Agreement between Heidelberg Retina Tomograph-I and -II in detecting glaucomatous changes using topographic change analysis

M Balasubramanian, C Bowd, RN Weinreb and LM Zangwill

CLINICAL STUDY

Abstract

Purpose   To evaluate the suitability of including both Heidelberg Retina Tomograph-I (HRT-I) and HRT-II examinations in the same longitudinal series for HRT topographic change analysis (TCA) and to evaluate parabolic error correction (PEC) to improve the agreement between HRT-I and HRT-II examinations.

Methods   A total of 66 eyes from the University of California San Diego Diagnostic Innovations in Glaucoma Study with baseline HRT-I and HRT-II examinations obtained on the same day and ≥ 3 HRT-II follow-up examinations were included. Two TCA analyses, HRT-I examination at baseline (HRT-I–mixed series) and HRT-II examination at baseline (HRT-II–only series) were compared. Agreement between the HRT-I–mixed and HRT-II–only series were estimated using Bland–Altman plots. Agreement was assessed: (1) using the current HRT software settings (PEC applied only to HRT-II–only series), and (2) modified HRT settings (PEC also applied to HRT-I–mixed series).

Results   With current HRT software settings, the HRT-I–mixed series significantly overestimated change locations (ie, red pixels) compared with the HRT-II–only series as indicated by statistically significant proportional biases in the Bland–Altman analysis. By applying PEC to HRT-I–mixed series there were no statistically significant biases in the TCA parameter estimates compared with the HRT-II–only series.

Conclusion   In some eyes, HRT-I and HRT-II baseline examinations are not interchangeable in TCA analysis without parabolic error correction. HRT-I–mixed series detected more changes characteristic of glaucoma when there were only minimal changes in the HRT-II–only series. Our results suggest that in the majority of cases, with PEC, HRT-I examinations may be included in a longitudinal series containing HRT-II examinations.

Introduction

Detecting glaucomatous progression is one of the most important, yet challenging, aspects of disease management. This task generally is accomplished by examining serial visual fields, serial optic disk photographs and, more recently, serial structural imaging measurements. The latter technique is challenging, in part, because glaucoma is a slowly progressing disease requiring years of follow-up. During this time, structural imaging instruments typically undergo hardware and software improvements. Because patients are followed for many years, they often undergo testing using more than one instrument version. These versions ideally should be backwards compatible and provide comparable measurements.

The Heidelberg Retina Tomograph (HRT, Heidelberg Engineering, GmbH, Heidelberg, Germany) is a confocal scanning laser ophthalmoscope used for topographic

Department of Ophthalmology, Hamilton Glaucoma Center, University of California San Diego, La Jolla, CA, USA

Correspondence: Dr LM Zangwill, Department of Ophthalmology, Hamilton Glaucoma Center, University of California, San Diego, 9500 Gilman Drive 0946, La Jolla, CA 92093, USA. Tel + 1 (858) 534-7686; Fax: + 1 (858) 822-0499. E-mail: zangwill@ glaucoma.ucsd.edu

assessment of the optic nerve head.[1] HRT-I (classic) was introduced for clinical use in 1991 and HRT-II in 1999. The scans acquired using the HRT-I at $10^\circ \times 10^\circ$ field of view and the HRT-II scans have the same transverse resolution of approximately $11\,\mu$m. Therefore, the topographies computed from the HRT-I and HRT-II scans are theoretically compatible and interchangeable. With current software, HRT-I scans are converted to HRT-II format and combined in a single longitudinal series with HRT-II scans to detect localized retinal height changes using topographic change analysis (TCA) and other change detection algorithms.[2–4]

We noticed that, in some eyes, TCA available in HRT software detected significantly larger number of locations with significant decrease in retinal height when using an HRT-I baseline examination compared to when using an HRT-II baseline examination conducted on the same day (for eg, see Figures 1a *vs* b and 2a *vs* b). In HRT software, a data normalization procedure corrects for any horizontal and vertical shifts, rotational and tilt misalignment, and differences in parabolic distortion (known as parabolic error correction (PEC), described in Subjects and methods section and in Appendix A) between baseline and each follow-up examinations in a longitudinal series. Because the field of view of HRT-I examinations ($10^\circ \times 10^\circ$) is smaller than that of HRT-II examinations, PEC is not applied when an HRT-I examination is included in the longitudinal series. We hypothesized that some of the TCA differences observed in Figures 1b and 2b in contrast to Figures 1a and 2a, respectively, are because of the lack of PEC when using HRT-I examination as baseline.

In this study, we compare the TCA results when HRT-I and HRT-II examinations were used as baseline with HRT-II examinations as follow-up. Specifically, we first assess agreement in the TCA results with an HRT-I at baseline (HRT-I–mixed series) without PEC to same eyes with an HRT-II at baseline (HRT-II–only series) with PEC as in the current HRT software. We then change the data normalization procedure to also correct for differences in parabolic distortion when HRT-I examinations are used at baseline (as in Figures 1c and 2c).
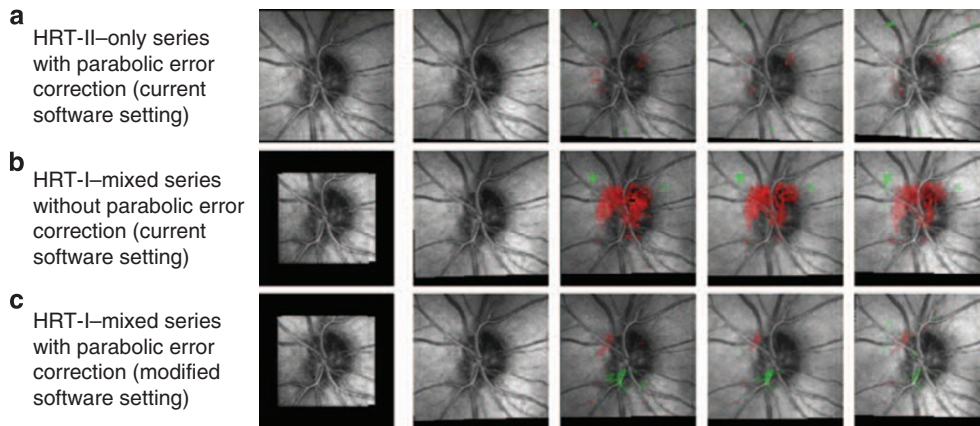
Subjects and methods

Subjects

All participants in the University of California San Diego (UCSD) Diagnostic Innovations in Glaucoma Study (DIGS) with good quality HRT-I and HRT-II examinations acquired on the same baseline visit and with at least three good quality HRT-II follow-up examinations were included in the study. HRT examinations with mean pixel height standard deviation (MPHSD) $<50\,\mu$m were considered to be of acceptable quality for analysis (after quality review according to standard protocols, HRT-I baseline examinations of five eyes in the HRT-I–mixed series and HRT-II baseline examinations of four eyes in the HRT-II–only series had MPHSD between 40 and $50\,\mu$m). The UCSD Institutional Review Board approved the study methodologies and all methods adhered to the Declaration of Helsinki guidelines for research in human subjects and the Health Insurance Portability and Accountability Act.

HRT instrumentation and data preparation

Because 10° HRT-I topographies have the same transverse resolution as 15° HRT-II topographies,[1] HRT software can convert 10° HRT-I topographies to HRT-II format for analysis. During conversion, 256×256 HRT-I
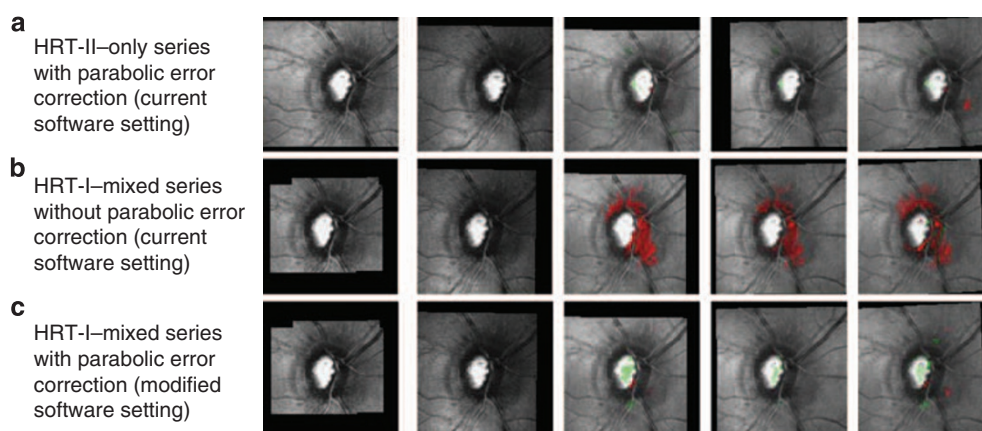
Figure 1 HRT TCA for the HRT-I–mixed series and HRT-II–only series of a UCSD DIGS participant (case 1). HRT-I–mixed series without parabolic error correction (b) detected more number of significant change locations compared with the HRT-II–only series (a). TCA agreement between HRT-I–mixed series (c) and HRT-II–only series (a) improved by applying parabolic error correction also to the HRT-I–mixed series.

a HRT-II–only series with parabolic error correction (current software setting)

b HRT-I–mixed series without parabolic error correction (current software setting)

c HRT-I–mixed series with parabolic error correction (modified software setting)

Figure 2 HRT TCA for the HRT-I–mixed series and HRT-II–only series of a UCSD DIGS participant (case 2). HRT-I–mixed series without parabolic error correction (b) detected more number of significant change locations compared with the HRT-II–only series (a). TCA agreement between HRT-I–mixed series (c) and HRT-II–only series (a) improved by applying parabolic error correction also to the HRT-I–mixed series.

retinal height measurements are embedded in the center of a 384×384 pixel area and measurements outside the central 256×256 region are marked as invalid.

Two separate HRT databases were constructed to evaluate TCA results. In the HRT-II–only series database, only HRT-II examinations were included for analysis. In the HRT-I–mixed series database, the baseline HRT-II examination was replaced with an HRT-I examination acquired on the same day and all HRT-II follow-up examinations were retained as in the HRT-II database. HRT software versions used in this study were HeyEx 1.6.1.0, HRTS 3.1.2.5, and HRTImport 1.3.0.0 (Heidelberg Engineering, GmbH, Heidelberg, Germany).

Optic disk contour margins in both databases were manually drawn by the same person. Contour lines on the HRT-I–mixed series were adjusted to closely match that of the HRT-II–only series (relative differences in disk area were within $\pm 5\%$). Topographies in relative-tilted coordinates were exported from the study databases.[1]
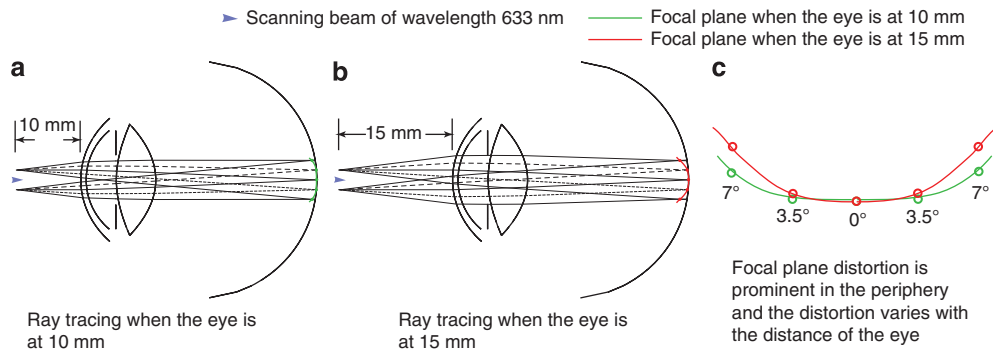
PEC for HRT follow-up topographies

For optic nerve head imaging, an optimal distance of 10 mm is recommended between HRT and the eye.[1] When eyes are imaged at this optimal distance, the focal plane of HRT lies parallel to the retinal surface and HRT images are optimal with minimal distortion. When this optimal imaging distance is not maintained, the focal plane becomes distorted especially in the peripheral region (G Zinser and M Reutter, Heidelberg Engineering, Personal Communication, 2007; also refer to Appendix A).

Effects of the distance between HRT and the eye on the retinal focal plane are illustrated using ray-tracing diagrams in Figure 3. A wide-angle schematic eye model based on the Gullstrand-Le Grand model eye and a

scanning beam of 2 mm diameter and 633 nm wavelength were used to construct the ray tracing diagrams using Optica 3 software (Barnhart Optical Research LLC, Urbana, IL, USA) in Mathematica 7 environment (Wolfram Research Inc, Champaign, IL, USA).[5] We used 633 nm wavelength scanning beams instead of 670 nm scanning beams used in HRT because more accurate refractive indices of cornea (1.3747), aqueous humor (1.336), lens (1.4183), and vitreous humor (1.3347) were available at 633 nm for more accurate ray tracing.[5]

As illustrated in Figure 3, differences in the distance between HRT and the eye result in differences in degrees of distortion of the retinal focal plane that are more prominent in the peripheral region. In HRT, retinal images corresponding to these retinal focal planes are acquired through an imaging pinhole located at a plane conjugate to these retinal focal planes.[6] This conjugate relationship between the retinal focal plane and the HRT pinhole located at the retinal conjugate plane is maintained at all working distances in HRT (G Zinser, personal communication, July 2010). Therefore, retinal measurements (at the pinhole) will have peripheral measurement distortions corresponding to the shape of these retinal focal planes. These differences in the retinal measurements result in differences in the curvature of the HRT retinal topographies that need to be corrected before analyzing HRT topographies for changes over time. Therefore, in addition to aligning topographies, HRT software also corrects for any difference in parabolic distortion of retinal measurements between baseline and follow-up examinations (parabolic error). Because parabolic distortion is more prominent in the peripheral region, HRT software corrects parabolic error only in 15° HRT-II topographies and not in 10° HRT-I topographies. Therefore, the current HRT software does not correct parabolic error when there is at least one HRT-I

► Scanning beam of wavelength 633 nm   ——— Focal plane when the eye is at 10 mm
                                                  ——— Focal plane when the eye is at 15 mm

Figure 3 Ray trace diagrams (a, b) illustrating that differences in the distance between HRT and the eye between any two scans result in differences in the shape of the focal plane traced by the HRT scanning beams (c). A wide-angle schematic eye based on the Gullstrand-Le Grand model eye and scanning beam of 2 mm diameter and 633 nm wavelength were used to generate the ray trace diagrams using Optica 3 software. For clarity, two rays per beam were used. Scanning beams of 633 nm wavelength were used for ray tracing instead of 670 nm scanning beams of HRT because more accurate refractive indices of the cornea, aqueous humor, lens, and vitreous humor were available for modeling at 633 nm. The retinal focal planes were constructed by fitting a curve through the focal points of the HRT scanning beams at $0°$, $\pm 3.5°$, and at $\pm 7°$. Differences in the focal plane distortion because of the differences in distance between HRT and the eye are more prominent in the peripheral region compared with the central region. By the conjugate focal principle, retinal structures at the focal plane are imaged through a pinhole located at a plane conjugate to the retinal focal plane. Therefore, retinal measurements will have peripheral measurement distortions corresponding to the shape of the retinal focal plane.

examination in a longitudinal series (as in the HRT-I–mixed series). The PEC procedure available in the HRT software (described in Appendix A) was implemented in MATLAB version 7.9 (The Mathworks Inc., Natick, MA, USA) and was separately applied to the HRT topographies in relative-tilted coordinates.

Topographic change analysis

HRT TCA is a three-way mixed effects ANOVA model for detecting localized (superpixel) retinal height changes and is commonly used for automated analysis of HRT topographies for detecting glaucomatous changes over time.[2,7–13] HRT TCA was implemented in MATLAB and superpixel change probabilities and mean difference topographies calculated. Four TCA change summary parameters (all units are in superpixels) were computed as described elsewhere:[8] (1) total number of superpixel locations with significant decrease in retinal height (red pixels), (2) size of the largest cluster of red pixels, (3) total number of superpixel locations with significant increase in retinal height (green pixels), and (4) size of the large cluster of green pixels, all within the disk margin.

TCA agreement

Agreement between the HRT-I–mixed and HRT-II–only series was determined using the latest follow-up examination of each participant. For all TCA parameters, difference between the HRT-I–mixed and HRT-II–only series, means, and 95% confidence intervals (CIs) were estimated. Bland–Altman mean *vs* difference plots were

generated and bias between the HRT-I–mixed and HRT-II–only series were estimated.[14,15] Observed bias is considered to be statistically significant only when the 95% CI of the bias does not include 0.

For quantitative evaluation of parabolic error in the HRT-I–mixed and HRT-II–only series, we estimated the mean parabolic error at a radius of $500\,\mu m$ from the center of the parabolic error maps (Figure 7c) generated by the PEC procedure described in Appendix A.

Results

A total of 66 eyes from 66 UCSD DIGS participants were included. Participants ranged in age from 39.48 to 91.16 (median of 71.3) years. The median number of HRT follow-up examinations was five (range 4–7), representing a median of 5.27 years of follow-up. The mean spherical equivalent of refractive errors at baseline was -0.57 (95% CI: -1.08 to -0.07) diopters. There were no tilted optic disks in this study on the basis of the assessment of simultaneous stereophotographs by experienced certified graders. Demographics of the study participants and their baseline characteristics are presented in Table 1.

Agreement between the HRT-I–mixed and HRT-II–only series is summarized in Table 2. Using the current HRT software settings, the TCA red pixel parameters were significantly higher in the HRT-I–mixed series (without PEC) compared with the HRT-II–only series (with PEC). After applying PEC to the HRT-I–mixed series, there were no statistically significant differences between the HRT-I–mixed and HRT-II–only series.

Table 1 Baseline characteristics of the study participants

No. of eyes (no. of subjects)		66 (66)
Baseline age (years)	Mean (95% CI)	69.42 (66.77, 72.08)
	Median (range)	71.33 (39.48, 91.16)
No. of HRT exams	Median (range)	5 (4, 7)
HRT follow-up years	Mean (95% CI)	4.69 (4.23, 5.15)
	Median (range)	5.27 (0.24, 7.98)
Baseline refraction—spherical equivalent (diopters)	Mean (95% CI)	−0.57 (−1.08, −0.07)
	Median (range)	0.00 (−6.25, 2.75)
Baseline disk area (mm^2) from HRT-I exams in the HRT-I–mixed series	Mean (95% CI)	1.92 (1.81, 2.04)
	Median (range)	1.91 (1.03, 3.32)
Baseline disk area (mm^2) from HRT-II exams in the HRT-II–only series	Mean (95% CI)	1.92 (1.80, 2.03)
	Median (range)	1.91 (1.10, 3.46)
Baseline SAP mean deviation (dB)	Mean (95% CI)	−1.06 (−1.88, −0.25)
	Median (range)	−0.20 (−13.73, 2.20)
Baseline SAP pattern standard deviation (dB)	Mean (95% CI)	2.53 (1.84, 3.22)
	Median (range)	1.53(0.99, 13.01)
Baseline % abnormal disk[a] from photo evaluation		27.3% (18 out of 66 eyes)
Baseline % abnormal visual field[b]		4.5% (3 out of 66 eyes)[c]
Baseline % of both abnormal disk[a] from photo evaluation and abnormal visual field[b]		22.7% (15 out of 66 eyes)[c]

Abbreviations: CI, confidence interval; HRT, Heidelberg Retina Tomograph.
[a]Optic disk with excavation, neuroretinal rim thinning or retinal nerve fiber defects indicative of glaucoma by review of simultaneous stereophotographs by two experienced graders.
[b]Visual field pattern standard deviation with $P \leq 0.05$ and/or glaucoma hemifield test outside normal limits by STATPAC analysis.
[c]Visual field exams were not available for 2 out of 66 eyes; these 2 eyes had normal disk from photo evaluation.

Figure 4 shows the Bland–Altman plots for the TCA parameters, wherein parabolic error was corrected only in the HRT-II–only series as in the current HRT software. The estimated biases in Figure 4a and b had statistically significant positive slopes indicating that (1) the HRT-I–mixed series of the study participants significantly overestimated TCA red pixel parameters compared with the HRT-II–only series and (2) the difference between the HRT-I–mixed and HRT-II–only series increased in proportion to the TCA changes detected (ie, proportional bias). For the TCA green pixel parameters, observed biases and mean differences between the HRT-I–mixed and HRT-II–only series were of borderline statistical significance.

Figure 5 shows the Bland–Altman mean *vs* difference plots of the TCA parameters estimated after applying PEC to both the HRT-I–mixed and HRT-II–only series. For all TCA parameters, there was no statistically significant bias between the HRT-I–mixed and HRT-II–only series. PEC to the HRT-I–mixed series removed proportional biases in the red pixel parameter estimates (Figure 5a and b), and slightly increased the variability of differences in the green pixel parameters (Figure 5c and d).

Figure 6 shows the Bland–Altman mean *vs* difference plots of the TCA parameters estimated from the HRT-I–mixed and HRT-II–only series both without PEC. It can be observed that Figure 6 is very similar to Figure 4 indicating that HRT-I–mixed series without PEC has poor agreement with both HRT-II–only series with and

without PEC. The mean parabolic error estimated at a radius of 500 μm from the center of the parabolic error maps (Figure 7c) was 27.42 (95% CI = 23.44, 31.41) μm for the HRT-I–mixed series and −0.07 (95% CI = −0.56, 0.42) μm for the HRT-II–only series. Therefore, the parabolic error observed in the HRT-I–mixed series was significantly larger than in the HRT-II–only series further supporting the similarity observed between Figures 4 and 6.

For all of TCA parameters, there was no statistical significant association between the observed differences in the TCA parameter values and the differences in disk area (in pixels) between the HRT-I–mixed and HRT-II–only series (Kendall's Tau *P*-value > 0.30).

Discussion

This analysis indicates that, with the current software settings, TCA detects more topographic locations with significant decrease in retinal height in longitudinal series with a baseline HRT-I examination (HRT-I–mixed series) without PEC compared with the same series with a baseline HRT-II examination (HRT-II–only series) with PEC. In some eyes, TCA detected changes in the HRT-I–mixed series suggest neuroretinal rim changes characteristic of glaucoma which are not apparent in the HRT-II–only series (Figures 1b and 2b). The differences and biases in the green pixel parameters observed between the HRT-I–mixed and HRT-II–only series were

Table 2 Agreement between the TCA parameters of HRT-I–mixed series (with and without parabolic error correction—PEC), and HRT-II–only series (with PEC; n = 66 eyes)

| Parameter (within optic disk)	Mean (95% CI)			Difference (95% CI)	
	HRT-II-only series with PEC	HRT-I-mixed series without PEC	HRT-I-mixed series with PEC	HRT-I-mixed series without PEC minus HRT-II-only series with PEC	HRT-I-mixed series with PEC minus HRT-II-only series with PEC
Implemented with software version 3.1.2.5	Yes	Yes	No	Yes	No
Total no. of TCA red pixels (superpixels)	49.95 (32.55, 67.36)	143.99[a] (108.15, 179.82)	41.06 (27.15, 54.97)	94.03[b] (58.57, 129.49)	−8.89 (−25.29, 7.51)
Size of the largest cluster of TCA red pixels (superpixels)	32.77 (19.81, 45.74)	106.50[a] (74.40, 138.60)	25.62 (15.75, 35.49)	73.73[b] (41.30, 106.16)	−7.15 (−20.36, 6.06)
Total no. of TCA green pixels (superpixels)	26.29 (17.16, 35.41)	11.52 (3.89, 19.14)	29.24 (19.42, 39.07)	−14.77[b] (−23.38, −6.17)	2.96 (−9.53, 15.44)
Size of the largest cluster of TCA green pixels (superpixels)	14.68 (9.70, 19.67)	6.79 (2.80, 10.77)	18.38 (11.37, 25.38)	−7.89[b] (−12.74, −3.05)	3.70 (−4.49, 11.89)

Abbreviations: CI, confidence interval; HRT, Heidelberg Retina Tomograph; TCA, topographic change analysis.
[a]HRT-I–mixed series parameters significantly different from the respective HRT-II–only series parameters.
[b]Difference between TCA parameters of HRT-I–mixed series and HRT-II–only series statistically significant (ie different from 0); refer to Figures 4-6 for the observed biases.

of borderline significance. Summary measures of both red and green pixel-based TCA parameters (Table 2) and their Bland–Altman plots (Figures 4–6) suggest that correcting parabolic error in the HRT-I–mixed series improves TCA agreement between the HRT-I–mixed and HRT-II–only series. Improved agreement observed for both red and green pixel-based TCA parameters indicates that the PEC procedure did not introduce significant number of erroneous green pixels while controlling red pixels and vice versa.

From the Bland–Altman plots in Figure 5, we observed the presence of a few outlying extreme differences even after applying the PEC to the HRT-I–mixed series. For example, in the largest cluster of red pixels plot in Figure 5b, differences > 100 superpixels are present in five eyes even after applying PEC. There are several possible explanations for the differences in TCA results between the HRT-1–mixed series and HRT-II–only series. There may be differences between 'eye data' of the baseline HRT-I and HRT-II examinations, such as corneal curvature radius, astigmatic corrective lenses, glasses, and contact lenses that effect the magnification and scaling of the topographies and thus may significantly affect any topographic analysis. In this study, however, we observed no causal relationship between the eye data and the extreme outliers (data not shown). In addition, although HRT-I and HRT-II instruments have similar optical design, and acquire scans with compatible transverse resolution, there are several differences in their scan protocols. For example, in HRT-I, three separate optic disk scans are manually acquired in succession to constitute an optic disk examination, whereas three optic disk scans are automatically acquired in succession in HRT-II/3 instrument. Therefore, some differences are expected in the intraexam variability of HRT-I examinations and of HRT-II examinations. Moreover, the HRT-I acquires a fixed set of 32 optical section images irrespective of the scan depth, thus, has a variable axial resolution of 62–128 μm between optical sections. The HRT-II acquires a variable set of 16–64 optical sections and maintains a constant axial resolution of 62.5 μm between optical sections. It is expected that the differences in axial resolution between HRT-I and HRT-II instruments could be a source of variability between the topographies constructed from the HRT-I and HRT-II optic disk scans. There are other general improvements in the HRT-II instruments (such as tilted head rest, prescanning of the optic nerve head, faster scans, and so on.) that may also make it easier to take images at a consistent distance from the eye. The effects of these and other sources of measurement variability than can result in differences in the HRT-I and HRT-II images are difficult to characterize in this small cohort.
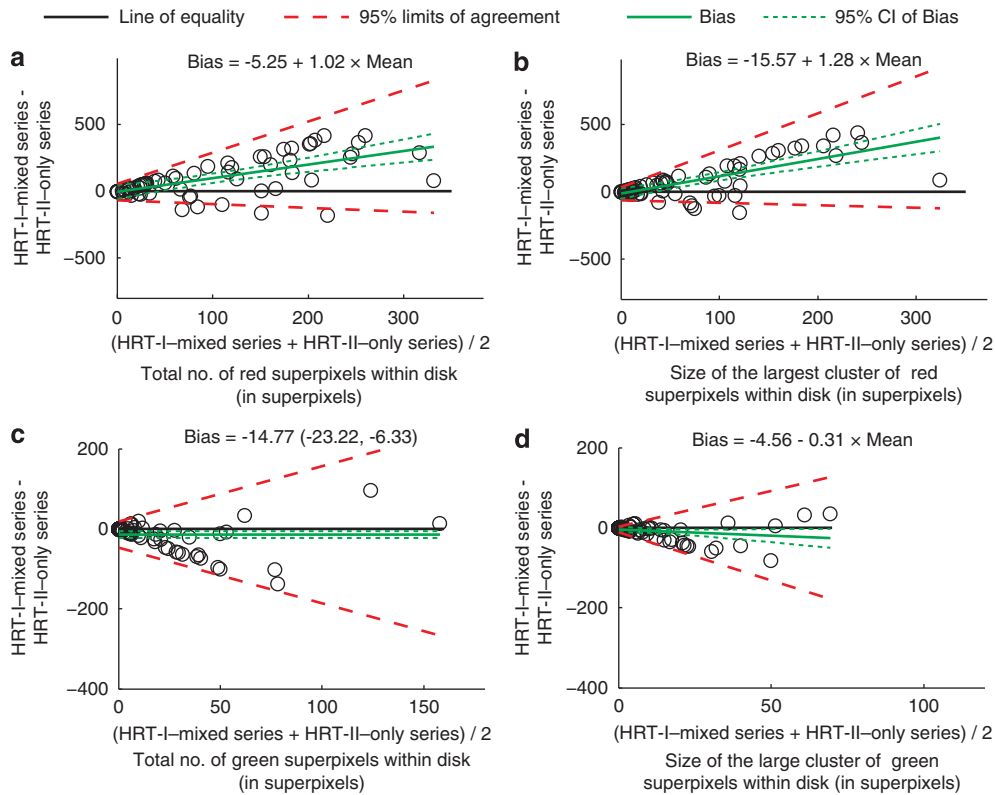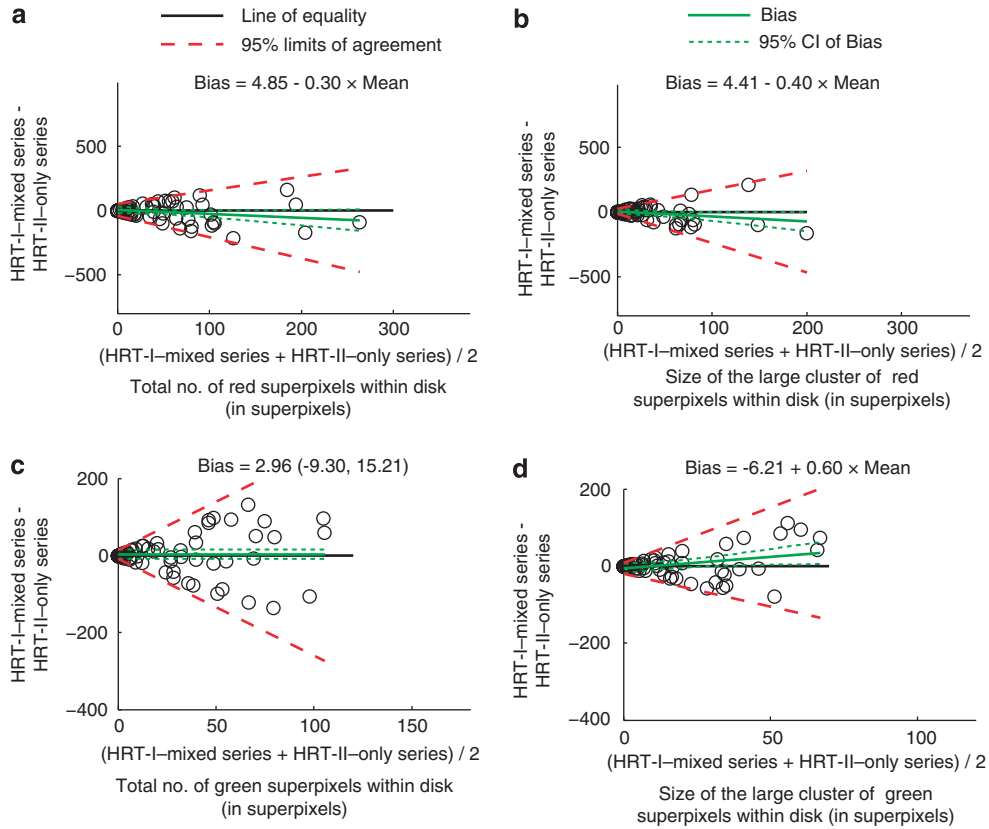
Figure 4 Bland–Altman plots of TCA parameter mean *vs* difference constructed from the HRT-I–mixed series and HRT-II–only series using the current HRT software setting, wherein parabolic error correction is applied only to the HRT-II–only series. Panels a and b (number of red superpixels within the disk) indicate that significant proportional biases exist in the TCA red pixel parameters because HRT-I–mixed series detected more TCA red pixels than the HRT-II–only series. Panels c and d (number of green superpixels) indicate that the differences observed between the TCA green pixel parameters of HRT-I–mixed series and HRT-II–only series are of borderline statistical significance.

Currently, there are no studies on the repeatability of the HRT progression parameters for HRT-I–only longitudinal series and HRT-II–only longitudinal series (ie, how variable the progression parameters are if we replace the baseline examination of a longitudinal series comprised of only HRT-I examinations with another baseline HRT-I examination acquired on the same day; similarly, how variable the progression parameters are for the longitudinal series comprised of HRT-II examinations only). Therefore, it is not clear if the extreme outlying differences observed between the HRT-I–mixed series and HRT-II–only series are also likely between any two repeated HRT-I–only series or between any two repeated HRT-II–only series of an eye.

In addition, we modified the PEC calculations to include two additional strategies to improve the PEC procedure currently available in HRT software (data not shown). In brief, the modified PEC (1) excludes measurements within the optic disk to remove possible influence of optic disk measurements and any glaucomatous changes over time in eyes with disk

diameter $>1000\,\mu$m and (2) adaptively adjusts the dimensions of the peripheral retinal ring to correct for residual error during PEC. The modified procedure provided good agreement between HRT-I–mixed series and HRT-II–only series, similar to the current PEC procedure available in HRT software. By excluding optic disk measurements while estimating differences in parabolic distortion in a follow-up examination, the modified PEC procedure may reduce the possibility of correcting for changes in the follow-up examinations due to significant rim thinning or deepening of the optic cup.

Including an HRT-I examination at baseline with HRT-II follow-up examinations represents a common clinical scenario. In clinical practice, a TCA series may include several HRT-I examinations followed by HRT-II examinations because of software/hardware upgrades. The number of patients with HRT-I baselines and HRT-II follow-up is growing as patients continue to be examined with HRT-II and monitored using TCA analysis. It should be noted that the differences in TCA-detected retinal height decrease results between the

Figure 5 Bland–Altman plots of TCA parameter mean *vs* difference constructed from the HRT-I–mixed series and HRT-II–only series of the study participants after changing the current HRT software setting to apply parabolic error correction to both the HRT-I–mixed and HRT-II–only series. Panels a–d indicate that no statistically significant bias exists between the TCA parameters estimated from the HRT-I–mixed series and HRT-II–only series (line of equality within the 95% CI limits of the bias estimate).

HRT-I–mixed and HRT-II series did not occur in all eyes. In eyes with differences in TCA results, it is likely that the distance between the eye and HRT was different for the baseline HRT-I examinations and the follow-up HRT-II examinations, the condition that PEC was designed to correct. It is therefore difficult to estimate the proportion of eyes that can be expected to show the larger decrease in retinal height with an HRT-I as baseline compared with the HRT-II as baseline because the number can vary by clinic and operator. However, on the basis of significantly lower parabolic error observed in the HRT-II–only series, it can be inferred that HRT-II–only series may provide a more specific detection of glaucomatous changes.

We also investigated the possibility that the PEC procedure may be overcorrecting for differences in the HRT-II–only series thereby causing significant differences with the HRT-I–mixed series without PEC. Bland–Altman plots of TCA parameters estimated from the HRT-I–mixed and HRT-II–only series, both without PEC, also exhibited significant proportional biases in the TCA parameters (Figure 6). In addition, parabolic error

present in the HRT-I–mixed series was significantly higher ($27.42\,\mu$m) than the HRT-II–only series ($-0.07\,\mu$m). Therefore, it is essential to correct parabolic error in the HRT-I–mixed series.

It is not readily apparent why HRT-II–only series have significantly lower parabolic error compared to the HRT-I–mixed series. In HRT-I topographies, peripheral retinal region outside 10° are not available. Therefore, during PEC, residual errors (calculated from retinal measurements outside the central 10°) are not corrected in the HRT-I–mixed series. We considered the absence of residual error correction in HRT-I–mixed series as a possible reason for the significantly higher parabolic error observed in the HRT-I–mixed series. To test this hypothesis, we calculated the parabolic error maps for several examples in the HRT-II–only series using only the measurements within the central 10° retinal topographic measurements and thus without correcting for the residual error. However, the parabolic errors were approximately the same for both with and without correcting for residual error in all of the HRT-II–only series examples tested. Further investigation in a
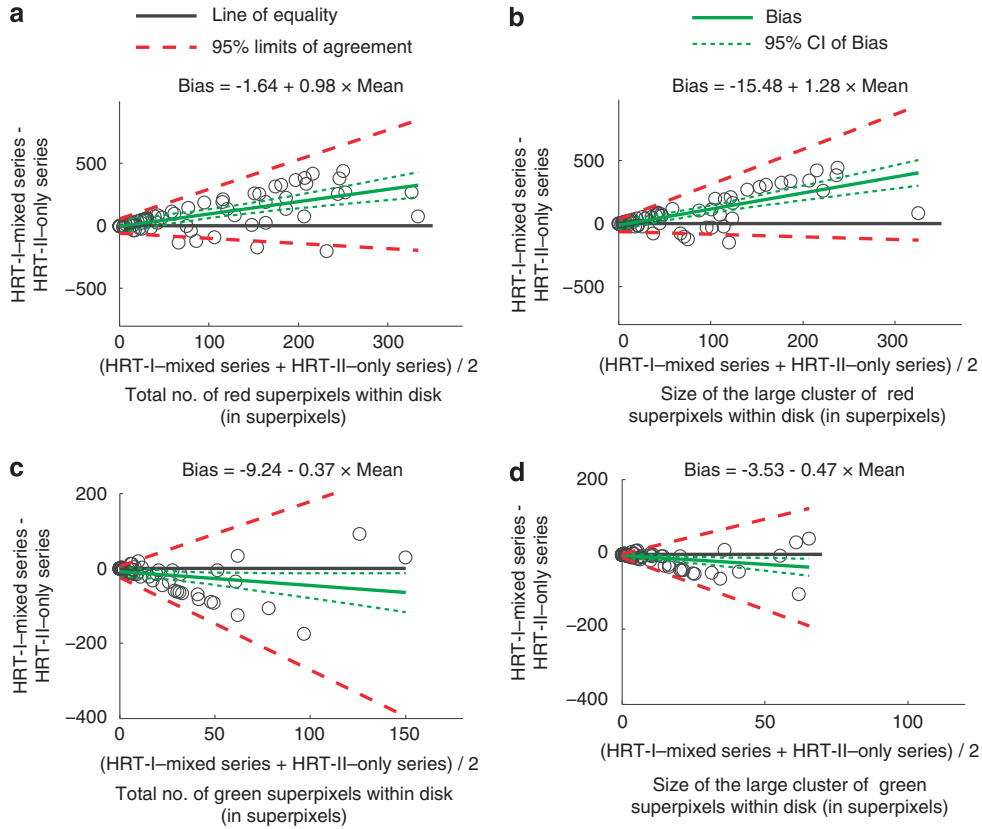
Figure 6 Bland–Altman plots of TCA parameter mean *vs* difference constructed without correcting for differences in parabolic distortion in both the HRT–I–mixed series and HRT–II–only series. Panels a–d indicate that significant proportional biases exist between the estimated TCA parameters because HRT–I–mixed series overestimated red superpixels and underestimated green superpixels compared to the HRT–II–only series.
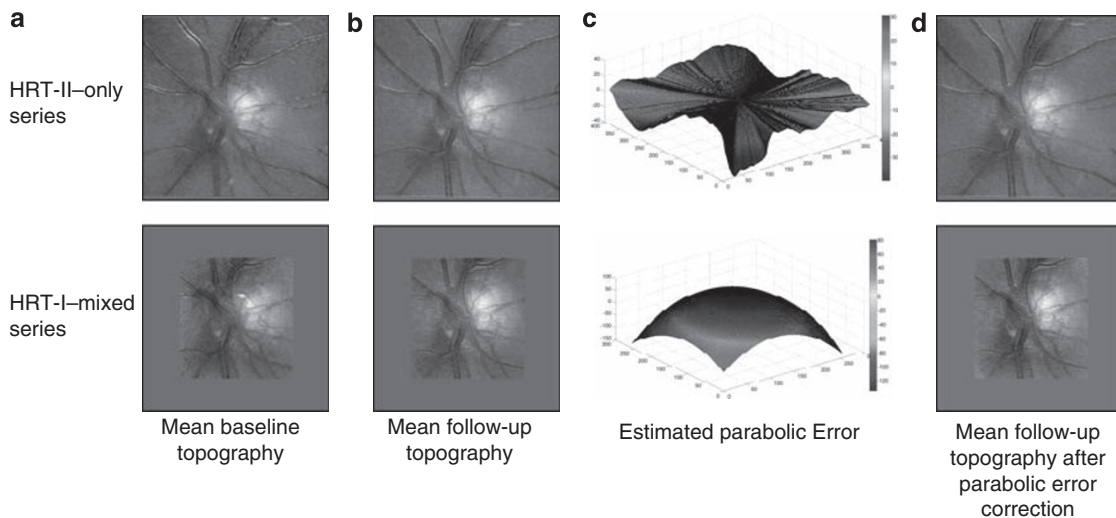
Figure 7 Parabolic error correction applied to an HRT–II–only series (top row) as in the current HRT software (HRTS version 3.1.2.5) and to an HRT–I–mixed series (bottom row; not currently available in HRT software). The total parabolic error estimate shown in c was calculated from the difference between mean baseline topography (a) and the mean follow-up topography (b). The estimated total parabolic error (c) was subtracted from the follow-up topographies (b) to derive the parabolic error corrected follow-up topographies (d). Changes in the mean HRT–II follow-up topography after parabolic correction in panel d (compared with panel b) are not visually apparent because of the wide range of height measurements in the mean topographies. In panel c, the mean parabolic error estimate at a radius of $500\,\mu m$ from the center was $-0.54\,\mu m$ for the HRT–II–only series and $68.6\,\mu m$ for the HRT–I–mixed series.

controlled setting is necessary to understand the sources of higher parabolic error between HRT-I and HRT-II topographies.

Because the TCA disagreement observed between the HRT-I–mixed and HRT-II–only series are due at least in part to parabolic error, such disagreement is also expected when using other change detection algorithms.[3,4] In HRT, some of the rim, cup, and optic disk-related stereometric parameters are estimated using optic disk margin transferred from baseline to each follow-up after normalizing the topographies. Therefore, it is also important to evaluate agreement between the stereometric parameters of HRT-I and HRT-II examinations and the effects of PEC. This is studied in a separate work.[16]

One of the limitations of this study is the relatively small population size—a result of the strict eligibility criterion used for participant selection, where at least one good quality baseline HRT-I examination and one good quality HRT-II examination were required on the same day with at least three good quality HRT-II follow-up examinations for TCA analysis. A possible source of error in this agreement analysis is due to the necessity to manually draw optic disk contour lines separately on the HRT-I–mixed and HRT-II–only series.[17] Possible contour line differences were minimized by closely replicating the contour line of the HRT-I–mixed series on to the HRT-II–only series by the same person (relative disk area differences were within ±5%). Moreover, we found no association between differences in disk size and differences in the TCA parameter estimates. Although this report focuses on the number of red and green pixels within the disk margin, we observed similar results for other TCA parameters including size of the largest red/green clusters relative to disk size, and area and volume of the red/green clusters within disk.

The PEC used in the HRT software is one of the data normalization procedures applied to HRT topography series before analyzing for changes over time. The other commonly known topography normalization procedures correct for horizontal and vertical shifts and rotational and tilt alignment among topographies in a longitudinal series. This study further highlights the importance of the data normalization procedures for preparing the high-resolution optical imaging datasets for progression analysis and their effects in the clinical diagnostic accuracy of the computational methods, such as TCA.

In summary, using the current HRT software, retinal height decreases in some eyes were overestimated by TCA when an HRT-I examination was used at baseline instead of an HRT-II examination. Application of PEC improved the agreement between the HRT-I–mixed and HRT-II–only series with few extreme outlying differences. Therefore, our results suggest that, with PEC, in the majority of cases,

HRT-I examinations may be included in a longitudinal series containing HRT-II examinations.

Summary

What was known before
- Heidelberg Retina Tomograph (HRT)-II/3 instruments can analyze and process optic disk scans acquired with the HRT-1 instruments and the optical system of HRT-I and HRT-II/3 are theoretically backwards compatible.

What this study adds
- We report our findings that in some eyes, glaucomatous progression may be overcalled when HRT-1 scans are used as baseline images with HRT-II scans as follow-up images in the topographic change analysis. This current paper describes our findings about the lack of agreement between HRT-1 scans and HRT-II/3 scans when combined in a single longitudinal series for detecting glaucomatous progression using the current HRT software. Moreover, in this manuscript, we have suggested possible solutions to the problem that can be easily implemented to improve agreement between HRT-1 and HRT-II/3 scans for their combined use in a single longitudinal series.

Conflict of interest

Carl Zeiss Meditec: RNW (F, C), LMZ (F); Heidelberg Engineering: MB (F), LMZ (F), RNW (F); Lace Elettronica: CB (F), Paradigm: RNW (F); Topcon: RNW (F), LMZ (F); Optovue: LMZ (F).

References

1 Heidelberg Engineering, GmbH. *Heidelberg Retina Tomograph: Operating Manual Software Version 3.0* 2005.

2 Chauhan BC, Blanchard JW, Hamilton DC, LeBlanc RP. Technique for detecting serial topographic changes in the optic disc and peripapillary retina using scanning laser tomography. *Invest Ophthalmol Vis Sci* 2000; **41**(3): 775–782.

3 Patterson AJ, Garway-Heath DF, Strouthidis NG, Crabb DP. A new statistical approach for quantifying change in series of retinal and optic nerve head topography images. *Invest Ophthalmol Vis Sci* 2005; **46**(5): 1659–1667.

4 Balasubramanian M, Zabic S, Bowd C, Thompson HW, Wolenski P, Iyengar SS *et al.* A framework for detecting glaucomatous progression in the optic nerve head of an eye using proper orthogonal decomposition. *IEEE Trans Inf Technol Biomed* 2009; **13**(5): 781–793.

5 Escudero-Sanz I, Navarro R. Off-axis aberrations of a wide-angle schematic eye model. *J Opt Soc Am A Opt Image Sci Vis* 1999; **16**(8): 1881–1891.

6 Zinser G, Wijnaendts-van-Resandt RW, Ihrig C. *Confocal Laser Scanning Microscopy for Ophthalmology. in SPIE Scanning Imaging.* SPIE: Hamburg, Federal Republic of Germany, 1988.

7 Artes PH, Chauhan BC. Longitudinal changes in the visual field and optic disc in glaucoma. *Prog Retin Eye Res* 2005; **24**(3): 333–354.

8 Bowd C, Balasubramanian M, Weinreb RN, Vizzeri G, Alencar LM, O'Leary N *et al.* Performance of confocal scanning laser tomograph Topographic Change Analysis (TCA) for assessing glaucomatous progression. *Invest Ophthalmol Vis Sci* 2009; **50**(2): 691–701.

9 Chauhan BC, Hutchison DM, Artes PH, Caprioli J, Jonas JB, LeBlanc RP *et al.* Optic disc progression in glaucoma: comparison of confocal scanning laser tomography to optic disc photographs in a prospective study. *Invest Ophthalmol Vis Sci* 2009; **50**(4): 1682–1691.

10 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**(10): 1492–1499.

11 Kalaboukhova L, Fridhammar V, Lindblom B. Glaucoma follow-up by the Heidelberg Retina Tomograph–new graphical analysis of optic disc topography changes. *Graefes Arch Clin Exp Ophthalmol* 2006; **244**(6): 654–662.

12 Kourkoutas D, Buys YM, Flanagan JG, Hatch WV, Balian C, Trope GE. Comparison of glaucoma progression evaluated with Heidelberg Retina Tomograph II versus optic nerve head stereophotographs. *Can J Ophthalmol* 2007; **42**(1): 82–88.

13 Vizzeri G, Weinreb RN, Martinez de la Casa JM, Alencar LM, Bowd C, Balasubramanian M *et al.* Clinicians agreement in establishing glaucomatous progression using the Heidelberg Retina Tomograph. *Ophthalmology* 2009; **116**(1): 14–24.

14 Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999; **8**(2): 135–160.

15 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; **1**(8476): 307–310.

16 Balasubramanian M, Bowd C, Weinreb RN, Zangwill LM. Agreement between the Heidelberg Retina Tomograph (HRT) stereometric parameters estimated using HRT-I and HRT-II. *Optom Vis Sci* 2011; **88**: (in press).

17 Garway-Heath DF, Poinoosawmy D, Wollstein G, Viswanathan A, Kamal D, Fontana L *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**(6): 664–669.

Appendix A

Parabolic error correction applied only to Heidelberg Retina Tomograph-II (HRT-II)-only series (as in the current HRT software)

Algorithmic details presented here are based on an unpublished technical document from Heidelberg Engineering, GmbH, Heidelberg, Germany describing parabolic error correction procedure available in the HRT-3 software (HRTS module version 3.1.2.5) and through personal communication (Gerhard Zinser and Michael Reutter, December 2006 through January 2008).

The focal plane distortion because of the differences in the distance between HRT and the eye and the resulting retinal measurement distortion in the HRT topographies are approximately parabolic in shape as shown in Figure 3. Therefore, the curvatures of any two retinal topographies of an eye will be different if the same HRT–eye distance is not maintained between examinations. The resulting non-disease related curvature differences between any two retinal topographies of an eye are referred to as the parabolic error between the two examinations. Before conducting change analysis, such as HRT topographic change analysis, curvature differences (ie, parabolic error) among topographies should be removed to ensure that accurate difference images are constructed. Because the curvature differences are not purely parabolic, the parabolic error correction algorithm available in the HRT software updates an initial paraboloid fit for the curvature differences using an estimate of any residual error. The residual error reflects curvature differences present between the two topographies after the initial paraboloid fit, especially in the peripheral retinal region. Algorithmic details of the parabolic error correction algorithm are as follows.

Figure 8 shows the schematic layout of the parabolic error correction available in the current HRT software (HRTS module version 3.1.2.5) and Figure 7 (top row) shows an example parabolic error correction applied to an HRT-II baseline and follow-up examination pair. To characterize any differences in parabolic distortion between the baseline and each follow-up examination, a mean difference topography $D(x, y)$ is calculated as the difference between the mean baseline topography and the mean follow-up topography. After excluding a central circular region of diameter $1000 \, \mu m$ (with center pixel coordinate: 193, 193) and the locations with invalid topographic height measurements (ie, topograph height $= -32768$ in the .RAW files), a paraboloid $P(x, y) = a_1 + a_2 x + a_3 y + a_4 xy + a_5 x^2 + a_6 y^2$ is fit to describe the shape of the difference in the parabolic distortion between baseline and the follow-up examination. The mean difference topography after an initial parabolic error correction is given by $\hat{D}(x,y) = D(x,y) - P(x,y)$. Any residual difference in the parabolic distortion between
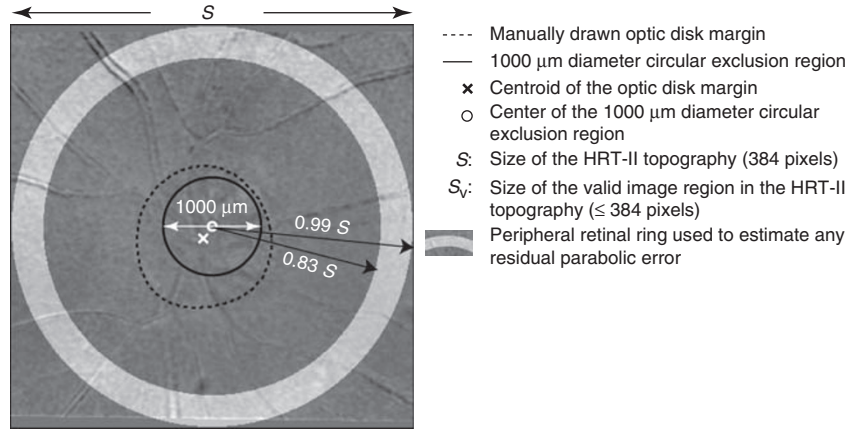
Figure 8 Schematic layout of the parabolic error correction procedure overlaid on an example HRT-II difference topography. Parabolic error correction is a data normalization procedure used in HRT software to correct for any differences in parabolic distortion between a baseline and a follow-up examination. In the current HRT software (HRTS version 3.1.2.5), measurements within the central circular region of diameter 1000 μm are excluded while estimating the parabolic error between a baseline and a follow-up examination. Measurements within a peripheral retinal ring were used to correct for any residual error.

baseline and the follow-up after the initial parabolic error correction is corrected using a residual error correction procedure as follows. The residual retinal height difference at a radius of R_{avg} ($=175$ pixels) from the center of the image between the baseline and follow-up examination is calculated from the parabolic error corrected retinal height differences $\hat{D}$ at an angular interval of $1°$ within a peripheral ring of inner radius $R_{min} = 160$ pixels and outer radius $R_{max} = 190$ pixels, where $R_{avg} = (R_{min} + R_{max})/2$. Thus, the residual retinal height difference (RD) is calculated as

$$RD(R_{avg}, \varphi) = \frac{\sum_{r=R_{min}}^{R_{max}} \sum_{\theta=\phi}^{\varphi+1} \hat{D}(r, \theta)}{\text{Number of pixels within the } 1° \text{ angular segment of } \varphi° \text{ and } (\varphi+1)°}$$

The residual retinal height difference estimates are further spatially smoothed using a moving average filter with a spatial support of $23°$. Using the angularly smoothed peripheral residual retinal height difference estimates (ie, RD), the final residual error correction matrix for the entire imaging area is calculated as follows.

Residual error correction matrix

$$RC(r, \varphi) = RD(r, \varphi) \cdot \left(\frac{r}{R_{avg}}\right)^2$$

Based on the residual error correction matrix estimate $RC(r,\varphi)$, the initial parabolic error correction estimate $P(x,y)$ is revised as $\hat{P}(x,y)$ by fitting a paraboloid to the residual error corrected mean difference topography $D(x,y) - RC(x,y)$.

Total parabolic error $= \hat{P}(x,y) + RC(x,y)$

Finally, differences in parabolic distortion between the baseline and the follow-up HRT examination are removed by subtracting the estimated total parabolic error from each of the 3 single topographies in the follow-up HRT examination.

Parabolic error correction applied to both HRT-I–mixed series and HRT-II–only series (not available in the current HRT software)

The parabolic error correction procedure described above for HRT-II–only series was also applied to all of the HRT-I–mixed series in the HRT-I–mixed database. The only difference is that because there is no valid pixels in the HRT-I topographies converted to HRT-II format in the peripheral retinal region (ie, at radius $R_{avg} = 175$ pixels), the residual error correction matrix $RC = 0$ and therefore, the revised paraboloid fit is same as the initial fit. The total parabolic error is given by the paraboloid fit $P(x,y)$ for the mean height difference topography $D(x,y)$. Figure 7 (bottom row) shows the application of parabolic error correction to an HRT-I–mixed series.