tvst

Article

Increased Depth, Reduced Extent, and Sharpened Edges of Visual Field Defects Measured by Compass Fundus Perimeter Compared to Humphrey Field Analyzer

Ping Liu[1], Bao N. Nguyen[1], Andrew Turpin[2], and Allison M. McKendrick[1]

[1] Department of Optometry and Vision Sciences, The University of Melbourne, Victoria, Australia
[2] School of Computing and Information Systems, The University of Melbourne, Victoria, Australia

Correspondence: Allison M. McKendrick, Department of Optometry and Vision Sciences, The University of Melbourne, Victoria, Australia.
e-mail: allisonm@unimelb.edu.au

Received: March 1, 2021
Accepted: September 12, 2021
Published: October 25, 2021

Keywords: visual field; eye-tracking; glaucoma; structure-function

Citation: Liu P, Nguyen BN, Turpin A, McKendrick AM. Increased depth, reduced extent, and sharpened edges of visual field defects measured by compass fundus perimeter compared to humphrey field analyzer. Transl Vis Sci Technol. 2021;10(12):33, https://doi.org/10.1167/tvst.10.12.33

Purpose: The purpose of this study was to compare visual field results of the COMPASS fundus perimeter (CMP) and the Humphrey Field Analyzer (HFA) in the same eyes; to compare structure-function concordance between circumpapillary retinal nerve fiber layer (Cp-RNFL) profiles and the two perimetry results; and to evaluate whether differences between the two results reflect postulated advantages of real-time eye movement compensation during perimetry.

Methods: We retrospectively analyzed 24-2 visual field data measured with CMP and HFA together with Cp-RNFL optical coherence tomography (OCT) scan data from 95 eyes of 65 people with glaucoma. We defined visual field locations with total deviation (TD) less than −5 dB as defective. The CMP and HFA fields were compared on measures of: spatial extent (number of defective locations); depth (TD values); and sharpness of scotomata edges (maximum TD difference between defective locations and their neighbors). Structure-function concordance between Cp-RNFL profile and respective visual field was also compared.

Results: Compared to the HFA, scotomata measured by CMP were of reduced spatial extent (mean difference = −3.14 locations, $p < 0.001$), greater depth (median TD of CMP = −17 dB versus HFA = −13 dB, $p = 0.029$) and steeper edges (median of maximum TD difference of CMP = 10.6 dB versus HFA = 6 dB, $p < 0.001$). Structure-function concordance between Cp-RNFL profile and either visual field were comparable despite the reduced scotoma spatial extent measured by CMP.

Conclusions: Glaucomatous visual fields measured by CMP displayed characteristics consistent with expected effects of using real-time eye movement compensation technology compared to the widely used HFA.

Translational Relevance: Glaucomatous visual field defects measured by the CMP are more localized, deeper, and steeper than those of the HFA.

Introduction

Accurate and reliable measurements of the location, spatial extent, and depth of scotomata (visual field defects) by static automated perimetry (SAP) are critical for quantifying functional damage in glaucoma and for measurement of the progression of visual damage. Moreover, clinical diagnosis and ongoing management of glaucoma is based on a combination of information from functional and structural measures among other critical factors. Clinicians often look for structure-function matches between the location of anatomic change at the optical nerve head (ONH) and/or circumpapillary retinal nerve fiber layer (Cp-RNFL) thinning in spectral-domain optical coherence tomography (OCT) scans and visual field defects.[1] Corroborating structure-function evidence facilitates clinical decision making. Consequently, accurate and reliable visual field results are crucial.

The accuracy of visual field results is dependent on the test taker maintaining stable central

fixation throughout the test, among other factors. Eye movements as small as 2 degrees have been shown to increase measured sensitivity/dB in scotomatous regions when stimulus grids of 2 degrees inter-stimulus spacing are used.[2,3] Eye movements of 5 degrees have been shown to result in the blindspot being measured as of reduced depth, increased extent, and less steep scotoma edges.[2] As a result, Vingrys and Demirel have cautioned clinicians that fixation instability likely makes shallow and/or small scotomata harder to identify.[2] Importantly, eye movements in the order of $\pm$ 5 degrees are commonly found even in well trained observers.[4] Although there is no direct evidence that eye movements of this magnitude have similar effects on the measurement of scotomata in grids of 6 degrees inter-stimulus spacing, it is probable that methods of perimetry that do not compensate for fixation instability are similarly susceptible to fixation drifts, particularly when the underlying true edge of the scotoma in visual field space lies close to the position of a test point.

Recently, the COMPASS fundus perimeter (CMP; CenterVue, Padua, Italy) has integrated a live infrared retinal tracking technology in its central 30 degrees visual field assessment.[5,6] The integrated live retinal tracking of CMP suspends testing when eye movements are detected. CMP also actively compensates for small eye movements during testing by recalculating the required stimulus projection position prior to each presentation based on the current position of the retina. Hence, in the presence of non-uniform visual field sensitivity, this feature should accordingly improve the accuracy of perimetry tests, with more precise measurement of the location, spatial extent, and depth of scotomata compared to the Humphrey Field Analyzer (HFA).

A second major difference between the CMP and HFA is the testing strategy. CMP uses an implementation of the Zippy Estimation by Sequential Testing (ZEST) strategy that does not exploit spatial correlations between adjacent locations.[7] In contrast, the Swedish Interactive Thresholding Algorithm (SITA) standard strategy used by the HFA incorporates interpoint correlations in its prior models.[8,9] Assuming that visual field damage due to glaucoma manifests as discrete scotomata (as opposed to diffuse loss, which can also occur), we would expect the lack of interpoint correlations in the testing algorithm and eye movement compensation in the CMP to: (1) portray established scotomata with sharper edges; and (2) detect small, isolated scotomata. Whereas recent studies have reported that CMP yields aggregated results comparable to the HFA in many respects,[5,10] whether the CMP realizes these two behaviors has not been systematically investigated. We do so here by quantitatively comparing CMP against HFA visual field results in the same cohort of patients with glaucoma.

A further key difference between the CMP and HFA is that the CMP produces a color retinal photograph and aligns the visual field test locations against the image. This creates the potential to register the CMP retinal photograph against fundus images collected during OCT. That is, the exact location on the retina for each visual field location is known using the CMP but is only assumed using the HFA.

Here, we apply this image registration and quantitatively compare the structure-function concordance between the OCT derived Cp-RNFL thickness profile and CMP versus HFA. A recent study has reported a similar strength of both global and cluster structure–function correlations between Cp-RNFL thickness and the visual field results returned by either the CMP or the HFA.[11] We chose to use a concordance approach rather than compute strengths of correlations, in line with our recent argument that the most reliable relationship to explore structure-function relationships is through the analysis of mapping damaged visual field locations to damage on the Cp-RNFL.[12] The logic for this statement is as follows. First, visual field locations within the central 30 degrees with normal sensitivity could genuinely map to either a normal or abnormal sector of the Cp-RNFL, because an abnormal Cp-RNFL sector does not necessarily derive from axons that represent the central retina. Similarly, there is also no guarantee that damage in a Cp-RNFL sector will appear in the measured central visual field, if the damage in the Cp-RNFL sector arises from axons that represent the more peripheral visual field. Consequently, computing simple correlations between Cp-RNFL thickness and visual field sensitivity is confounded. Here, we use an alternate concordance approach to compare the two visual field tests in this paper.[12] Briefly, a concordance is defined when a defective location (total deviation [TD] values more severe than -5 dB) and the minimum percentile of Cp-RNFL at the mapped location $\pm$ 15 degrees on the ONH is below the bottom first percentile ($P < 0.01$) of the normative reference values. A concordance ratio was then calculated for an eye as the number of concordances divided by the total number of defective locations.

In summary, given the three main differences between CMP and HFA (eye movement compensation, test algorithm, and retinal image registration for structure-function mapping) in the context of an

otherwise equivalent test grid, we tested the following hypotheses:

1. The difference in the number of defective locations measured by CMP versus HFA differs depending on the number of defective locations within the visual field. Specifically, CMP will report more locations as defective when the number of defective locations is small (corresponding to an improved ability of CMP to detect isolated scotomata). In contrast, CMP will report fewer defective locations than HFA for visual fields where the number of defective locations is large (established scotomata will appear smaller, with sharper edges, with CMP).
2. The TD values of defective locations will be generally lower (deeper defects) measured by CMP compared to HFA (lack of spatial smoothing of scotomata using the CMP due to both eye movement compensation and algorithm differences).
3. The maximum gradient between a defective location and its neighbors in visual fields will be higher using CMP than HFA (sharper scotoma edges for CMP).
4. The structure-function concordance ratio between the OCT Cp-RNFL thickness profile and CMP versus HFA visual fields will be higher for CMP than HFA due to improved measurement of discrete scotomata in addition to more accurate registration of the retinal images to the OCT for CMP relative to HFA.

Methods

Database

Retrospective data of 124 eyes of 79 patients with glaucoma recruited as part of a larger validation study of the CMP fundus perimeter were reviewed.[10] The diagnoses of all eyes were made by expert evaluation of glaucomatous optic neuropathy, that is, the optic disc and/or Cp-RNFL appearance in fundus photographs, or OCT Cp-RNFL scans as per the original study protocol.[10] Patients with fewer than two defective locations (TD < -5 dB) in both the CMP and HFA visual fields were excluded (i.e. patients with two or more defective locations in either field were included). This is a conservative choice that weakens the evidence base for supporting hypothesis 1 (that CMP will detect more isolated scotomata than HFA) but makes it more likely that visual field loss is genuinely present. This study adhered to the tenets of the Declaration of Helsinki, and all data included was collected at the Melbourne test site of the larger study.

Visual Field Data

All eyes completed one CMP New Grid test with ZEST strategy and one HFA 24-2 test using SITA standard in random order. Both devices have a background luminance of 31.5 asb and a maximum luminance of 10,000 asb. The CMP tested all 52 non-blindspot locations of the 24-2 grid and 12 additional macular locations. However, only the 52 locations in common with the 24-2 grid were analyzed herein. Patients' pointwise sensitivity thresholds from both devices were compared to the respective perimeter's normative dataset and then converted to TD values, calculated for each of the 52 locations as the difference from the normal sensitivity. A test location was classified as defective if the TD value was less than -5 dB. A color fundus photograph was taken as a standard part of the CMP test.

OCT Data

Patients' Cp-RNFLs were scanned with the Glaucoma Module of the Spectralis SD-OCT (Heidelberg Engineering, Heidelberg, Germany). Manufacturer's segmentations were extracted for a 3.5 mm diameter Cp-RNFL scan of 768 A-scans in a 360 degrees ring. Automated Cp-RNFL segmentation was inspected and confirmed by a trained operator. The Cp-RNFL thickness measurements were exported post segmentation using proprietary tools (Heidelberg Eye Explorer; Heidelberg Engineering), with the location of the ONH relative to the fovea automatically determined. The Cp-RNFL thickness measurements were compared with the normative values (modeled as a normal distribution using means and standard deviations from the Heidelberg Engineering US Mixed Database available for the Spectralis OCT) and classified as within the normal reference range, between the first and the fifth percentile of the normal reference values, or below the first percentile of the normal reference values. The normative database and calculations used are equivalent to those in the commercial instrument.

The position of the temporal raphe, required for the individualized structure-function mapping, was estimated from a nominally vertical macular cube scan protocol using the method of Bedggood et al.[13] The temporal raphe angle was set at the population norm of 174 degrees if a patient's vertical macular cube scan was not available.[14]

Analyses

Visual Field Results

For the direct comparison between CMP and HFA visual field results, we selected three main measures corresponding to the first three hypotheses: (1) the number of defective locations (hypothesis 1); (2) the magnitude of TD values of defective locations (hypothesis 2); and (3) maximum TD difference between defective locations and their surrounding locations (hypothesis 3). Visual field locations were classified as an "agreed defective" location if the TD values of both fields were lower than −5 dB and as a "nonagreed defective" location for the respective field if the TD value of one field was lower than −5 dB but not for the other. Comparisons were made for all locations, agreed and nonagreed defective locations. Paired and independent *t*-test and their nonparametric alternative Wilcoxon rank-sum and Wilcoxon signed-rank test were used as the inferential statistics test where appropriate. The 95% confidence intervals (CIs) of fitted model parameters were calculated by bootstrapping. All analyses were performed using R version 3.5.3.[15]

The raw number of defective locations per eye does not provide information regarding the number of discrete scotomata nor the area of scotomata measured by CMP and HFA. To further quantify the number and area of scotomata as measured by each machine, we defined a scotoma to be any group of defective locations that were adjacent to each other in any of the up to eight possible neighboring directions, but not crossing the midline. We refer to this as connected component analysis in the results below (see Fig.1 for an example).

Structure-Function Mapping and Concordance Calculation

An individualized structure-function mapping scheme was used to calculate the point-wise concordance between visual field locations and OCT.[12,16] The individualized model incorporates parameters including: (1) the vertical and horizontal position of the ONH relative to the fovea; (2) the orientation of the temporal raphe with respect to the fovea–Bruch's membrane opening axis. Recently, Turpin et al. have demonstrated that the individualized mapping approach has resulted in significantly improved structure-function concordance between structural and functional data in an empirical dataset relative to a commonly used population average map.[12]

Anatomic features for each eye in our empirical dataset were input to the model to derive individualized maps for both the CMP and HFA visual field locations. Note although the CMP and HFA 24-2 patterns share identical coordinates in terms of visual angle for the 52 non-blindspot locations, the fundus image taken by CMP allows alignment between CMP fundus image with the OCT infra-red image to compensate for any misalignments. This process slightly modified CMP test location coordinates and is described in detail in the following section. A structure-function concordance was defined when a visual field location was defective and the minimum Cp-RNFL thickness within ± 15 degrees of the mapped location on the ring scan was below the first percentile of the population norms. The choice of sector size was based on our previous work investigating practical limitations to the resolution of structure-function mapping once the effects of imprecision in the input clinical measures are considered.[16] The concordance ratio of an eye was then calculated as the number of concordances in an eye divided by the total number of defective visual field locations.

The structure-function concordance ratios between Cp-RNFL thickness profile and CMP versus Cp-RNFL thickness profile and HFA fields were compared for all, agreed and nonagreed defective locations using paired and independent *t*-test and their nonparametric

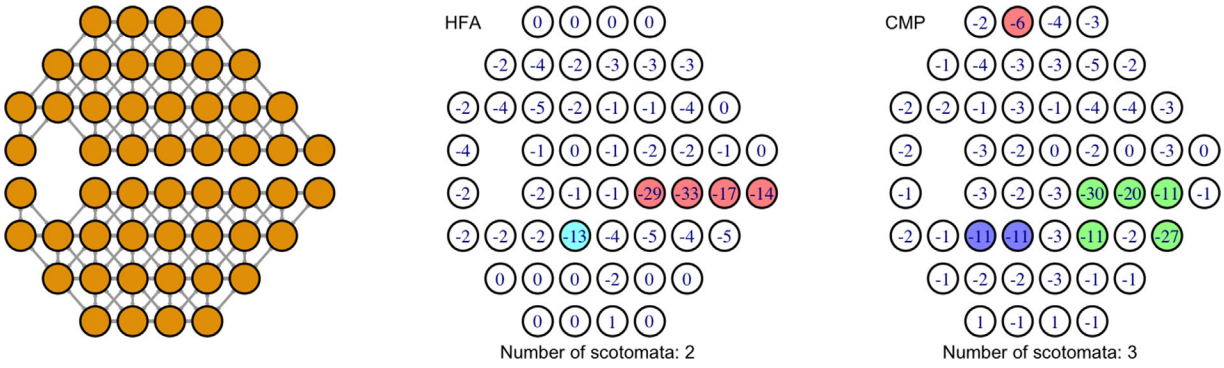

Figure 1. *Left*: An illustration of the neighborhood connections in the 52 non-blindspot locations of 24-2 grid. *Middle and right*: The connected component analysis results of an example patient's Humphrey Field Analyzer (*middle*) and COMPASS (*right*) fields with printed TD values. Each connected component (scotoma) is uniquely colored in each field.

alternative Wilcoxon rank-sum and Wilcoxon signed-rank test where appropriate. The hypothesized higher structure-function concordance ratio for CMP versus HFA (hypothesis 4) was only predicted to manifest for all and nonagreed defective locations as the agreed defective locations between CMP and HFA visual fields share almost identical coordinates and therefore mapped ONH sectors.

CMP Fundus Image Alignment With OCT ONH Image (Visual Field Locations Rotation)

For each set of CMP fundus image and OCT ONH image, an automated image alignment using the scale-invariant feature transform algorithm was applied to the images to obtain the rotation degree of the fundus image relative to the OCT ONH image.[17,18] The OCT ONH image was overlaid on top of the CMP fundus image with the fundus image being rotated by the rotation degree calculated by the automated image alignment. Author P.L. visually inspected all resulting overlaid images to determine whether the ONH, upper and lower branches of the central retinal artery and vein were aligned. The automated alignment successfully aligned 67 out the 124 available image sets. The remaining 57 sets were manually aligned by two researchers independently. The final rotation of each fundus image was calculated as the mean of the two values obtained from the manual alignments. The coordinates of the CMP 52 non-blindspot locations were subsequently recalculated. Figure 2 illustrates this alignment for an example patient. For easy presentation of structure-function concordance, we use Cartesian coordinates centered on the fovea because these are the conventional coordinate systems for visual field

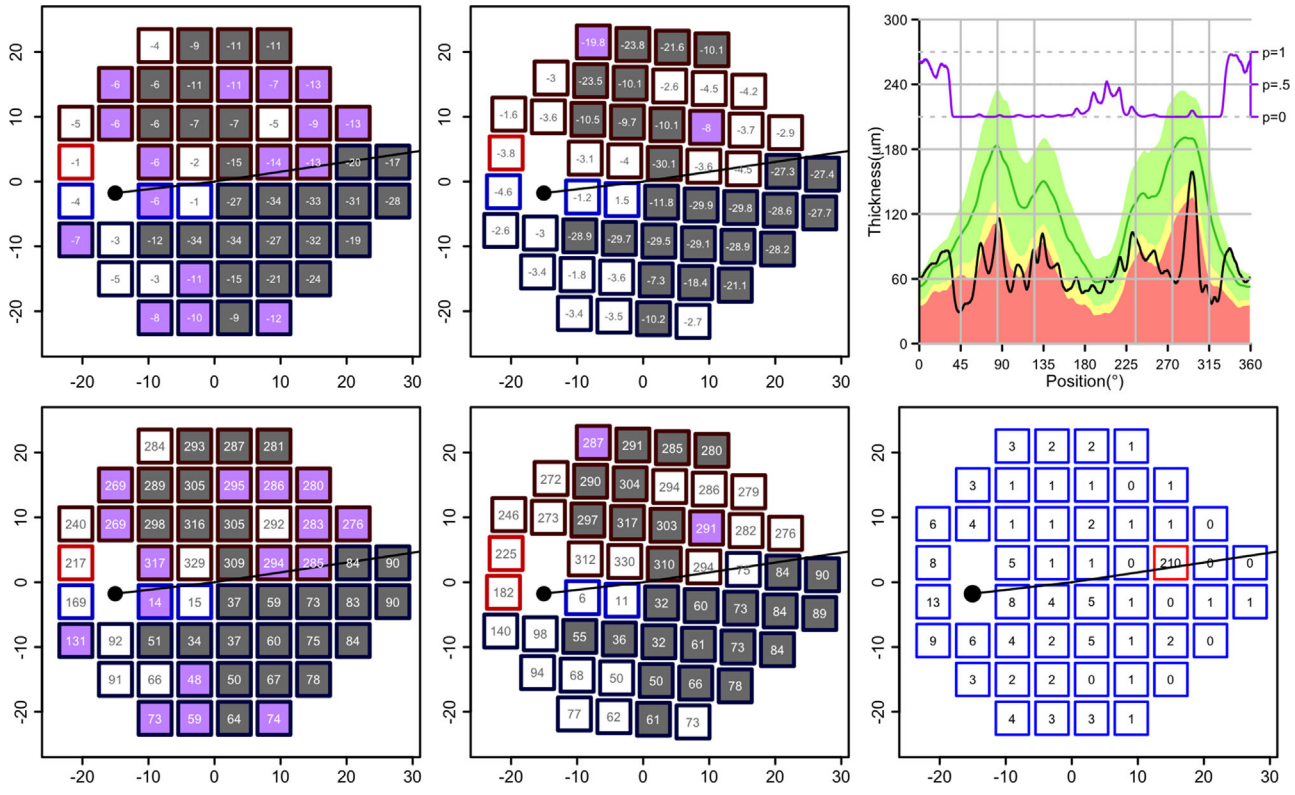

Figure 2. The Humphrey Field Analyzer (HFA) and COMPASS (CMP) 24-2 total deviation (TD) outputs, OCT Cp-RNFL thickness profile of a patient. *Left and middle column*: 24-2 TD output of HFA and CMP respectively with printed TD values on the top and mapped ONH sector on the bottom. Note the CMP visual field is rotated due to alignment process with the OCT infra-red image. The *square symbols* are color coded in these four plots with *dark-grey/purple background* colors representing agreed versus nonagreed defective locations and *red/ blue* foreground colors representing locations mapped onto the inferior versus superior ONH sectors respectively. *Light* versus *dark red* and *blue* represents locations mapped onto a 30 degrees ONH sector with the minimum $P > 1\%$ and $< 1\%$, respectively. The *solid black circle* represents the location of the ONH and the *black line* the Disk-Fovea-Raphe angle. *Right column top* plot: Cp-RNFL thickness profile of the patient (*black line*) and the population mean (mean normal: *green line*) and confidence intervals in shaded areas with *green*, *red*, and *yellow* indicating $P >= 5\%$, $<5\%$, and $<1\%$, respectively. *Right column bottom* plot: Differences (in absolute degrees) in mapped ONH sector between CMP and HFA corresponding visual field locations using individualized structure-function mapping with *red* square(s) highlighting the location(s) whose mapped ONH sector between the CMP the HFA results differs more than 15 degrees apart due to CMP fundus image and OCT infra-red image alignment.

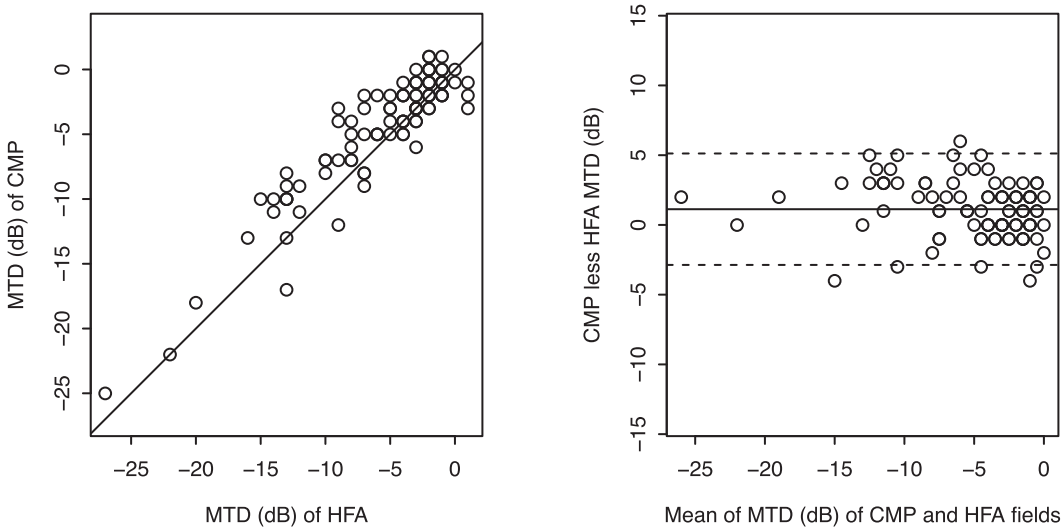

Figure 3. *Left*: Scatter plot of mean total deviations (MTDs) of the COMPASS (CMP) and the Humphrey Field Analyzer (HFA) results. The *black solid line* indicates the situation if agreement were perfect. *Right*: Bland-Altman plot of the same data. The *black solid line* and *dashed lines* indicate the mean difference (1.1 dB) of MTD measured by CMP versus HFA, and 95% confidence intervals, respectively.

location specification. All results are presented in the left eye format. Positive latitude refers to the superior visual field therefore inferior retina, and longitude increases in the nasal direction. For Cp-RNFL thickness profiles, 0 degrees sector starts from the temporal side going anticlockwise (3, 12, 9, and 6 o'clock corresponding to 0 degrees, 90 degrees, 180 degrees, and 270 degrees, respectively) in the retinal space.

Results

Summary Statistics

Out of the 124 eyes of 79 patients with glaucoma with CMP test results, HFA 24-2 test results, and SD-OCT scans, 95 eyes of 65 patients were included in the final analysis. The rest were excluded due to having less than two defective locations in both visual fields. Manual alignment of fundus and ONH OCT images were performed for 48 out of the included 95 eyes. The position of the temporal raphe was estimated from the nominally vertical macular cube scan protocol for 87 of the 95 eyes. Relevant summary statistics of the included patients are reported in Table 1, and further explained below. The agreement of the mean of total deviation (MTD) of all 52 non-blindspot locations between CMP and HFA is presented in Figure 3 with the MTD of CMP being significantly greater (less damaged) than that of HFA ($P < 0.001$, $Z = 2386$, Wilcoxon signed rank test). The Bland-Altman agreement analysis revealed that the mean difference of MTD between

Table 1. Summary Statistics of the Cohort

	Mean	SD									
	71.30	6.60									
			CMP				HFA				
Age	Min	Q1	Median	Q3	Max	Min	Q1	Median	Q3	Max	*P* Value
MTD (dB)	−25	−7	−3	−1.5	1	−27	−8	−4	−2	1	<0.001
Number of Defects	0	4	10	16	49	0	4	11	23	47	<0.001
MTD of Defects (dB)	−26	−19	−13	−9	−5	−29	−18	−12	−8	−6	0.043
S-F Concordance	0	0.97	1	1	1	0	0.75	1	1	1	0.136

Mean and standard deviation (SD) for age and minimum, lower quartile (Q1), median (Q2), upper quartile (Q3), maximum values for the COMPASS (CMP) and the Humphrey Field Analyzer (HFA) visual fields, and structure-function concordance ratios of the cohort.

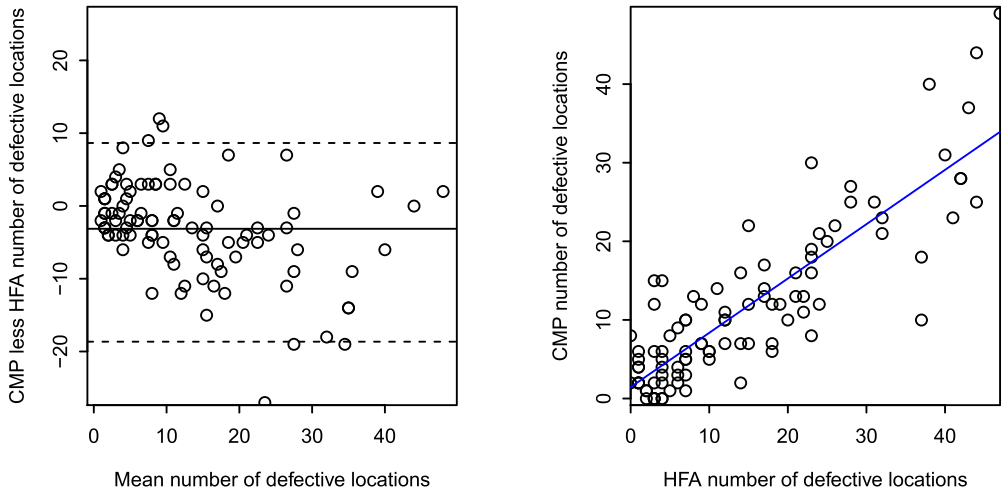

Figure 4. *Left*: Bland-Altman plot of the total number of defective locations per eye measured by the COMPASS (CMP) and Humphrey Field Analyzer (HFA) fields. The *black solid line and dashed lines* indicate the mean difference (−3.14) locations between the total number of defective locations measured by CMP versus HFA, and 95% confidence intervals, respectively. *Right*: The linear regression of the total number of defective locations per eye measured by the CMP versus HFA fields with a slope of 0.69.

CMP and HFA was statistically different from zero (mean ± standard error = 1.13 ± 0.21 dB, t (94) = 5.39, $P < 0.001$; see right panel of Fig.3).

Number of Defective Locations

Overall, out of the total 4940 locations across all eyes, there were 818 agreed defective locations (defective in both CMP and HFA) and 294 locations defective locations in CMP and not HFA, and 592 vice versa. Thus, 22.5% of all locations were classified as defective by CMP compared to 28.5% by HFA. There were significantly fewer defective locations per eye measured by CMP than by HFA ($P < 0.001$, $Z = 1027$, Wilcoxon signed rank test), with a mean difference of −3.14 locations, as shown in the Bland-Altman plot of the data (see the left panel of Fig.4). A linear regression model of the number of defective locations measured by CMP versus HFA yielded a slope of 0.69 (95% CIs = 0.58 to 0.81, $P < 0.001$) and an intercept of 1.45 (95% CIs = −0.04 to 2.94, $P = 0.08$) with a coefficient of determination of 0.74 (see right panel of Fig.4). The slope value of less than 1 returned by the linear model suggests that CMP generally measured fewer defective locations than HFA as the number of defective locations reported by the HFA increased. Given the 95% CIs of the intercept just include 0, we cannot conclusively demonstrate that CMP measures more defective locations than HFA for small scotomata.

The Bland-Altman agreement analysis of the number of defective locations within a scotoma using connected component analysis indicated that, overall, the area of scotomata measured by CMP was smaller than by HFA with the mean difference (± standard error) of −1.76 (± 0.53) defective locations per scotoma being significantly different from 0 (t (86) = −3.34, $P < 0.001$).

Depth of Defective Locations

Summary statistics of TDs of defective locations are presented in Table 2. Across all eyes, the TDs of all defective locations were significantly lower (deeper) for CMP than for HFA ($P = 0.029$, $Z = 744252$, Wilcoxon rank sum test). The TDs of agreed defective locations were significant lower (deeper) for CMP than HFA ($P = 0.002$, $Z = 188474$, Wilcoxon signed rank test). There was no significant difference in TDs for nonagreed defective locations between the two fields ($P = 0.782$).

Maximum Gradient Between Defective Locations and Their Neighbors

Summary statistics of maximum TD gradient of defective locations and their neighbors are presented in Table 3. The maximum gradient is calculated as maximum of the TD values of neighboring locations minus the TD value of a central defective location. Across all eyes, the maximum gradients for all defective locations were significantly greater (steeper) for CMP than for HFA ($P < 0.001$, $Z = 1005945$, Wilcoxon rank sum test). For agreed defective locations, the maximum gradients were significantly greater (steeper) for CMP

Table 2. Summary Statistics of TD for Defective Locations

	CMP					HFA					
	Min	Q1	Median	Q3	Max	Min	Q1	Median	Q3	Max	*P* Value
All defects	−31	−28	−17	−8.4	−5	−35	−27	−13	−8	−6	0.029
Agreed defects	−31	−28	−25	−11	−5	−35	−30	−21	−11	−6	0.002
Nonagreed defects	−29	−12	−8	−6.4	−5.1	−34	−12	−8	−7	−6	0.782

Minimum, lower quartile (Q1), median (Q2), upper quartile (Q3), maximum values of total deviations (TDs) in dB for all, agreed and nonagreed defective locations for COMPASS (CMP) and Humphrey Field Analyzer (HFA), respectively.

Table 3. Summary Statistics of Maximum TD Gradient

	CMP					HFA					
	Min	Q1	Median	Q3	Max	Min	Q1	Median	Q3	Max	*P* Value
All defects	0	5.9	10.6	21	32	0	3	6	12	36	<0.001
Agreed defects	0	5.8	13.4	24	32	0	3	8	16	36	<0.001
Nonagreed defects	0	5.9	8.4	12	30	0	3	5	9	31	<0.001

Minimum, lower quartile (Q1), median (Q2), upper quartile (Q3), maximum values of maximum total deviations (TDs) gradient (in dB) between defective locations and their neighbors for all, agreed and nonagreed defective locations for CMP and HFA, respectively.

than for HFA ($P < 0.001$, $Z = 224113$, Wilcoxon signed rank test). A Bland-Altman plot of the data demonstrates that there are more defective locations with greater maximum gradients measured by CMP than by HFA in the dynamic range of the maximum gradient range (about 5 to 25 dB), whereas the opposite can be observed for the low and high extremes, which is likely due to the floor and ceiling effects given the TD range (−5 to −34 dB) of defective of locations (Fig.5). Overall, the mean difference of maximum gradient measured by CMP and HFA was statistically different from zero (mean ± standard error = 3.86 ± 0.34 dB, $t\,(817) = 11.38$, $P < 0.001$). The maximum TD gradients for nonagreed defective locations were also significantly greater for CMP than for HFA ($P < 0.001$, $Z =$ 90868, Wilcoxon rank sum test).

Structure-Function Concordance

Summary statistics of structure-function concordance ratios are presented in Table 4. There was no significant difference between the overall structure-function concordance ratios between the two devices ($P = 0.136$, $Z = 4582$, Wilcoxon rank sum test). The concordance ratios were high for agreed defective locations and not significantly different between the two devices ($P = 0.664$, $Z = 2774$, Wilcoxon rank sum test). The high structure-function concordance ratios were expected given the agreed defective locations are more likely to represent established scotomata. The concordance ratios were not significantly different for respective nonagreed defective locations between the two devices ($P = 0.189$, $Z = 3708$, Wilcoxon rank sum test).

Discussion

The COMPASS perimeter differs from the widely used HFA in three main aspects: (1) the integrated retinal tracking; (2) an implementation of a testing algorithm that does not exploit spatial correlations; and (3) the ability to accurately register test locations to a retinal image. The findings of the current study are largely consistent with the expected effects of the first two factors. First, the number of defective locations and the spatial extent of scotomata measured by CMP is reduced compared to HFA. Second, CMP tends to measure scotomata as deeper (lower TDs of defective locations) and with sharper edges (higher maximum TD gradients between defective locations and their neighboring locations) compared to HFA in relation to comparison with the instrument specific normative databases. Associated with the reduced number of defective locations, the MTD over all locations is higher for CMP compared to HFA. Structure-function concordance ratios are comparable between CMP and HFA despite the restricted scotoma spatial extent measured by CMP.

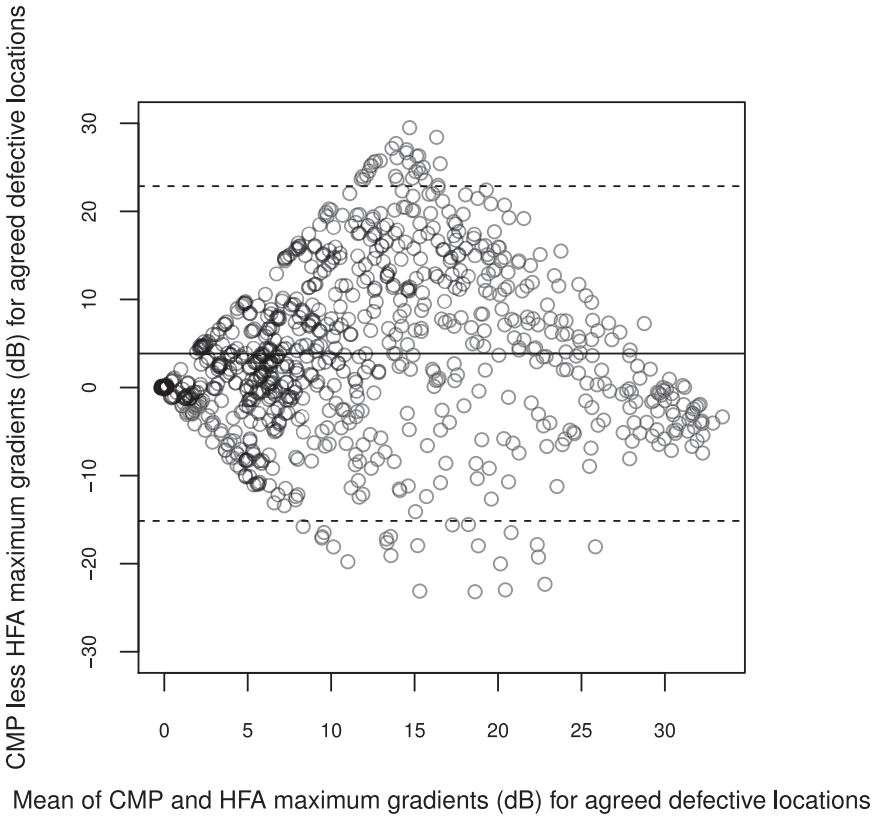

Figure 5. Bland-Altman plot of the maximum gradient (dB) between all agreed defective locations and their surrounding locations for the COMPASS (CMP) versus Humphrey Field Analyzer (HFA) results. The mean difference between the two measurements is 3.3 dB as indicated by the *black solid line* and its 95% confidence intervals are indicated by the *dashed lines*.

Table 4. Summary Statistics of Structure Function Concordance

	CMP					HFA					
	Min	Q1	Median	Q3	Max	Min	Q1	Median	Q3	Max	*P* Value
All defects	0	0.97	1	1	1	0	0.75	1	1	1	0.136
Agreed defects	0.81	1	1	1	1	0.81	1	1	1	1	0.664
Nonagreed defects	0	0.94	1	1	1	0	0.75	1	1	1	0.189

Minimum, lower quartile (Q1), median (Q2), upper quartile (Q3), maximum values of structure-function concordance ratios for all, agreed and nonagreed defective locations for the COMPASS (CMP) and Humphrey Field Analyzer (HFA), respectively.

The differences between CMP and HFA visual field patterns found in the current study are largely consistent with the expected effects of retinal tracking on reducing fixation instability. The *a priori* expectation was that integration of retinal tracking and real-time compensation for eye movements in CMP would result in measurements of scotomata that were deeper, with reduced spatial extent and steeper edges, which we found in our data. We have found that CMP measures fewer defective locations with increasing number of defective locations but could not conclusively demonstrate that CMP measures more locations as defective when the number of defective locations is small (see the right panel of Fig.4). One possible explanation for this finding is that we excluded patients with fewer than two defective locations in both fields. Furthermore, the 6 degrees of spacing of the 24-2 test pattern does not lend itself to precise spatial definition of small scotomata. Indeed, this specific hypothesis may be better assessed with data collected on a more closely spaced test grid, such as 10-2 data.

The results of our dataset in terms of MTD differ slightly to what has been reported in the larger validation study[10] despite identical methods of calculation. The MTDs of CMP (median = −4.37 dB, interquartile range = 8.92 dB) and HFA (median = −4.73 dB,

interquartile range = 9.19 dB) were not significantly different for the glaucoma cohort in the larger validation study whereas MTD for CMP (median = −3 dB, interquartile range = 5.5 dB) was significantly higher than HFA (median = −4 dB, interquartile range = 6 dB) in the current study. The MTD agreement analysis in figure 2 of Montesano et al.,[10] and Figure 3 of the current study suggests that this is likely because there were more patients with small scotomata in our cohort compared to the larger sample in the validation study. It should be noted that the participants included here were a subset of the validation study, where we had additional OCT data not included in the study protocol that enabled determination of the position of the temporal raphe for the individualized structure-function mapping.

We have found comparable structure-function concordance ratios between the CMP and HFA despite the reduced spatial extent of scotomata measured by the CMP. Matching concordance between the CMP and HFA is expected for those locations that were classified as defective on both tests (so called "agreed" locations), because these share almost identical spatial locations between the two devices (whereas the alignment of the CMP and OCT fundus images modified the coordinates of test locations slightly for CMP fields, this only shifted the mapped ONH sectors significantly; i.e. mapped onto sectors of more than 15 degrees apart, for 66 of the total 4940 locations across all eyes). Indeed, this analysis demonstrates that the re-alignment of the images is unlikely to significantly impact on structure-function concordance, at least for visual field defects spaced at 6 degrees and using current mapping techniques. However, if visual field locations could be mapped to ONH with higher precision, or visual field defects were to be measured on a finer spatial scale, or at individualized test locations,[19] then alignment between the CMP and the OCT might yield benefits in some situations. It is also possible that different results might be obtained for mapping to RNFL bundle defects on wide-field imaging, rather than using Cp-RNFL scans.

It is worth noting that the true spatial extent, depth, and sharpness of edges of scotomata is unknown and we make no attempt to draw any conclusion that CMP measured scotomata in these aspects more accurately than HFA. Test accuracy of visual field procedures can really only be assessed via computer simulation where there is a known ground truth.[20,21] However, simulating the effects of eye movements is complex, and may not represent true behavior across the duration of a visual field test. The differences between CMP and HFA visual field patterns measured herein are consistent with the expected effects of improved fixation stability, and hence provide evidence for the advantages of integrating retinal tracking technology into SAP. It is, however, important to acknowledge that the retinal tracking does add some testing time, particularly for people who make a lot of eye movements. We were unable to directly compare the clinically measured test time between the instruments for the 24-2 grid alone because the CMP measured additional macular locations. The original validation study has found similar testing times between the two devices.[10]

We applied a simple inclusion criterion for visual fields (two locations of TD of <−5 dB in either field). We deliberately did not choose a commonly used visual field staging system, because these are mostly derived from empirical databases of HFA data, and therefore incorporate measurement biases that are specific to HFA algorithms and other aspects of the HFA test procedure. Our approach avoids bias in the visual field selection process and allows fair comparison of visual field defect size/depth between the two devices in the subsequent analyses. Importantly, all participants were recruited based on optic disc/Cp-RNFL parameters as per the original study protocol.[10] Visual fields were not an original inclusion criterion of the study to avoid visual field instrument specific bias in the database.

One additional possible contributor to bias in the outcomes is that the instruments do not share a common normative population, hence it is possible that if significant differences exist in the normative database, then different visual field locations may be identified using the total deviation metric. We have no specific reason to suspect this to be the case. An alternate approach could have been for us to compute defects relative to a novel group of healthy controls who had been assessed on both devices (i.e. create our own normative database). We chose not to apply this approach because of the risk of revealing differences that are not a feature of the actual clinical implementations of these devices.

Our study suggests that the novel features of the CMP should enhance its ability to accurately track the position and depth of defects over follow-up visits. Validating whether this prediction is realized is complicated due to the absence of a gold standard definition of visual field progression that is also device independent. Notably, the varying existent criteria for classifying visual field progression have been derived using data collected on clinical implementations of SAP,[22–26] for example, using the HFA. Consequently, the measured visual fields used to devise these criteria have incorporated any biases or algorithmic misrepresentations of visual field defects that exist in those procedures.

In summary, relative to the HFA, the CMP tended to report visual field defects as deeper with sharper

edges and reduced spatial extent. This observation aligns with the predicted effects of several features of the CMP (fundus tracking and absence of spatial neighborhood logic in the test algorithm). Both of these features are designed to improve test accuracy for the measurement of discrete scotomata, although we recognize that accuracy cannot be measured directly in human observers. Accurate measurement of visual field status is important not only for visual field defect detection, but also for tracking of progression.[22–26] The presence of new visual field-testing modalities that show some disagreement with commonly used versions of SAP is potentially important to further our understanding of the true nature of visual field defects and their progression.

Acknowledgments

Disclosure: **P. Liu**, None; **B.N. Nguyen,** None; **A. Turpin**, CenterVue SpA (C, F), Haag-Streit (F), Heidelberg Engineering (F); **A.M. McKendrick**, CenterVue SpA (C, F), Haag-Streit (F), Heidelberg Engineering (F)

References

1. Wall M, Lee EJ, Wanzek RJ, Chong LX, Turpin A. Temporal wedge defects in glaucoma: Structure/function correlation with threshold automated perimetry of the full visual field. *J Glaucoma.* 2020;29:191–197.
2. Vingrys AJ, Demirel S. The effect of fixational loss on perimetric thresholds and reliability. In Mills R., ed. *Perimetry Update 1992/1993*. Amsterdam, UK: Kugler & Ghendini; 1993:521–526.
3. Wyatt HJ, Dul MW, Swanson WH. Variability of visual field measurements is correlated with the gradient of visual sensitivity. *Vision Res.* 2007;47:925–936.
4. Demirel S, Vingrys AJ. Eye movements during perimetry and the effect that fixational instability has on perimetric outcomes. *J Glaucoma.* 1994;3:28–35.
5. Rossetti L, Digiuni M, Rosso A, et al . Compass: Clinical evaluation of a new instrument for the diagnosis of glaucoma. *PLoS One.* 2015;10:e0122157.
6. Fogagnolo P, Modarelli A, Oddone F, Digiuni M, Montesano G, Orzalesi N, Rossetti L. Comparison of compass and Humphrey perimeters in detecting glaucomatous defects. *Eur J Ophthalmol.* 2016;26:598–606.
7. Turpin A, McKendrick AM, Johnson CA, Vingrys AJ. Development of efficient threshold strategies for frequency doubling technology perimetry using computer simulation. *Invest Ophthalmol Vis Sci.* 2002;43:322–331.
8. Bengtsson B, Olsson J, Heijl A, Rootzén H. A new generation of algorithms for computerized threshold perimetry, SITA. *Acta Ophthalmol Scand.* 1997;75:368–375.
9. Olsson J, Rootzén H. An image model for quantal response analysis in perimetry. *Scand J Stat.* 1994;21:375–387.
10. Montesano G, Bryan SR, Crabb DP, et al. A comparison between the COMPASS fundus perimeter and the Humphrey field analyzer. *Ophthalmology.* 2019;126:242–251.
11. Montesano G, Rossetti LM, McKendrick AM, et al. Effect of fundus tracking on structure-function relationship in glaucoma. *Br J Ophthalmol*. 2020;104:1710–1716.
12. Turpin A, McKendrick AM. Improving personalized structure to function mapping from optic nerve head to visual field. *Transl Vis Sci Technol.* 2021;10:19.
13. Bedggood P, Tanabe F, McKendrick AM, Turpin A. Automatic identification of the temporal retinal nerve fiber raphe from macular cube data. *Biomed Opt Express.* 2016;7:4043–4053.
14. Bedggood P, Nguyen B, Lakkis G, Turpin A, McKendrick AM. Orientation of the temporal nerve fiber raphe in healthy and in glaucomatous eyes. *Invest Ophthalmol Vis Sci.* 2017;58:4211–4217.
15. R Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. 2019; Available at: https://www.R-project.org/.
16. Denniss J, Turpin A, McKendrick AM. Individualized structure–function mapping for glaucoma: Practical constraints on map resolution for clinical and research applications. *Invest Ophthalmol Vis Sci.* 2014;55:1985–1993.
17. Lowe DG. Distinctive image features from scale-invariant keypoints. *Int J Comput Vis.* 2004;60:91–110.
18. Bradski G. The OpenCV Library. *Dr. Dobb's Journal of Software Tools.* 2000;120:122–125.
19. Muthusamy V, Turpin A, Walland MJ, Nguyen BN, McKendrick AM. Increasing the spatial resolution of visual field tests without increasing test duration: An evaluation of ARREST. *Transl Vis Sci Technol.* 2020;9:24.
20. Turpin A, McKendrick AM, Johnson CA, Vingrys AJ. Properties of perimetric threshold estimates

from full threshold, zest, and SITA-like strategies, as determined by computer simulation. *Invest Ophthalmol Vis Sci.* 2003;44:4787–4795.
21. Chong LX, Turpin A, McKendrick AM. Assessing the GOANNA visual field algorithm using artificial scotoma generation on human observers. *Transl Vis Sci Technol.* 2016;5:1.
22. De Moraes CG, Liebmann JM, Levin LA. Detection and measurement of clinically meaningful visual field progression in clinical trials for glaucoma. *Prog Retin Eye Res.* 2017;56:107–147.
23. Bengtsson B, Heijl A. A visual field index for calculation of glaucoma rate of progression. *Am J Ophthal.* 2008;145:343–353.
24. Gardiner SK, Crabb DP. Examination of different pointwise linear regression methods for determining visual field progression. *Invest Ophthalmol Vis Sci.* 2002;43:1400–1407.
25. Heijl A, Leske MC, Bengtsson B, Bengtsson B, Hussein M, Group E. Measuring visual field progression in the early manifest glaucoma trial. *Acta Ophthalmol Scand.* 2003;81:286–293.
26. Iester M, Capris E, De Feo F, et al. Agreement to detect glaucomatous visual field progression by using three different methods: A multicentre study. *Br J Ophthalmol.* 2011;95:1276–1283.