

Original Research Article

Perimetry comparisons for Octopus G Top and Dynamic programmes versus Humphrey 24-2 SITA fast and SITA standard programmes

ABSTRACT

Aims: Previous comparisons of Octopus and Humphrey threshold programmes have shown differences, particularly for depth of defects. Recent developments have altered the Octopus background illumination. Thus, we sought to determine the subsequent extent of differences in results from commonly used central static programmes using Octopus and Humphrey perimeters.

Study design: Diagnostic accuracy study.

Place and duration of study: Department of Ophthalmology, Warrington and Halton Hospitals NHS Foundation Trust, between June 2009 and July 2012.

Methodology: We utilised Octopus G-TOP or G-dynamic programmes and Humphrey 24-2 sita fast or standard programmes. Poor reliability was taken as errors >25%. Comparison was made for global indices (mean sensitivity, mean deviation, pattern standard deviation/loss variance) and test duration from each perimeter using unpaired t tests, and Bland Altman plots for correlation of individual differences between perimeters. Clinical agreement for presence/absence of visual field loss was determined by independent assessment of results by two authors (kappa test).

Results: Eighty patients (159 eyes) were recruited. 38 eyes were excluded due to poor reliability. 34 patients (55 eyes) underwent fast static visual field assessment: 17 males and 17 females with a mean age of 63 years (SD 14). 38 patients (64 eyes) underwent standard static visual field assessment: 18 males and 20 females with a mean age of 65 years (SD 13). For both full and fast perimetry, both perimeters showed good agreement for inter observer comparison (K=0.5-0.6). SITA standard perimetry showed a higher mean sensitivity value (+2.5dB±1.0) and SITA fast perimetry showed a higher mean sensitivity value (+3.4dB±1.1) than Octopus values. Bland-Altman analysis showed increasing variability with increasing severity of defect indicating a lack of correlation for moderate to severe visual field results.

Conclusions: Visual field results were comparable for mild defects. Octopus perimetry continues to show greater defect size where there is more extensive visual field loss relating to the different strategies used by both perimeters.

Keywords: *Octopus; Humphrey; G programme; 24-2 programme; threshold perimetry*

1. INTRODUCTION

Visual field assessment is an important clinical tool in the assessment of patients with acute and chronic ocular and/or neurological diseases. Automated static visual field assessment

17 has been undertaken since the 1970's using Humphrey and Octopus perimeters, with both
18 perimeters undergoing extensive hardware and software upgrades over the past two
19 decades.

20 The literature reports a number of comparisons of Humphrey and Octopus automated static
21 perimetry using various programme strategies. Comparisons of fast versus slow threshold
22 strategies have been undertaken to evaluate whether an altered strategy that significantly
23 reduces test time loses detection sensitivity[1-5]. Comparisons of Humphrey SITA fast (SF)
24 with SITA standard (SS) have been shown to have excellent sensitivity and specificity for
25 glaucomatous visual field loss with significantly faster test durations for SF[6]. Scherrer and
26 colleagues[7] compared the results of Octopus tendency oriented perimetry (TOP) versus
27 dynamic strategies and found the faster test strategy also provided accurate and reliable
28 results.

29 Intra-perimeter comparisons have been undertaken for Octopus and Humphrey automated
30 perimeters. One difficulty that has arisen with comparison of these perimeters in the past is
31 the effect of differing background illuminations on threshold values which impacts on mean
32 deviation and pattern standard deviation/loss variance indices. Thus, conversion calculations
33 and tables have been provided in the literature to address this issue[8]. King et al[9]
34 compared the Humphrey 24-2 SF programme to the G1-TOP programme on the Octopus
35 101 perimeter. They found high correlation coefficients between both comparisons for MD
36 and LV/PSD. However, the authors acknowledged that both perimeters operated different
37 background illuminations and thus they converted the Humphrey values for direct
38 comparison to the Octopus results. Following conversion, it was reported that G-TOP
39 estimated MD as greater than Humphrey and LV as less than Humphrey values. The
40 increased difference was particularly seen with increased magnitude of defect. Papp et
41 al[10] compared the G1 programme on Octopus 500 versus Humphrey 30-2 for MD and
42 LV/PSD values and found these to be significantly different. Again, the authors converted the
43 Humphrey values taking into account the differences between the Humphrey and Octopus
44 luminance values. Following conversion a good correlation was achieved.

45 Furthermore, differences between perimeters may also be due to the method in which the
46 test programme is undertaken. A stair step or bracketing process is used for Humphrey
47 central threshold strategies. Initially four points are tested in order to determine the threshold
48 level at those points which are then used as a starting level for neighboring points and so on
49 until the entire visual field has been completed. Points are tested twice where the anticipated
50 response is outside 5 decibels of that expected and the second response obtained is
51 bracketed on the final printout of results[11]. For full threshold testing, a 4-2 decibel step
52 bracketing process is used. The fastpac strategy uses a 3-1 bracketing process which
53 speeds up the test. Octopus TOP is a systematic method that assesses the thresholds of
54 neighbouring points by interpolation[12]. However, boundary edges of visual field defects are
55 slightly rounded as a result and thus the tendency is to give shallower field defects than with
56 Humphrey[12]. For the dynamic test strategy, step sizes adapt to the slope of the frequency
57 of seeing curve using a bracketing process of 2dB to 10dB[12].

58 Langorhorst[13] reported that the dynamic strategy showed more mild defects than on SS.
59 However, with early or moderate defects, Octopus indicated a higher local defect value.
60 Langorhorst postulated that this difference may relate to fluctuations of thresholds in areas of
61 relative local defects in glaucoma patients. Octopus dynamic strategy determines the
62 threshold value after the threshold has been crossed only once. SS is more complicated in
63 its estimation of threshold based on comparison to normative data, intratest responses and
64 neighbouring responses. For more severe field defects, Octopus showed less deep defects
65 than SS which may be due to determination of defect depth by large threshold step sizes.

In recent years, Octopus perimetry has undergone further development culminating in the release of the Octopus 900 perimeter which combines static and kinetic operation modes. The Octopus now operates at the same background illumination as Humphrey perimetry for static visual field assessment. In this study, we sought to determine what differences persist between the Octopus and Humphrey perimeters given the, now, same background illumination for both perimeters and utilising commonly used and previously compared threshold strategies on Humphrey and Octopus static perimetry: SITA-standard (SS) or SITA-fast (SF) and G-dynamic (GD) or G-Tendency Oriented Perimetry (G-TOP) respectively.

2. MATERIAL AND METHODS

A prospective cohort study for diagnostic evaluation was undertaken in accordance with the Tenets of the Declaration of Helsinki. Institutional Research and Development unit approval was obtained.

2.1 Patients

The study comprised of two groups. Group a included patients with diagnoses of glaucoma, ocular hypertension, glaucoma suspects or neurological disease undergoing visual field assessment using SF 24-2 perimetry (Humphrey Instruments, Dublin, CA, USA). Group b included patients with the above diagnoses undergoing visual field assessment using SS 24-2 perimetry.

Participants were not pre-selected for the study but were identified randomly, i.e. notes were taken consecutively from the list waiting for visual field assessment without prior knowledge of patient ability and cognition. A selection bias existed in that the patients recruited to this study had been booked to an out-patient visual field clinic for static perimetry. Thus, there was an assumption that these patients had sufficient ability and cognition to undertake standard automated perimetry.

2.2 Visual field protocol

The 24-2 programme was utilised on Humphrey perimetry as this programme is the standard choice in our clinics for visual field assessment requiring threshold evaluation. This programme consists of 54 stimulus locations offset from the vertical and horizontal meridia and interspaced by 6 degree intervals. The programme assesses the visual field out to 30 degrees nasally and out to 24 degrees superiorly, inferiorly and temporally. The background illumination is set at 31.5asp.

For 24-2 programme full threshold testing, the stimulus intensity is increased in 4 decibel steps until recorded and then decreased to below the threshold level and increased again until recorded, in 2 decibel steps, to confirm the threshold level at that point. The 24-2 fastpac strategy for threshold testing determines threshold sensitivity in 3 decibel steps, thus speeding up the test process. The Swedish Interactive Thresholding Algorithm (SITA) considers many factors in determining what stimuli to present at each point during the test. These factors include age, normative data, detailed characteristics of abnormal and normal tests and patient responses so far in the test. They are combined and weighted into the SITA visual field model which continually produces updated calculations of the threshold at each point.

The G programme was utilised on Octopus 900 perimetry (Haag Streit International, Switzerland) as this programme is advised for general threshold evaluation of the visual

field. This programme consists of 59 stimulus locations offset from the vertical and horizontal meridia and which assess a 30 degree visual field area. The test is a physiology related test pattern that offers a higher density of stimuli in the centre (interspaced at 2.8 degrees). Distribution of test locations follow the nerve fibre bundle layer to allow improved correlation of functional and structural results. Background illumination is set at 31.4asp.

The Octopus TOP method takes into account that the threshold values of neighbouring locations are correlated. It assesses the thresholds of neighbouring points by interpolation. During the test every answer at a particular point is also taken into account in the adjustments of the neighbouring locations[12]. However, abrupt edges are slightly rounded as a result and thus the tendency is to give shallower field defects than with Humphrey[12]. For the dynamic test strategy, step sizes adapt to the slope of the frequency of seeing curve. With increased depth of defect the stimulus luminance step size increases from 2dB to 10dB. The final measured value is calculated as the mean between the two last stimuli[12]. The threshold is only crossed once and all locations are subsequently retested in a second phase.

The study protocol consisted of visual field assessment with both Humphrey perimetry and Octopus perimetry on the same day. The order of testing was randomised as to which of the two assessment types was used first in order to take fatigue effect in to consideration. A short break of 5-10 minutes was allowed between testing on either perimeter. Randomisation was not undertaken using a computer generated table. Patients were assigned to one perimeter or another according to which perimeter was available for use at the time the patient was called for assessment. Inclusion criteria were adult patients attending for visual field assessment, able to sit at the perimeter unaided, able to press the response button, able to understand and follow instructions for performing the test. The exclusion criteria were patients who were unable to sit for the duration of perimetry assessment, follow instructions for performing the test or too ill to complete the full assessment. All patients were provided with full explanation of the purpose of the test and procedure prior to undertaking any perimetry examination. After explanation, patients were given the choice to opt out if they chose (for example if the Humphrey visual field test was performed first and it was of a particularly long duration). However, in practice, very few patients declined to take part in the study.

2.3 Comparison of results and analysis

We analysed the results from each eye of patients (where available). As our study is of diagnostic accuracy and, as visual field loss can affect each eye to quite different extents, we wished to compare the accuracy of each perimeter programme to detect the same type and location of field loss.

Reliability was determined by fixation loss, false positive and false negative responses on Humphrey perimetry and by false positive and false negative responses on Octopus perimetry. Poor reliability was deemed present with fixation losses of >25% and false positive and negative responses of >25%[9].

In order to determine clinical agreement for the two perimeters, qualitative comparisons of the visual fields were carried out by two authors (FR, MW) for inter observer agreement. The criteria for abnormality on Octopus perimetry included:

- MD>2dB and/or sLV value of >6dB
- ≥3 contiguous points <5dB forming a focal defect

157 • A combination of the above two

158 The criteria for abnormality on Humphrey perimetry included:

159 • MD>2dB and/or PSD value of >6dB

160 • ≥ 3 contiguous points at $p < 5\%$ forming a focal defect

161 • Glaucoma hemifield test outside normal limits

162 • A combination of any two of the above

163 Visual field results in both groups were assessed as binary responses for presence or
164 absence of visual field defects (normal or abnormal) and for the latter, were further assessed
165 for type of visual field defect (glaucoma or neurological). One author assessed the results of
166 Octopus perimetry (FR) and the second author assessed the results of Goldmann perimetry
167 (MW). A 'match' between perimeters was subsequently determined as 'Yes/No' for presence
168 of defect.

169 A score of visual field defect size and severity was generated for all results as described by
170 King et al[9]. For defect size, significantly depressed points ($< 1\%$ and $< 0.5\%$) on the total
171 deviation and pattern deviation plots were counted and summed. A percentage score was
172 obtained by dividing the sum by 52 (number of points on Humphrey probability plots) and by
173 59 (number of points on Octopus probability plots). For calculation of the defect severity
174 each abnormal point on the total deviation and pattern deviation plots was assigned a defect
175 score: 1 for $< 5\%$, 2 for $< 2\%$, 3 for $< 1\%$ and 4 for $< 0.5\%$. All point scores were counted and
176 summed. A percentage severity score was obtained by dividing the sum by the maximum
177 score possible (208 on Humphrey perimetry: 52 points $\times 4$; 236 on Octopus perimetry: 59
178 points $\times 4$).

179 The mean deviation on Octopus perimetry is recorded as a positive value but may be
180 positive or negative on Humphrey results. Therefore the Humphrey results were multiplied
181 by -1 to yield a positive value for statistical analysis[9]. Values for global indices (mean
182 deviation (MD), pattern standard deviation (PSD)/loss variance (sLV)) were taken for direct
183 comparison as were values of mean sensitivity (MS). MS was automatically calculated on
184 Octopus results. However, this was manually calculated for Humphrey perimetry as the
185 mean of the sum of all calculated threshold points.

186 A direct comparison of results was made for Humphrey and Octopus perimetry results using
187 the statistical package SPSS version 15 [IBM, USA]. Comparison of similar measures (e.g.
188 global indices, defect size and severity scores, duration of assessment) between Humphrey
189 and Octopus results were evaluated by unpaired t test. A 2-tailed p value of < 0.05 was
190 considered to be statistically significant.

191 Bland-Altman strategy was used to compare the differences between two independent
192 measurements for global indices versus the average of both. When analysing the Bland-
193 Altman results, we expected most of the differences to lie within $\pm 1.96SD$ if normally
194 distributed. Provided the differences within $\pm 1.96SD$ would not be clinically important, we
195 considered that the two results of both methods could be considered as similar. We
196 therefore set a clinical cut-off as to what difference we would regard as clinically acceptable.

197 In order to determine clinical significance of differences between Octopus and Humphrey
198 perimetry results, we determined that a departure of $> 5dB$ on MS would be significant. This

cut-off was based on the static perimetry programming that threshold values exceeding 5dB from age-expected values, are automatically re-tested. We determined our cut-off as 2.5dB for a clinically significant departure for MD and PSD.

Kappa (K) evaluation of agreement was used to correct the proportion of agreement between perimeters due to chance when evaluating intraobserver interpretation of visual field results. K values range from 0 to 1. A K value of 1 was defined as perfect agreement and a value of >0.7 was deemed a strong agreement[14].

3. RESULTS AND DISCUSSION

Eighty patients (159 eyes) were recruited to this study with equal division of 40 patients across two groups. The visual field results were first evaluated for reliability to exclude results that exceeded the cut-off of $>25\%$ errors. Of 159 visual field results, 38 were excluded with 121 visual field results suitable for analysis.

3.1 Octopus G-TOP versus Humphrey SITA fast 24-2 perimetry

Group a consisted of 34 patients (55 eyes) undergoing Humphrey 24-2 SF versus Octopus G-TOP perimetry. There were 17 males and 17 females with a mean age of 63 years (SD 14: range 29-89).

3.1.1 Comparison of visual field indices

The mean values for MS, MD, sLV/PSD and duration of test are shown in table 1. There is a significant difference between G-TOP and SF results for MS with a difference between means of -3.4 ± 1.1 , $p=0.004$ (unpaired t test) and mean higher values for SF results. There is no significant difference between G-TOP and SF results for MD ($p=0.516$), sLV/PSD ($p=0.579$) and test duration ($p=0.107$) with a difference between means of 0.71 ± 1.1 , -0.34 ± 0.61 and 0.04 ± 0.22 respectively.

Bland-Altman analysis (figure 1) shows a wide distribution of differences for MS differences for G-TOP and SF results with variability across comparisons. Despite no significant differences for MD, sLV/PSD and test duration, Bland-Altman analysis shows increasing variability in comparisons with increasing average values.

3.1.2 Comparison of defect size and severity

The mean values for defect size and defect severity for total deviation and pattern deviation plots are shown in table 2. There is a significant difference between G-TOP and SF results for defect size on total deviation plots with a difference between means of -25.16 ± 5.87 , $p=0.001$ (unpaired t test) and mean higher values for G-TOP results. There is also a significant difference between G-TOP and SF results for defect size on pattern deviation plots with a difference between means of -8.07 ± 3.93 , $p=0.043$ (unpaired t test) with mean higher values for G-TOP results. There is no significant difference between G-TOP and SF results for defect severity on either total deviation ($p=0.06$) and pattern deviation ($p=0.97$) plots with a difference between means of -10.85 ± 8.69 and -0.11 ± 3.84 respectively.

238
239
240
241
242

243
244
245
246

Table 1 G-TOP and SITA-fast 24-2 mean values: visual field indices

	G-TOP		SF 24-2	
	Mean	SD	Mean	SD
MS vs MS (dB)	21.73	6.10	25.09	5.80
MD vs MD (dB)	5.27	5.54	4.57	5.81
sLV vs PSD	3.96	2.50	4.30	3.80
Test duration (minutes)	3.35	1.29	3.71	1.02

247
248
249
250

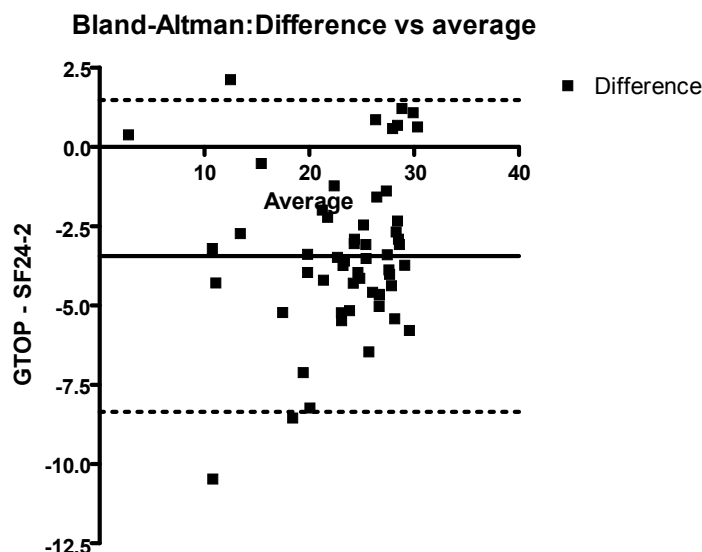
MS: mean sensitivity; dB: decibel; MD: mean deviation; sLV: standard loss variance
PSD: pattern standard deviation; SD: standard deviation

251

Figure 1 Bland-Altman analysis of G-TOP versus SF 24-2 results

252

A: Mean sensitivity differences

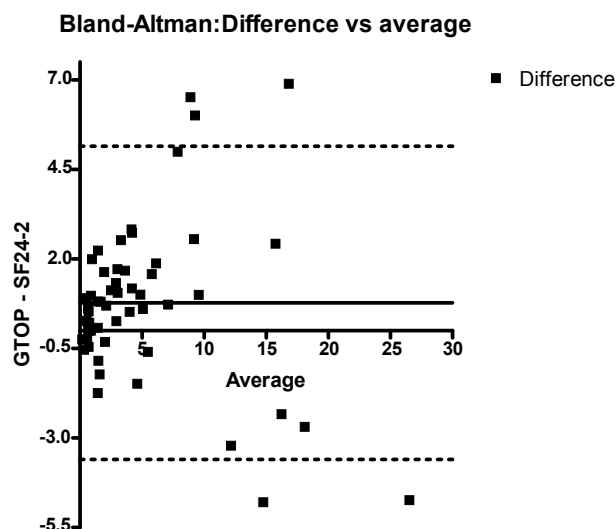


253
254
255
256
257
258

259
260
261
262
263

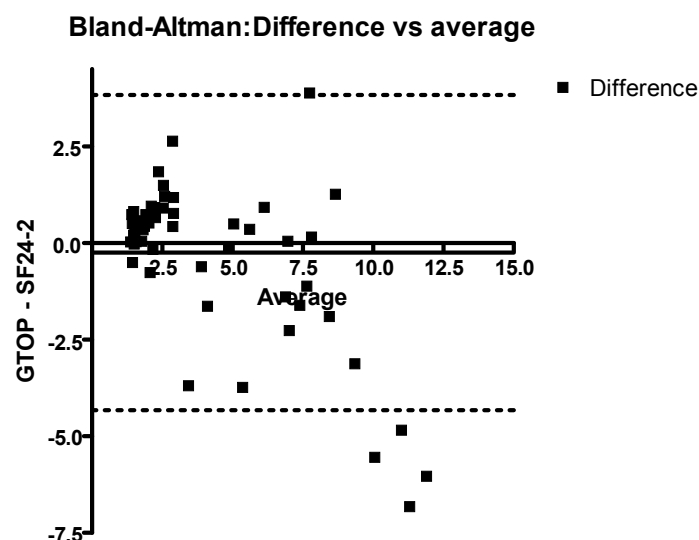
The solid line represents the mean bias of -3.363 with a higher mean MS from SF 24-2 results compared to G-TOP results. The dotted lines represent limits of agreement (LoA) $\pm 1.96SD$ (-8.28 to 1.55). There is no consistent bias of one method versus the other with variability noted across all comparisons. The LoA are within $\pm 5dB$ from the mean bias suggesting that the differences are not clinically significant.

264 B: Mean deviation differences
265



266
267
268 The solid line represents the mean bias of +0.706 with a lower mean MD from SF 24-2
269 results compared to G-TOP results. The dotted lines represent $\pm 1.96SD$ (-3.68 to 5.09). At
270 low values there is good agreement between both measures but greater variability at
271 increasing differences between both measures with differences exceeding a cut-off for
272 clinical significance of $\pm 2.5dB$.
273

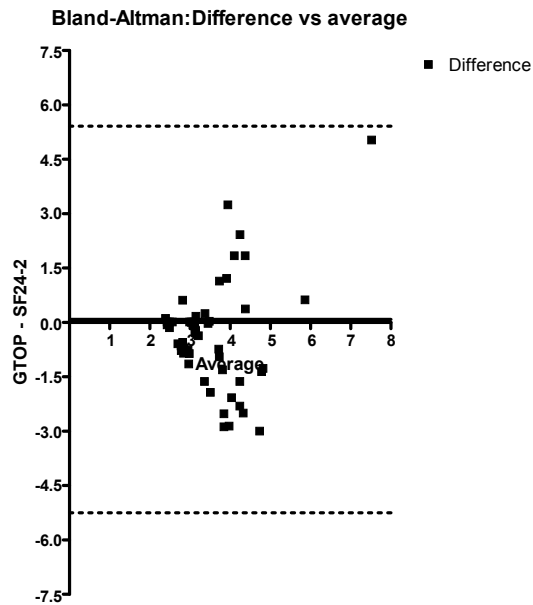
274 C: Loss variance versus Pattern standard deviation differences



275
276 The solid line represents the mean bias of -0.341 with a higher mean PSD from SF 24-2
277 results compared to the mean sLV on G-TOP results. The dotted lines represent $\pm 1.96SD$ (-4.38 to 3.69). There is a proportional response with an increase in variability as
278 the difference from average comparison of sLV and PSD values increases. At low values,
279 differences are within a cut-off for clinical significance of $\pm 2.5dB$. However, at higher values,
280 differences exceed the cut-off.
281

282
283
284

D: Test duration differences



285
286

287 The solid line represents the mean bias of -0.362 with a higher mean test duration for SF 24-
288 2 compared to G-TOP. The dotted lines represent LoA $\pm 1.96SD$ (-3.23 to 2.50). Variability
289 increases with longer test duration averages.

290 **Table 2 G-TOP and SITA-fast 24-2 mean values: Defect size and severity**

	G-TOP		SF 24-2	
	Mean	SD	Mean	SD
Total deviation plot Defect size	47.03	4.40	21.87	3.89
Total deviation plot Defect severity	36.52	4.24	25.67	3.80
Pattern deviation plot Defect size	24.39	2.85	16.33	2.71
Pattern deviation plot Defect severity	19.17	2.79	19.06	2.63

291

292

293 **3.1.3 Comparison of inter observer agreement**

294 Intra-observer evaluation of presence or absence of visual field defect was found to be
 295 moderate ($K=0.5$). Six G-TOP results were classed as abnormal by examiner 1 (FR) where
 296 corresponding SF results were classed as normal by examiner 2 (MW). Five SF results were
 297 classed as abnormal by examiner 2 where corresponding G-TOP results were classified as
 298 normal by examiner 1. The remaining 44 sets of results were equally classed as normal or
 299 abnormal by both examiners.

300 **3.2 Octopus G-dynamic versus Humphrey SITA standard 24-2 perimetry**

301 Group b consisted of 38 patients (64 eyes) undergoing Humphrey 24-2 SS versus Octopus
 302 GD perimetry. There were 18 males and 20 females with a mean age of 65 years (SD 13:
 303 range 23-81).

304 **3.2.1 Comparison of visual field indices**

305 The mean values for MS, MD, sLV/PSD and duration of test are shown in table 3. There is a
 306 significant difference between GD and SS results for MS with a difference between means of
 307 -2.5 ± 1.0 , $p=0.015$ (unpaired t test) and mean higher values for SS results. There is also a
 308 significant difference for test duration with a difference between means of $+0.89 \pm 0.26$,
 309 $p=0.001$ (unpaired t test) and mean longer durations for GD results. There is no significant
 310 difference between GD and SS results for MD ($p=0.828$) and sLV/PSD ($p=0.648$) with a
 311 difference between means of $+0.21 \pm 0.97$ and $+0.26 \pm 0.57$ respectively.

312 Bland-Altman analysis (figure 2) shows no correlation between MS differences for GD and
 313 SS results with variability across comparisons but with increasing variability for test duration
 314 with longer test times. Despite no significant differences for MD and sLV/PSD, Bland-Altman
 315 analysis shows increasing proportional variability in comparisons with increasing average
 316 values.

317 **Table 3 G-dynamic and SITA-standard 24-2 mean values**

318

	G-DYN		SS 24-2	
	Mean	SD	Mean	SD
MS vs MS (dB)	22.63	5.16	25.14	6.35
MD vs MD (dB)	4.42	4.59	4.21	6.29
sLV vs PSD	4.24	2.62	3.98	3.69
Test duration (minutes)	6.45	1.82	5.55	1.04

319
 320 MS: mean sensitivity; dB: decibel; MD: mean deviation; sLV: standard loss variance
 321 PSD: pattern standard deviation; SD: standard deviation
 322

323 **3.2.2 Comparison of defect size and severity**

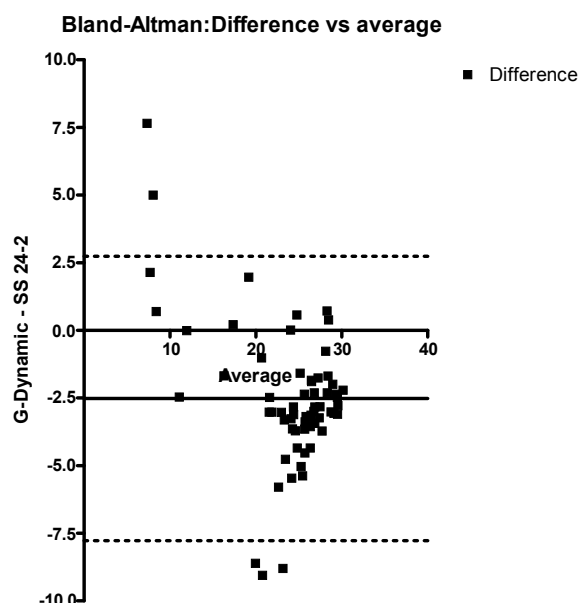
324 The mean values for defect size and defect severity for total deviation and pattern deviation
 325 plots are shown in table 4. There is a significant difference between GD and SS results for

defect size on total deviation plots with a difference between means of -18.35 ± 4.87 , $p=0.001$ (unpaired t test) with mean higher values for GD results. There is no significant difference between GD and SS results for defect size on pattern deviation plot ($p=0.13$: difference between means of -5.41 ± 3.55) or for defect severity on either total deviation ($p=0.06$) and pattern deviation ($p=0.97$) plots with a difference between means of -10.85 ± 8.69 and -0.11 ± 3.84 respectively.

Figure 2 Bland-Altman analysis of G-Dynamic versus SS 24-2 results

A: Mean sensitivity differences

334



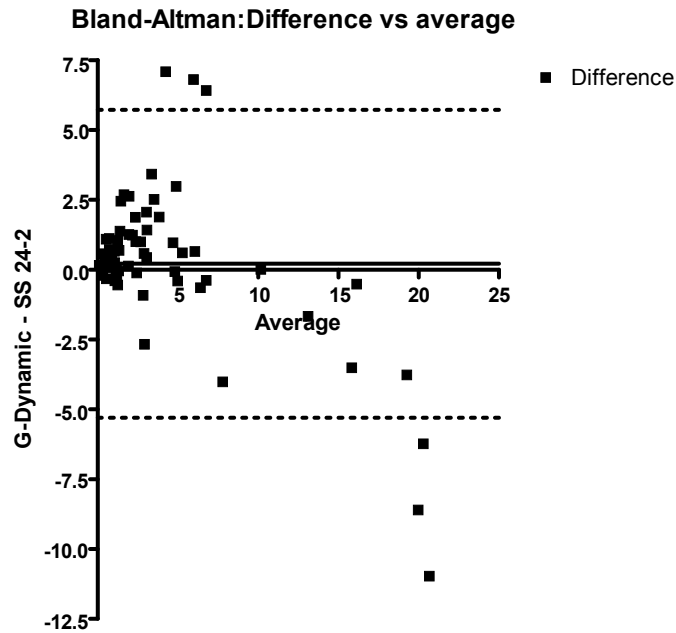
335
336

The solid line represents the mean bias of -2.519 with a higher mean MS from SS 24-2 results compared to G-Dyn results. The dotted lines represent limits of agreement (LoA) $\pm 1.96SD$ (-7.77 to 2.73). There is no consistent bias between mean sensitivities with variability noted across all comparisons. The LoA are within $\pm 5dB$ from the mean bias suggesting that the differences are not clinically significant.

342
343
344
345
346
347
348
349
350
351
352
353
354
355
356

357
358
359

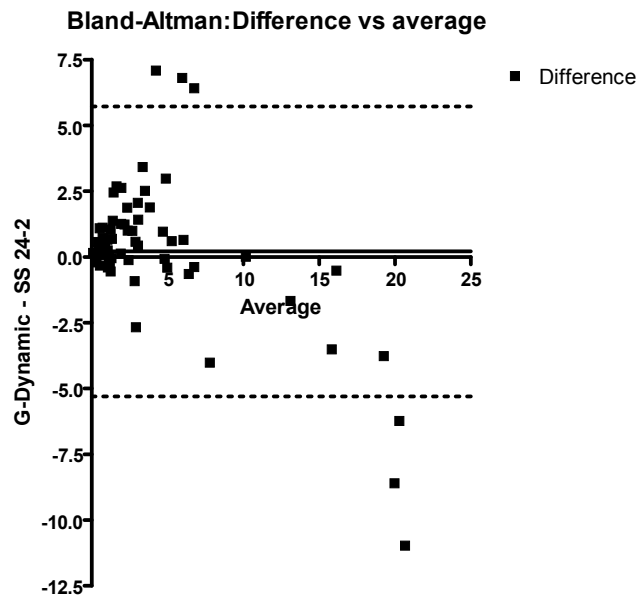
B: Mean deviation differences



360
361

362 The solid line represents the mean bias of +0.211 with a lower mean MD from SS 24-2
363 results compared to G-Dyn results. The dotted lines represent LoA $\pm 1.96SD$ (-5.29 to 5.72).
364 At low values there is good agreement between both measures but greater variability at
365 increasing differences between both measures with differences exceeding a cut-off for
366 clinical significance of $\pm 2.5dB$.

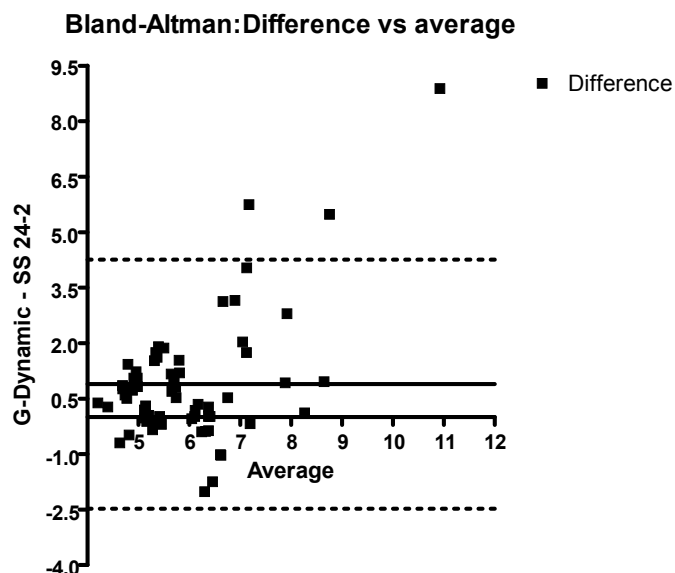
367 C: Loss variance versus Pattern standard deviation differences



368

369 The solid line represents the mean bias of +0.259 with a lower mean PSD from SS 24-2
 370 results compared to the mean sLV on G-Dyn results. The dotted lines represent LoA
 371 $\pm 1.96SD$ (-3.22 to 3.74). There is a proportional response with an increase in variability as
 372 the difference from average comparison of sLV and PSD values increases. At low and high
 373 values, most differences are within a cut-off for clinical significance of $\pm 2.5dB$. However,
 374 some differences exceed the cut-off.

375 D: Test duration differences



376
 377

378 The solid line represents the mean bias of +0.892 with a lower mean test duration for SS 24-
 379 2 compared to G-Dyn. The dotted lines represent LoA $\pm 1.96SD$ (-2.48 to 4.26). Variability
 380 increases with longer test duration averages.

381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400

401
402
403
404
405

Table 4 G-Dyn and SITA-standard 24-2 mean values: Defect size and severity

	G-TOP		SF 24-2	
	Mean	SD	Mean	SD
Total deviation plot Defect size	35.8	3.54	17.46	3.35
Total deviation plot Defect severity	26.56	3.27	20.89	3.32
Pattern deviation plot Defect size	20.44	2.22	15.03	2.77
Pattern deviation plot Defect severity	15.76	2.13	17.36	2.72

406

407 **3.2.3 Comparison of inter observer agreement**

408 Intra-observer evaluation of presence or absence of visual field defect was found to be
409 substantial (K=0.6). Five GD results were classed as abnormal by examiner 1 (FR) where
410 corresponding SS results were classed as normal by examiner 2 (MW). Six SS results were
411 classed as abnormal by examiner 2 where corresponding GD results were classified as
412 normal by examiner 1. The remaining 53 sets of results were equally classed as normal or
413 abnormal by both examiners.

414

415 This study has aimed to compare two different strategies commonly used for general
416 assessment of static perimetry of the central visual field: the 24-2 programme versus the G
417 programme. Our comparison of SF and G-TOP results and SS and GD results found a
418 significant difference for MS with Humphrey perimetry values being on average 3.4dB higher
419 for SF and 2.5dB higher for SS. Our findings are comparable to previous studies reporting
420 differences in mean sensitivity. Humphrey perimetry has a higher maximum stimulus
421 luminance (range up to 51dB) compared to Octopus perimetry (range up to 47dB) and as a
422 result the normal values are higher by 3-4dB compared to Octopus perimetry[12]. Table 5
423 shows a comparison of Octopus 900 and Humphrey parameters.

424 MS is automatically calculated on Octopus 900 results whereas it is not provided as a value
425 on Humphrey perimetry. Therefore, we sought to evaluate the values that would be
426 immediately available for comparison including the MD and PSD/sLV. A number of
427 explanations have previously been suggested for the differences found for MD and PSD/sLV
428 at higher value differences which include the number, location and weighting of test locations

within the different programmes, the strategy used for each programme (bracketing versus tendency-oriented) and differences in background illumination for each perimeter. The latter explanation can be removed for this study as both Humphrey and Octopus perimeters operated at the same background illumination. Thus the differences reported in past literature have been minimized. We did not have to convert our Humphrey values for direct comparison and found no significant difference between MD and PSD/sLV on initial statistical analysis comparing overall MD and PSD/sLV values.

Table 5 Octopus versus Humphrey perimeter parameters

	Octopus perimetry	Humphrey perimetry
Background luminance	31.4asb (10cd/m ²)	31.5asb (10cd/m ²)
Luminance for 0dB	4000asb	10000asb
Stimulus exposure time	100ms	200ms
Full threshold bracketing strategy	4-2-1	4-2
dB range	0 to 47	0 to 51

However, Bland Altman evaluation showed variability in differences. MD showed increasing variability with greater magnitudes of difference for both SF versus G-TOP and SS versus GD. At lower values, the differences were within 2.5dB which we considered a clinically acceptable difference. At higher values, the differences were outside the cut-off. This was further shown when comparing results for the calculated defect size from total deviation plots. Mean values for Octopus results were significantly higher for both G-TOP and GD. Thus for normal visual field results or results with slight visual field impairment, results were similar. However, for higher values, results from both perimeters were considerably different.

The same was found for PSD/sLV measurements in that a proportional agreement was found on Bland Altman analysis. At low values, differences largely fell within the clinically acceptable difference of 2.5dB but not with higher value differences. This was further shown for G-TOP as there was a significant difference for defect size on pattern deviation plots with mean higher values for G-TOP.

It is important to consider the alternative explanations to explain the differences found between both perimeters including the number, location and weighting of test locations within the different programmes plus the strategy used for each programme (bracketing versus tendency-oriented).

The 24-2 and G programmes are both assessments within the central 30 degree visual field but the 24-2 programme only extends to 30 degrees nasally and extends to 24 degrees for inferior, superior and temporal visual field. Conversely, the G programme extends out to 30 degrees in all areas of the visual field. The more extensive peripheral assessment of visual field by the G programme may partly explain the significant difference between programmes in the presence of more severe visual field loss as the G programme may pick up larger defects in more advanced ocular disease.

463 The 24-2 and G programmes also have different patterns of stimulus location. The 24-2
464 programme has equal spacing between stimuli which are located 6 degrees apart. The G
465 programme has a greater concentration of stimuli in the central 10 degrees of visual field
466 and less towards the outer 30 degrees. These stimuli follow arrangements to mirror nerve
467 fibre bundle pathways. The Octopus mean defect is the average of the total deviation values
468 in the field test. The more important and less variable deviations are near the centre of the
469 field and are weighted more than those at the periphery[12]. Humphrey MD is the overall
470 departure of the average deviation of the visual field result from that expected of a normal
471 field of the same age group. It is derived from the results of the total deviation plot and is the
472 weighted mean of all numbers in the plot which are the deviation from the average normal
473 values for age[11]. Each deviation from normal is weighted according to the variance of
474 normal values at that location. Therefore points with low variance closer to fixation are more
475 important than points with higher variance more peripherally.

476 The Octopus loss variance is the local heterogeneity of a defect. It is small in fields with
477 generalized damage and increased with the number of depth of localized scotomas[12]. The
478 Humphrey PSD is a global index that measures the degree to which the shape of the
479 patient's field departs from normal age-corrected reference field model. It is determined by
480 variation from the normal hill of vision. A normal hill of vision will give a low PSD as will a
481 generalized elevation or depression of sensitivity across the entire visual field. Areas of focal
482 loss of vision give rise to higher PSD values[11]. It is possible that the differences in
483 weighting of these calculations give rise to the variances seen with greater magnitudes of
484 defect. Funkhouser & Fankhauser[15] previously reported that unweighted and fluctuation-
485 weighted indices measured by Octopus perimetry show negligible differences for the G1
486 programme.

487 Humphrey 24-2 SS and SF programmes use, respectively, a full threshold 4-2 decibel step
488 bracketing process and a fastpac threshold 3-1 bracketing process. Octopus dynamic test
489 strategy also uses a bracketing process in which step sizes ranging from 2dB to 10dB adapt
490 to the slope of the frequency of seeing curve[12]. However, Octopus TOP does not use a
491 bracketing process but assesses the thresholds of neighbouring points by interpolation, i.e. it
492 uses the subject's response at each point to estimate the sensitivity at that point but also to
493 contribute to the sensitivity estimation at neighbouring points[12].

494 Langorhorst[13] reported that the dynamic strategy showed more mild defects than on SS.
495 However, with early or moderate defects, Octopus indicated a higher local defect value.
496 Langorhorst postulated that this difference may relate to fluctuations of thresholds in areas of
497 relative local defects in glaucoma patients. Octopus dynamic strategy determines the
498 threshold value after the threshold has been crossed only once. SS is more complicated in
499 its estimation of threshold based on comparison to normative data, intratest responses and
500 neighbouring responses. For more severe field defects, Octopus showed less deep defects
501 than SS which may be due to determination of defect depth by large threshold step sizes.
502 Similar differences have been noted when utilizing the TOP strategy with under-estimation of
503 sLV in comparison to traditional bracketing processes[1]. Anderson reported poor spatial
504 resolution characteristics for TOP with the depth of defect typically over-estimated and with
505 blurring of defect boundaries due to the use of linear interpolation between test locations[16].

506 In addition to our comparisons of global indices and test duration, we evaluated inter-
507 observer variances. Comparison of inter-observer differences was found to be moderate on
508 Kappa analysis for G-TOP versus SF and substantial agreement for GD and SS. A similar
509 number of normal/abnormal results were found for Octopus versus Humphrey perimetry
510 results as determined independently by two of the authors. Using a similar method, King and
511 colleagues[9] reported a very high level of clinical agreement between two investigators in

judging their results as normal or glaucoma. They suggested that this high agreement indicated the identical qualitative similarity between assessments in a clinical environment.

Duration of perimetry testing is an important aspect of achieving accurate and reliable visual field results. An excessive period of test duration can give rise to increased variability and fatigue-related defects. G-TOP was found to have a significantly quicker test time in the study by King et al[9] whereas there was no significant difference in duration of test in our study for SF versus G-TOP. A further study compared SS 30-2 with Octopus 101 32 dynamic programme[13]. The test time was comparable but the Octopus strategy took less time for more severe defects. This was also found in our study for GD versus SS. It is important to consider the trade-off between faster test duration and accuracy and reliability of the visual field result.

4. CONCLUSION

Following comparison for detection of global or localised visual field deficit and quantitative evaluations of global indices and defect size and severity, moderate to substantial inter-observer agreement was found, and visual field results showed similar results for normal or mildly affected visual fields. Although there were no significant differences on overall analysis of MD and PSD/sLV when sub-analysed for defect severity, greater deficits were recorded for Octopus perimetry in more severely affected visual fields. Thus, despite both perimeters operating the same background illuminations, differences persist and particularly for moderate to severe visual field loss. This relates to test programmes (locations and density of stimuli) and test strategy (conventional 4-2 and 3-1 bracketing steps versus variable 2 to 10 bracketing steps or interpolation process).

ACKNOWLEDGEMENTS

Financial Disclosure: Haag Streit has provided travel expenses for conference attendance to FR and the loan of the Octopus 900 perimeter for the conduct of this research study. The sponsor or funding organization had no role in the design or conduct of this research.

544 **COMPETING INTERESTS**

545

546 The authors have declared that no competing interests exist.

547

548

549 **AUTHORS' CONTRIBUTIONS**

550

551 Author FR designed the study, performed the statistical analysis, wrote the protocol, and
552 wrote the drafts of the manuscript. Author MW provided oversight for the study and
553 contributed to the drafts of the manuscript. Author SS managed the recruitment and data
554 collection for the study. All authors read and approved the final manuscript.

555

556 **CONSENT**

557

558 All authors declare that informed consent was obtained from the patient prior to taking part in
559 this study.

560

561

562 **ETHICAL APPROVAL**

563

564 All authors hereby declare that all experiments have been examined and approved by the
565 appropriate ethics committee and have therefore been performed in accordance with the
566 ethical standards laid down in the 1964 Declaration of Helsinki.

567

568 **REFERENCES**

569

570 1. Maeda H, Nakaura M, Nagi. New perimetric threshold test algorithm with dynamic
571 strategy and tendency oriented perimetry (TOP) in glaucomatous eyes. *Eye*. 2000; 219: 373-
572 378

573 2. Gonzalez-Hernandez M, Morales J, Azuara-Bianco A, Sanchez JG, de la Rosa MG.
574 Comparison of diagnostic ability between a fast strategy, tendency-oriented perimetry, and
575 the standard bracketing strategy. *Ophthalmologica*. 2005; 219:373-8.

576 3. Artes PH, Iwase A, Ohno Y, Kitazawa Y, Chauhan BC. Properties of perimetric
577 threshold estimates from Full Threshold, SITA Standard, and SITA Fast strategies.
578 *Investigative Ophthalmology Vision Science*. 2002; 43:2654-9.

579 4. Wadood AC, Azuara-Bianco A, Aspinall P, Taguri A, King AJ. Sensitivity and
580 specificity of frequency-doubling technology, tendency-oriented perimetry, and Humphrey
581 Swedish interactive threshold algorithm-fast perimetry in a glaucoma practice. *American
582 Journal of Ophthalmology*. 2002; 133:327-32.

583 5. Pierre-Filho Pde T, Schimiti RB, de Vasconcellos JP, Costa VP. Sensitivity and
584 specificity of frequency-doubling technology, tendency-oriented perimetry, SITA Standard
585 and SITA Fast perimetry in perimetrically inexperienced individuals. *Acta Ophthalmologica
586 Scandinavica*. 2006; 84:345-50.

587 6. Budenz DL, Rhee P, Feuer WJ, McSoley J, Johnson CA, Anderson DR. Sensitivity
588 and specificity of the Swedish interactive threshold algorithm for glaucomatous visual field
589 defects. *Ophthalmology*. 2002; 109:1052-8.

- 590 7. Scherrer M, Fleischhauer JC, Helbig H, Johann Auf der Heide K, Sutter FK.
591 Comparison of tendency-oriented perimetry and dynamic strategy in octopus perimetry as a
592 screening tool in a clinical setting: a prospective study. *Klin Monbl Augenheilkd*. 2007;
593 224:252-4.
- 594 8. Monhart M, Zulauf M. Conversion of the Octopus visual-field index LV to the Zeiss-
595 Humphrey field analyzer index PSD for the glaucoma risk calculator S.T.A.R. *Ophta*. 2006;
596 1: 15-16
- 597 9. King AJ, Taguri A, Wadood AC, Azuara-Blanco A. Comparison of two fast
598 strategies, SITA Fast and TOP, for the assessment of visual fields in glaucoma patients.
599 *Graefes Archives Clinical Experimental Ophthalmology*. 2002; 240:481-7.
- 600 10. Papp A, Kis K, Németh J. Conversion formulas between automated-perimetry
601 indexes as measured by two different types of instrument. *Ophthalmologica*. 2001; 215:87-
602 90.
- 603 11. Rowe FJ. *Visual fields via the visual pathway*. Wiley-Blackwell Publishing. 2004
- 604 12. Weijland A, Fankhauser F, Bebie H, Flammer J. *Automated perimetry. Visual field*
605 *digest*. 2004, 5th edition. Haag Streit AG.
- 606 13. Langerhorst CT, Carenini LL, Bakker D, van den Berg TJTP, de Bie-Raakman MAC.
607 Comparison of SITA and dynamic strategies with the same examination grid. *Perimetry*
608 *Update*, pp. 17–24. *Proceedings XIIIth International Perimetric Society Meeting, Gardone*
609 *Riviera (BS), Italy, September 6–9, 1998*. Eds: WallM, Wild JM. 1999 Kugler Publications,
610 The Hague, The Netherlands
- 611 14. Nayak BK, Hazra A. How to choose the right statistical test? *Indian Journal of*
612 *Ophthalmology*. 2011; 59: 85-6
- 613 15. Funkhouser A, Fankhauser F. The effects of weighting the "mean defect" visual field
614 index according to threshold variability in the central and midperipheral visual field. *Graefes*
615 *Archives Clinical Experimental Ophthalmology*. 1991; 229:228-31.
- 616 16. Anderson AJ. Spatial resolution of the tendency-oriented perimetry algorithm.
617 *Investigative Ophthalmology and Visual Science*. 2003; 44: 1962-68
618