optical lateral resolution and lateral sampling (∼ 10 μm).[6] The remainder of the differences between the 2 devices largely relates to the operating software and the degree of automation in image acquisition. The HRT Classic was designed as a research instrument, which allows the operator to adjust a number of parameters, such as depth of focus and number of single topography images acquired. The HRT II was designed as a clinical instrument optimized for ease of use and speed of image acquisition. Many of the operator-dependent acquisition variables have been automated for the HRT II.

The HRT Classic uses an MS-DOS operating platform, which was upgraded to a Windows-based program, called Heidelberg Eye Explorer (HEE) for the HRT II.[6] The HEE software is backward-compatible, which means that topographies obtained with the HRT Classic can be assessed using this software. As such, a patient's HRT Classic and HRT II mean topographies can be examined in the same series. The level to which this backward-compatibility performs in practice has a clinically important impact, as many of the reported HRT Classic longitudinal studies were analyzed using HEE.[7,8] Many clinical practices and research groups began examining patients using the HRT Classic and then switched to the HRT II when it was introduced. The power to detect true change in the longitudinal data of each patient may be greatly improved if data from the 2 devices could be analyzed in the same series.[9] However, it is unknown how measurements made using the HRT Classic and the HRT II might differ and how the follow-up assessment could be affected if a longitudinal series was obtained initially with the earlier version of the HRT, and then subsequently with the later version.

Detecting glaucomatous progression by examining changes in the stereometric morphological parameters of the ONH is a common approach. Neuroretinal rim area (RA) is an important index for glaucoma detection and monitoring as a reduction in RA is associated with the loss of retinal ganglion cell axons typical of glaucoma.[10,11] It has been shown that RA exhibits less test-retest variability than other HRT stereometric parameters[12] and is useful for discriminating between glaucomatous and normal eyes.[13] Given these factors, as well as the more straightforward clinical correlate of neuroretinal rim narrowing in glaucoma, HRT RA may be considered a meaningful parameter for detecting glaucomatous progression. Good agreement between RA measurements calculated from HRT Classic and HRT II is crucial to reliable assessment of changes in patients with image series acquired using both devices.
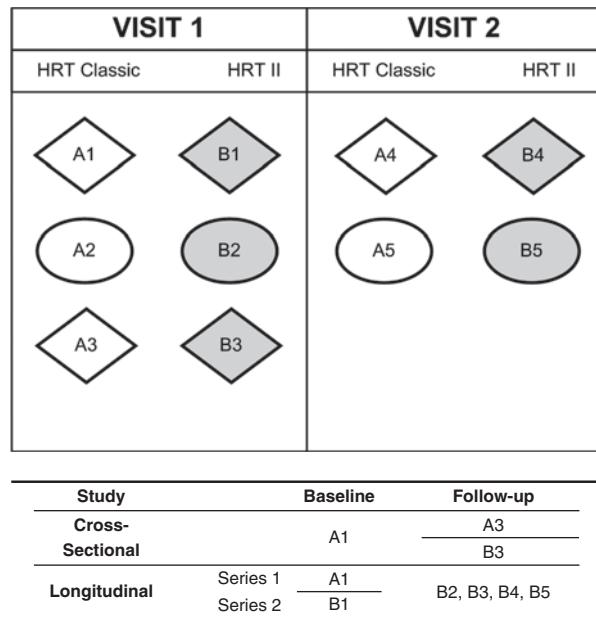
The purpose of this study was to investigate the agreement between RA measurements by the HRT Classic and those by the HRT II in a paired cross-sectional study and to assess the agreement of measured follow-up changes in "no-change" HRT II mean topography series using either an HRT Classic mean topography or an HRT II mean topography as the baseline image.

METHODS

Patient Selection

Seventy-four eyes of 74 subjects recruited from the ocular hypertension (OHT) and the adult general glaucoma clinics at Moorfields Eye Hospital were included in a prospective HRT test-retest study reported by Strouthidis et al.[14] In brief, 43 eyes with OHT and 31 with primary open angle glaucoma were selected. HRT Classic images and HRT II images were acquired by 2 experienced operators (E.T.W., N.G.S.) on the same date (visit 1), and on a second date (visit 2) within 6 weeks of visit 1. A total of 5 HRT Classic mean topographies and 5 HRT II mean topographies were obtained for each eye. A schematic for the visits and examinations is shown in Figure 1.

The subjects had no previous history of intraocular surgery and had all experienced ONH imaging with the HRT Classic. OHT was defined as intraocular pressure > 21 mmHg on 2 or more occasions and a baseline Humphrey 24-2 full threshold Advanced Glaucoma Intervention

Study		Baseline	Follow-up
Cross-Sectional		A1	A3
			B3
Longitudinal	Series 1	A1	B2, B3, B4, B5
	Series 2	B1	

FIGURE 1. Schematic of examinations and study protocols. For all patients, the 2 visits were within 6 weeks of each other. Every participant underwent 5 HRT Classic tests and 5 HRT II examinations. Examinations 1, 3, and 4 (diamond-shaped) were acquired by ETW; examinations 2 and 5 (ellipse-shaped) were acquired by NGS. A3 and B3 were compared in the cross-sectional part, with A1 as baseline topography. In the longitudinal part, series 1 was composed of A1 as baseline topography and B2 to B5 as follow-up topographies. In the same way, series 2 was composed of B1 as baseline examination, and B2 to B5 as follow-up topographies.

Study score of 0.[15] Primary open angle glaucoma was defined as a consistent Advanced Glaucoma Intervention Study visual field score > 0 and a pretreatment intraocular pressure > 21 mmHg on 2 or more occasions. Each eye was selected on the basis that it had a refractive error < 12 D of spherical power and no history of previous intraocular surgery. In subjects with lens opacity, the eye with the greater degree of opacity was preferentially selected although the presence of lens opacity itself was not a criterion for subject selection. This study adhered to the tenets of the Declaration of Helsinki and had local ethics committee approval and the subjects' informed consent.

Image Analysis

The acquired HRT Classic single topographies were imported as HRT-port files into HEE software (Version 1.6.1; Heidelberg Engineering, Heidelberg, Germany). Three HRT Classic single topographies were used to generate each mean topography. The same version of HEE software was used to analyze the HRT II mean topographies. Each mean topography was analyzed using 3 different reference planes (RP): the Moorfields RP, the 320 μm reference plane (320 RP), and the standard RP.[14,16,17] The contour line was drawn onto the baseline mean topography by a single observer (Y.X.W.), and was automatically exported to the follow-up mean topographies to ensure a constant disc area for all mean topographies within the same series. A manual alignment facility was used to optimize contour line position on an image-by-image basis if automatic alignment was judged to have

performed poorly, which was applied in 9 cases. The contour line was placed at the inner margin of the highly reflective halo at the boundary of the disc in the reflectance image (assumed to be the disc margin).[18]

"Cross-sectional" Assessment

This part of the study aimed to compare RA measurements calculated using the HRT Classic and the HRT II in a cross-sectional set of images acquired on the same visit. The protocol for the selection of images in this part of the study is summarized in Figure 1. The 2 selected mean topographies (A3 and B3) were the third HRT Classic and third HRT II mean topographies acquired consecutively by the same operator on the same visit (visit 1; Fig. 1). These were aligned to a baseline mean topography (A1) which was the first HRT Classic mean topography acquired on visit 1. Any systematic differences in RA measurements between devices could thus be appropriately investigated on the basis that they were measured from mean topographies acquired by the same operator within a short period of each other and were aligned in the same follow-up series and using the same contour line. A total of 74 HRT Classic RA measurements paired with 74 HRT II RA measurements were compared globally and in the 6 HRT sectors: temporal, superotemporal, superonasal, nasal, inferonasal, and inferotemporal.

Short "Time Series"

This part of the study assessed how rates of RA measurements may be affected by using either an HRT Classic mean topography or an HRT II mean topography as the baseline for longitudinal series, with follow-up HRT II mean topographies. Given that, in practice, HRT Classic data are more likely to have been acquired before HRT II data, these were considered the 2 most clinically relevant scenarios. The protocol is summarized in Figure 1. Two short time series were generated using either an HRT Classic or an HRT II mean topography as baseline (A1 or B1, respectively) and the same set of 4 follow-up HRT II mean topographies (B2 to B5). The 74 series with HRT Classic baselines were designated "series 1" and the paired 74 series with HRT II baselines were designated "series 2."

This study required 2 corresponding contour lines for baseline mean topographies in series 1 and 2. To ensure that these were as close as possible in shape and location with regard to the anatomy of the ONH, they were drawn in parallel. This was performed to ensure that the paired contour lines were as similar as possible; the control points of each contour line were visually matched to features of the optic disc margin to ensure better correspondence between the contour lines.

In the design of the original test-retest study, we assumed that there would be no underlying glaucomatous structural progression occurring between mean topographies acquired at visits 1 and 2 (within 6 wk). Any detected changes therefore would be as a result of the RA measurement variability, which could be random or potentially systematic in nature. With the short time series composed as above, linear rates of global RA change with time were calculated as the slope of the linear regression of global RA over time.[8,19] An arbitrary 1-year interval was imposed between each successive mean topography for analysis of the series, as this would be representative of typical clinical follow-up periods.

Statistical Analysis

The agreement of RA (global and sectoral) as measured by HRT Classic and by HRT II was assessed with mean differences (bias), 95% limits of agreement (LoA) and using Bland-Altman plots.[20] The Wilcoxon signed rank test was used to establish the statistical significance of differences between RA measurements in the "cross-sectional" study.

The rates of global RA change over time were compared between series 1 and 2. As with the cross-sectional analysis, agreement between groups in rates of RA change as measured by HRT Classic and by HRT II baselines was assessed with mean differences (bias) and 95% LoA. The statistical significance of differences in RA measurements was analyzed using the same methods as in the cross-sectional analysis. The number of statistically significant rates of RA change (linear regression slopes with $P < 0.05$) and the direction of these rates of change (negative or positive) were also determined.

All analyses were performed using SPSS for Windows (version 16.0; SPSS Inc., Chicago, IL).

RESULTS

The 74 subjects had a mean (SD) age of 68.2 (10.2) years, ranging from 20.4 to 84.8 years, and with a male: female ratio of 44:30. The subjects' eyes comprised 42 right eyes and 32 left eyes.

Cross-sectional Analysis

The mean (SD) global disc area (DA) for the HRT Classic and HRT II was 1.80 (0.37) mm². The mean (SD) global RA measurements obtained with HRT Classic (A3 with A1 baseline; Fig. 1) were 1.23 (0.35) mm², 1.17 (0.34) mm², and 1.23 (0.34) mm² using the Moorfields RP, 320 RP, and standard RP, respectively. The mean (SD) global RA measurements obtained with HRT II (B3 with A1 baseline; Fig. 1) were 1.14 (0.35) mm², 1.08 (0.35) mm², and 1.12 (0.35) mm² using the Moorfields RP, 320 RP, and standard RP, respectively. With an HRT Classic baseline topography, global and sector RA measurements were significantly greater in the HRT Classic follow-up topography compared with the HRT II follow-up topography ($P < 0.001$ for all comparisons, 2-way Wilcoxon signed rank test). Table 1 shows the mean RA for the HRT Classic and HRT II topographies and the between-device RA measurement differences for each RP. The mean RA difference (RA of HRT Classic − RA of HRT II) and 95% LoA was 0.09 (−0.17, 0.35) mm² using the Moorfields RP. Similarly, mean (95% LoA) of global RA differences were 0.09 (−0.13, 0.32) mm² and 0.11 (−0.24, 0.46) mm² with the 320 RP and standard RP, respectively. The between-device agreement of global RA measurements is illustrated for all RPs in the Bland-Altman plots in Figure 2. No statistically significant linear relationship was found between global RA differences and average measurements, $P = 0.73$, $P = 0.76$, and $P = 0.87$ for the Moorfields RP, 320 RP, and standard RP, respectively (ie, bias was not proportional).

Short Time Series Analysis

In series 1 (HRT Classic topography at baseline and 4 HRT II topography follow-ups), the mean (SD) rates of RA change were −0.02 (0.02) mm²/y, −0.02 (0.03) mm²/y, and −0.01 (0.04) mm²/y for the Moorfields RP, 320 RP, and standard RP, respectively. In series 2 (HRT II topography at baseline and 4 HRT II topography follow-ups),

TABLE 1. Mean Global and Sector Rim Area (RA) Measurements (mm^2) From the HRT Classic and HRT II Using the (A) Moorfields Reference Plane, (B) 320 μm Reference Plane, and (C) Standard Reference Plane

| Location	HRT Classic (Mean, SD)	HRT II (Mean, SD)	Paired Difference (Mean, 95% LoA)	P
(A)				
Global	1.23, 0.35	1.14, 0.35	0.09 (−0.17, 0.35)	< 0.001
Temporal	0.22, 0.09	0.19, 0.10	0.03 (−0.09, 0.15)	< 0.001
Superotemporal	0.15, 0.05	0.14, 0.06	0.01 (−0.04, 0.06)	< 0.001
Inferotemporal	0.16, 0.07	0.14, 0.07	0.02 (−0.04, 0.07)	< 0.001
Nasal	0.36, 0.09	0.35, 0.10	0.01 (−0.06, 0.08)	0.002
Superonasal	0.17, 0.06	0.16, 0.05	0.01 (−0.03, 0.04)	0.006
Inferonasal	0.17, 0.06	0.16, 0.07	0.01 (−0.03, 0.06)	0.001
(B)				
Global	1.17, 0.34	1.08, 0.35	0.09 (−0.13, 0.32)	< 0.001
Temporal	0.20, 0.10	0.16, 0.11	0.03 (−0.08, 0.14)	< 0.001
Superotemporal	0.14, 0.05	0.13, 0.05	0.01 (−0.04, 0.06)	< 0.001
Inferotemporal	0.15, 0.07	0.13, 0.07	0.02 (−0.03, 0.07)	< 0.001
Nasal	0.36, 0.10	0.34, 0.10	0.01 (−0.05, 0.08)	0.001
Superonasal	0.17, 0.06	0.16, 0.05	0.01 (−0.03, 0.04)	0.003
Inferonasal	0.17, 0.06	0.15, 0.07	0.01 (−0.04, 0.06)	< 0.001
(C)				
Global	1.23, 0.34	1.12, 0.35	0.11 (−0.24, 0.46)	< 0.001
Temporal	0.23, 0.08	0.19, 0.08	0.04 (−0.09, 0.17)	< 0.001
Superotemporal	0.15, 0.05	0.13, 0.06	0.01 (−0.04, 0.07)	< 0.001
Inferotemporal	0.16, 0.06	0.13, 0.07	0.02 (−0.05, 0.09)	< 0.001
Nasal	0.36, 0.09	0.34, 0.09	0.02 (−0.08, 0.11)	0.01
Superonasal	0.17, 0.06	0.16, 0.06	0.01 (−0.03, 0.05)	< 0.001
Inferonasal	0.17, 0.06	0.16, 0.06	0.01 (−0.04, 0.06)	< 0.001

Bias/paired difference values are shown (with 95% LoA) between devices.
The statistical significance of differences between RA values from each device was determined by the Wilcoxon signed rank test.
HR indicates Heidelberg Retina Tomograph; LoA, limits of agreement.

the mean (SD) rates of RA change were 0.00 (0.02) mm^2/y, −0.01 (0.02) mm^2/y, and 0.00 (0.03) mm^2/y for the Moorfields RP, 320 RP, and standard RP, respectively. Figure 3 illustrates the relationship between the rates of RA change in the 2 series. More negative rates were observed in series 1. The mean difference between rates of RA change (RA rate of change series 1 − RA rate of change series 2) and 95% LoA was −0.01 (−0.06, 0.03) mm^2/y using the Moorfields RP. Using the 320 RP and standard RP, mean differences between rates of RA change (95% LoA) were

−0.01 (−0.06, 0.04) mm^2/y and −0.02 (−0.09, 0.05) mm^2/y, respectively. The bias was significant ($P < 0.001$) using all the 3 RPs. No statistically significant linear relationship was found between rate of RA change differences and averaged rate of RA change (proportional bias), $P = 0.35$, $P = 0.09$, and $P = 0.06$ for global RA for the Moorfields RP, 320 RP, and standard RP, respectively.

Small differences were found between DA measurements for the different contour lines drawn onto the baseline mean topographies in series 1 and 2. The mean
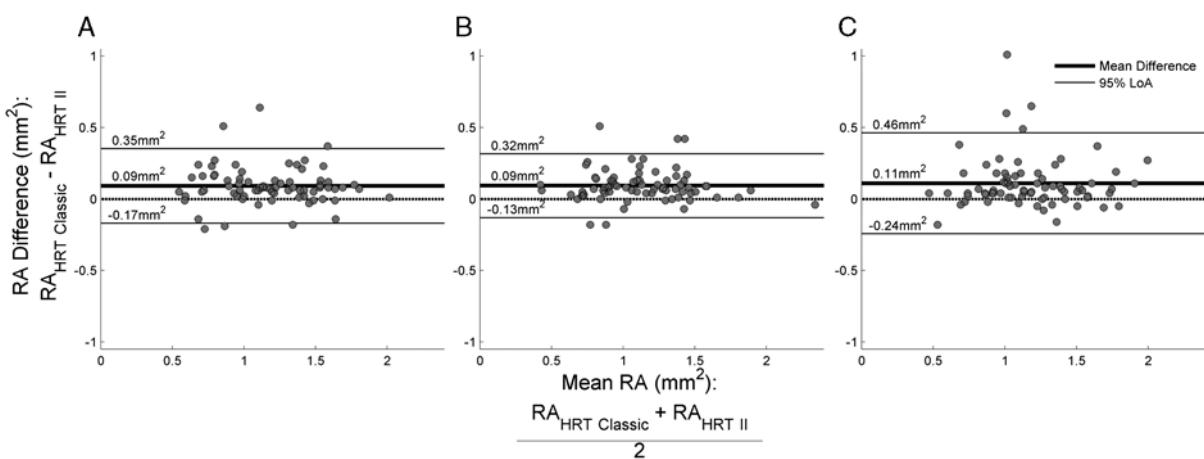
FIGURE 2. Bland-Altman plots of agreement between global rim area (RA) measurements with HRT Classic follow-up (A3) and HRT II follow-up (B3) topographies taken on the same day, by the same operator and sharing the same HRT Classic baseline topography (A1). RA measurement differences with 95% limits of agreement and mean differences are displayed for (A) Moorfields, (B) 320 μm, and (C) standard reference planes. HRT indicates Heidelberg Retina Tomograph.
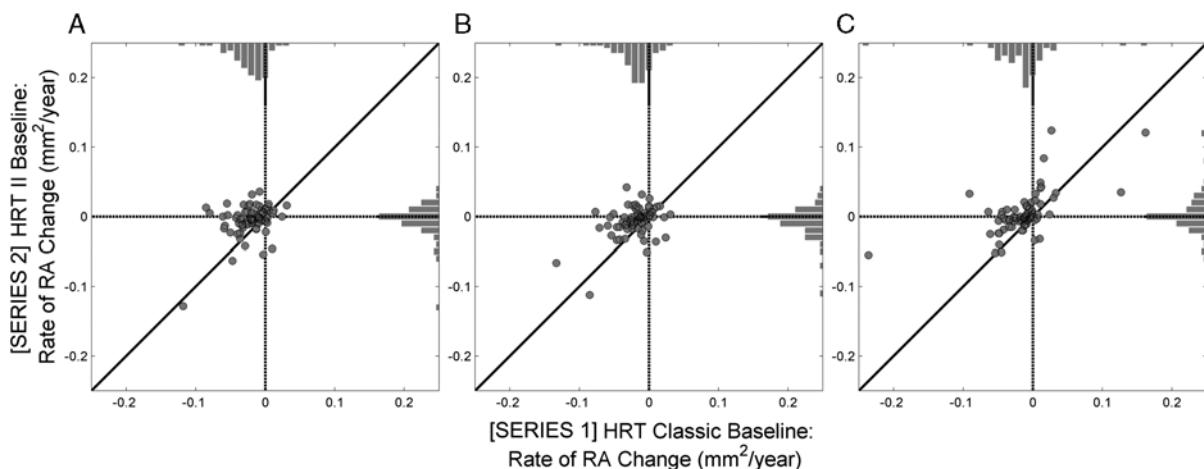
FIGURE 3. Scatter plots showing the relationship between the rates of rim area (RA) change in the short time series with identical follow-up HRT II topographies (examinations B2 to B5) but with HRT Classic (A1) or HRT II (B1) baseline topographies. Rates are displayed for (A) Moorfields, (B) 320 μm, and (C) standard reference planes. Marginal distributions are also shown on the upper horizontal and right vertical axes. HRT indicates Heidelberg Retina Tomograph.

difference (95% LoA) of DA between series 1 and 2 was 0.04 (−0.12, 0.20) mm² or approximately 2% larger on average for series 1. However, no correlation was observed between the difference in DA measurements and the differences in rates of RA change (Spearman $\rho = 0.09$, $P = 0.45$).

Table 2 summarizes the frequency of statistically significant slopes (rates of change). In series 1, marginally more significant negative rates of RA change were observed than in series 2 using the 320 RP and the standard RP and an equal number were observed in both series using the Moorfields RP. More significantly positive rates were found in series 2 than in series 1 for all RPs. None of these differences were statistically significantly different when the 95% confidence intervals were examined.

DISCUSSION

One of the key potential applications of imaging technology in glaucoma is the detection and monitoring of structural progression. Clinicians have been collecting longitudinal data with the HRT Classic since its introduction; it is likely that most users will have switched to the newer HRT II/III upon its introduction. Although the backward-compatibility of the HRT's operational software has allowed researchers to analyze HRT Classic images using the newer software, the rationale for performing the current study was to assess whether longitudinal analyses could be justifiably extended to include HRT II images and HRT Classic images within the same series. The results of this study suggest that combining both HRT Classic and HRT II mean topographies in the same longitudinal series is unwise.

We maintained the same operator, the same imaging session, and the same baseline topography contour line for both the HRT Classic and the HRT II mean topographies of each subject eye in the cross-sectional part of this study. Accordingly, any systematic differences (bias) between RA measurements would be due to either a systematic difference in the data or the HRT software analysis or a combination of both factors. In the "no-change" series, follow-up images were acquired in a short period such that

no measureable morphological (glaucomatous) ONH changes should occur between image acquisitions. It can therefore be safely assumed that systematic changes in RA/time were not pathologic in nature, but related to the software analysis.

In the short time series, baseline HRT Classic and HRT II images were taken by the same operator and follow-up images by both operators. However, any changes caused by different operators would be the same for both series 1 and 2. Thus any bias between the rates of RA change measured in series 1 and 2 would relate to some combination of a systematic difference between HRT Classic and HRT II topographies or the way in which HRT software analyzes these data types.

We found that follow-up RA measurements by the HRT Classic and those by the HRT II were not consistent when analyzed in the same series, with HRT Classic RA measurements being greater than those of HRT II by an average of 0.09, 0.09, and 0.11 mm² for the 3 different RPs. This between-instrument difference was not dependent on RA size and, as a consequence, has a proportionately greater effect on smaller discs. For example, using the Moorfields RP, if the actual RA measurement is 1.00 mm² there is a 9% bias but only a 6% bias if the RA is 1.50 mm². The most marked difference between the RA data from the 2 devices was found in the temporal sector. This is tempered by the wider 95% LoA in this sector which is in accordance with previous findings that RA variability is greatest in the temporal sectors.[21] We observed this RA measurement incompatibility, regardless of the RP used.

In the short time series of series 1, the systematic difference in RA measurements between baseline (HRT Classic) and follow-up (HRT II) impacts the estimated rates of RA change by trend analysis (such as linear regression). If baseline measurements are larger relative to follow-up measurements, then calculated rates are more likely to be negative. This was the case for our study. The biases in the rates of global RA change between series 1 and 2 were between −0.02 and −0.01 mm²/y for the 3 RPs. These differences are similar to observed rates of change in previous longitudinal studies in which the same HRT instrument was used throughout the series. Median rates of

TABLE 2. Direction, Value, and Total Number of Occurrences [With 95% Confidence Interval (CI)] of Statistically Significant ($P < 0.05$) Rates of Global Rim Area Change (mm^2/y), Calculated by Linear Regression, in Series 1 and 2 for Each Reference Plane (RP)

Series	Rate of Change Direction	Moorfields RP Rate	P	320 μm RP Rate	P	Standard RP Rate	P
1	Negative	−0.12	0.002	−0.13	0.02	−0.05	0.04
		−0.04	0.05	−0.09	0.01	−0.04	0.004
		−0.03	0.02	−0.05	0.03	−0.02	0.02
				−0.03	0.02		
	Total (95% CI)	3 (1, 8)		4 (1, 10)		3 (1, 8)	
	Positive	0.01	0.01	None		None	
	Total (95% CI)	1 (0, 5)		0 (0, 4)		0 (0, 4)	
2	Negative	−0.13	0.01	−0.11	0.03	−0.02	0.02
		−0.01	0.02	−0.04	0.01	−0.01	0.01
		−0.01	0.02	−0.01	0.01		
	Total (95% CI)	3 (1, 8)		3 (1, 8)		2 (0, 7)	
	Positive	0.01	0.04	0.03	0.01	0.01	0.01
		0.02	0.02	0.03	0.03	0.05	0.03
		0.02	0.03			0.12	0.01
	Total (95% CI)	3 (1, 8)		2 (0, 7)		3 (1, 8)	

RA loss have been reported as $0.012\,mm^2/y$ from the HRT Classic topography series of a group of OHT patients developing early glaucomatous visual field loss[22] and $0.005\,mm^2$ from the HRT II topography series of glaucoma patients.[23] In any potential studies examining longitudinal series of mixed HRT Classic and HRT II data, these subtle rates of RA change could well be confounded by this relatively large systematic difference.

It is difficult to fully explain the discrepancy between the RA measurements of HRT Classic and HRT II when analyzed as follow-up topographies to an HRT Classic topography. One reason may be the difference in the alignment algorithm used when analyzing a series with an HRT Classic baseline, compared with an HRT II baseline and HRT II follow-up mean topographies. In the latest HRT software version (HRT 3), the alignment algorithm corrects for horizontal and vertical shifts, rotational and tilt misalignment, and differences in parabolic distortion [known as parabolic error correction (PEC)], between baseline and each follow-up topography in a longitudinal series. Because the field of view of HRT Classic imaging is smaller than that of HRT II imaging, PEC is not applied when an HRT Classic examination is included in the longitudinal series.[24] In the image alignment algorithm between HRT II topographic images only, the HEE software performs a parabolic correction to account for the curved surface of the peripapillary retina. It follows from this that parameters obtained after HRT Classic to HRT II alignment, in comparison to HRT II to HRT II alignment, may well differ systematically.

In a recent paper, Balasubramanian et al[24] investigated the more frequent and larger magnitude of negative, topographic change in longitudinal series with combinations of HRT Classic and HRT II topographies as compared to series with HRT II topographies only. When the PEC was applied to the HRT Classic-HRT II combination series, the estimated topographic change in both sets of series was similar. The same group also investigated the difference in stereometric parameters between HRT Classic and HRT II topographies acquired on the same day.[25] Similar to our study, a finding of a statistically significant systematically larger global RA measurement, by an average of $0.06\,mm^2$, in HRT Classic topographies was re-

ported. This between-device RA difference remained even after the application of the PEC. However, the authors concluded that the discrepancy was not clinically problematic as it represented 5% or less of the typical global RA and as clinical LoA between same-day HRT II topography derived RA values were comparable to those limits in same-day HRT Classic and HRT II topographies.[20] Furthermore, statistical significance of bias was assumed to have arisen from the large sample size (n = 344). We have shown the same magnitude and a similar significance of bias in a smaller sample. Furthermore, we have demonstrated that the use of an HRT Classic mean topography at baseline to HRT II follow-up topographies results in a mean negative rate of change and a greater number of significant negative slopes compared to use of an HRTII baseline image. This tendency is likely to have a significant clinical impact on the HRT's ability to reliably detect disease progression if a "mixed" series of images is used.

Some limitations should be pointed out in the current study. Although all the contour lines were drawn by the same examiner and the 2 corresponding baseline mean topography contour lines for each short time series were redrawn very carefully as detailed in the methods, there were inevitably small random differences which could affect the measurements. Our analysis has shown, however, that there is no relationship between the differences in baseline DA between groups and the differences between rates of subsequent RA change. Another potential limitation is that, in the short time series, data were treated as a pseudo longitudinal time series, instead of a true longitudinal series. The advantage of this methodology is that it is certain that no disease progression took place over the image series. This approach has previously been used in the analysis of visual field tests.[26] Further studies will investigate the causes of this systematic difference in RA and examine the effects of combining HRT Classic and HRT II topographies on inferences from statistical change detection techniques designed to look for localized changes in topographic height over time.[27,28]

We have shown that RA measurements are larger with the HRT Classic than with the HRT II, when analyzed as follow-ups to an HRT Classic baseline, and we observed systematically more negative change in time series

composed of an HRT Classic baseline mean topography with HRT II follow-up mean topographies. The bias between HRT Classic and HRT II RA measurements is of a similar magnitude to the longitudinal RA changes reported in glaucomatous eyes and could present a major confounding factor in scientific studies or clinical judgments which combine data from these 2 devices. The results of the current study do not support the use of HRT Classic and HRT II mean topographies in the same longitudinal series. We recommend a new baseline should be created when switching to the newer device, until methods to correct for the current discrepancies become available.

REFERENCES

1. Quigley HA, Katz J, Derick RJ, et al. An evaluation of optic disc and nerve fiber layer examinations in monitoring progression of early glaucoma damage. *Ophthalmology*. 1992; 99:19–28.
2. Harwerth RS, Carter-Dawson L, Shen F, et al. Ganglion cell losses underlying visual field defects from experimental glaucoma. *Invest Ophthalmol Vis Sci*. 1999;40:2242–2250.
3. Garway-Heath DF, Holder GE, Fitzke FW, Hitchings RA. Relationship between electrophysiological, psychophysical and anatomical measurements in glaucoma. *Invest Ophthalmol Vis Sci*. 2002;43:2213–2220.
4. Mikelberg FS, Parfitt CM, Swindale NV, et al. Ability of the Heidelberg retina tomograph to detect early glaucomatous visual field loss. *J Glaucoma*. 1995;4:242–247.
5. Stouthidis NG, Garway-Heath DF. New developments in Heidelberg retina tomograph for glaucoma. *Curr Opin Ophthalmol*. 2008;19:141–148.
6. Heidelberg Engineering. *Operation Manual for the Heidelberg Retina Tomograph II. Vol 1.7.0*. Heidelberg, Germany: Heidelberg Engineering; 2003.
7. Artes PH, Chauhan BC. Longitudinal changes in the visual field and optic disc in glaucoma. *Prog Retin Eye Res*. 2005; 24:333–354.
8. O'Leary N, Crabb DP, Mansberger SL, et al. Glaucomatous progression in series of stereoscopic photographs and Heidelberg retina tomograph images. *Arch Ophthalmol*. 2010;128: 560–568.
9. Schlesselman JJ. Planning a longitudinal study. II. Frequency of measurement and study duration. *J Chronic Dis*. 1973;26: 561–570.
10. Varma R, Quigley HA, Pease ME. Changes in optic disk characteristics and number of nerve fibers in experimental glaucoma. *Am J Ophthalmol*. 1992;114:554–559.
11. Yücel YH, Gupta N, Kalichman MW, et al. Relationship of optic disc topography to optic nerve number in glaucoma. *Arch Ophthalmol*. 1998;116:493–497.
12. Tan JC, Garway-Heath DF, Hitchings RA. Variability across the optic nerve head in scanning laser tomography. *Br J Ophthalmol*. 2003;87:557–559.
13. Wollstein G, Garway-Heath DF, Fontana L, et al. Identifying early glaucomatous changes. Comparison between expert clinical assessment of optic disc photographs and confocal scanning ophthalmoscopy. *Ophthalmology*. 2000;107:2272–2277.
14. Strouthidis NG, White ET, Ho TA, et al. Factors affecting the test-retest variability of the Heidelberg retina tomograph and Heidelberg retina tomograph II. *Br J Ophthalmol*. 2005;89: 1427–1432.
15. Advanced Glaucoma Intervention Study. 2 Visual field test scoring and reliability. *Ophthalmology*. 1994;101:1445–1455.
16. Burk RO, Vihanninjoki K, Bartke T, et al. Development of the standard reference plane for the Heidelberg retina tomograph. *Graefes Arch Clin Exp Ophthalmol*. 2000;238:375–384.
17. Poli A, Strouthidis NG, Ho TA, et al. Analysis of HRT images: comparison of reference planes. *Invest Ophthalmol Vis Sci*. 2008;49:3970–3975.
18. Strouthidis NG, Yang H, Reynaud JF, et al. Comparison of clinical and spectral domain optical coherence tomography optic disc margin anatomy. *Invest Ophthalmol Vis Sci*. 2009; 50:4709–4718.
19. Strouthidis NG, Scott A, Peter NM, et al. Optic disc and visual field progression in ocular hypertensive subjects: detection rates, specificity, and agreement. *Invest Ophthalmol Vis Sci*. 2006;47:2904–2910.
20. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res*. 1999;8:135–160.
21. Tan JC, Garway-Heath DF, Fitzke FW, et al. Reasons for rim area variability in scanning laser tomography. *Invest Ophthalmol Vis Sci*. 2003;44:1126–1131.
22. Owen VM, Strouthidis NG, Garway-Heath DF, et al. Measurement variability in Heidelberg retina tomograph imaging of neuroretinal rim area. *Invest Ophthalmol Vis Sci*. 2006;47:5322–5330.
23. See JLS, Nicolela MT, Chauhan BC. Rates of neuroretinal rim and peripapillary atrophy area change: a comparative study of glaucoma patients and normal controls. *Ophthalmology*. 2009;116:840–847.
24. Balasubramanian M, Bowd C, Weinreb RN, et al. Agreement between Heidelberg retina tomograph-I and -II in detecting glaucomatous changes using topographic change analysis. *Eye*. 2011;25:31–42.
25. Balasubramanian M, Bowd C, Weinreb RN, et al. Agreement between the Heidelberg retina tomograph (HRT) stereometric parameters estimated using HRT-I and HRT-II. *Optom Vis Sci*. 2011;88:140–149.
26. Gardiner SK, Crabb DP. Frequency of testing for detecting visual field progression. *Br J Ophthalmol*. 2002;86:560–564.
27. Chauhan BC, Blanchard JW, Hamilton DC, et al. Technique for detecting serial topographic changes in the optic disc and peripapillary retina using scanning laser tomography. *Invest Ophthalmol Vis Sci*. 2000;41:775–782.
28. Patterson AJ, Garway-Heath DF, Strouthidis NG, et al. A new statistical approach for quantifying change in series of retinal and optic nerve head topography images. *Invest Ophthalmol Vis Sci*. 2005;46:1659–1667.