Journal of Glaucoma Publish Ahead of Print
DOI:10.1097/IJG.0000000000001856

Visual field endpoints based on subgroups of points may be useful in glaucoma clinical trials – a study with the Humphrey Field Analyzer and the Compass perimeter

Yaniv Barkana, MD, [1,2] Ari Leshno, MD, [1] Ori Stern, MD, [1] Reut Singer, MD, [1] Hermann Russ, MD, PhD,[2] Francesco Oddone, MD, PhD, [3] Paolo Lanzetta, MD,[4] Andrea Perdicchi, MD,[5] Chris A. Johnson, PhD,[6] David F. Garway-Heath, MD, FRCOphth,[7] Luca M. Rossetti, MD, FRCOphth,[8] Alon Skaat, MD[1]

[1]The Sam Rothberg Glaucoma Center, Goldschleger Eye Institute, Sheba Medical Center, affiliated to Sackler Faculty of Medicine Tel Aviv University, Tel Aviv, Israel

[2]Galimedix Therapeutics, 3704 Calvend Lane, Kensington, MD 20895

[3]Fondazione Bietti -IRCCS, Rome, Italy.

[4]Department of Medical and Biological Sciences, Ophthalmology Unit, University of Udine, Udine, Italy.

[5]Ophthalmology Unit, St. Andrea Hospital, NESMOS Department, University of Rome "Sapienza," Rome, Italy.

[6]Department of Ophthalmology and Visual Sciences, University of Iowa Hospitals and Clinics, Iowa City, Iowa.

[7]NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom.

[8]University of Milan e ASST Santi Paolo e Carlo, Milan, Italy.

Corresponding author: Yaniv Barkana, MD

The Sam Rothberg Glaucoma Center, Goldschleger Eye Institute

Sheba Medical Center, Tel Hashomer, Israel.

Tel: +972-3-530-2874

Fax: +972-3-530-2872

Email: yanivbarkana@gmail.com

Financial Support: None

The author(s) have made the following disclosure(s): F.O., A.M.M., P.L., C.A.J., D.F.G-.H., and L.M.R.: Consultants CenterVue, Padua, Italy.

Running head: High sensitivity glaucoma visual field new endpoints

Precis: Visual field endpoints based on average deviation of specific subsets of points rather than all points may offer a more homogenous dataset without necessarily worsening test-retest variability and so may be useful in clinical trials.

Disclosure: The authors declare no conflict of interest.

Abstract

Purpose: To characterize outcome measures encompassing particular subsets of visual field points and compare them as obtained with Humphrey (HVF) and Compass perimeters.

Methods: 30 patients with imaging-based glaucomatous neuropathy performed a pair of 24-2 tests with each of 2 perimeters. Non-weighted mean deviation (MD) was calculated for the whole field and separate vertical hemifields, and again after censoring of points with low sensitivity (MDc) and subsequently including only "abnormal" points with total deviation probability of <5% (MDc5%) or <2% (MDc2%). Test-retest variability was assessed using Bland-Altman 95% limits of agreement (95%LoA).

Results: For the whole field, using HVF, MD was -7.5 ± 6.9 dB, MDc -3.6 ± 2.8 dB, MDc5% -6.4 ± 1.7 dB and MDc2% -7.3 ± 1.5 dB. With Compass MD was -7.5 ± 6.6, MDc -2.9 ± 1.7 dB, MDc5% -6.3 ± 1.5, and MDC2% -7.9 ± 1.6. The respective 95% LoA were 5.5, 5.3, 4.6 and 5.6 with HVF, and 4.8, 3.7, 7.1 and 7.1 with Compass. The respective number of eligible points were 52, 42 ± 12, 20 ± 11 and 15 ± 9 with HVF, and 52, 41.2 ± 12.6, 10 ± 7 and 7 ± 5 with Compass. With both machines, standard deviation (SD) and 95%LoA increased in hemifields compared to the total field, but this increase was mitigated after censoring.

Conclusions: Restricting analysis to particular subsets of points of interest in the visual field after censoring points with low sensitivity, as compared with using the familiar total field mean deviation, can provide outcome measures with a broader range of mean deviation, a markedly reduced SD and therefore more homogenous dataset, without necessarily worsening test-retest variability.

Keywords – endpoints, outcome measures, glaucoma, field, sensitivity

Introduction

Glaucoma remains a leading cause of blindness. Although effective treatment is available that lowers intraocular pressure (IOP), treatment that directly targets the neurodegenerative process is not. In this respect the situation is similar to age-related macular degeneration, and other age-related neurodegeneration of the CNS such as Alzheimer's disease. There is active research for so-called neuroprotection treatment modalities that seeks to prevent death of retinal ganglion cells (RGCs).[1,2] This overlaps with "neuroenhancing" treatment that seeks to improve the function of RGCs that are sick and dysfunctional yet not irreversibly damaged; this concept is supported by previous observations of improved visual field performance after lowering of IOP[3,4], and such functional improvement treatment is approved for CNS diseases such as Alzheimer's and Parkinson's diseases[5]. For this ongoing research effort to be successful, there is need for endpoints that can maximize the likelihood of success of a clinical trial to show effectiveness of novel treatment.

Subjective visual field (VF) examination remains the mainstay of functional testing in glaucoma. However, the slow nature of glaucoma progression coupled with the suboptimal repeatability of field test results necessitate multiple testing over relatively long duration of many subjects in a typical trial. Strategies to minimize these requirements and make glaucoma trials more feasible have been proposed, including clustering of field tests rather than equal spacing[6], and looking for the *rate* of progression rather than yes-or-no pointwise criteria.[7] One group of researchers has proposed that when a visual field location has a sensitivity less than 15-19 dB, its repeatability becomes nearly random, and so does not contribute to the effective assessment of change and that investigators may wish to censor those locations from analysis. [8–10]

When assessing field change over time, the mean deviation (MD) of all field points is often used for analysis. The MD is considered a robust parameter for following glaucomatous progression.[3,7,11–13]An extensive analysis of available data that led to practical recommendations on measuring glaucomatous visual field progression recommended using MD.[14] The FDA has stated that it will consider a change of 7 dB in the mean sensitivity of the whole field as a clinically significant change.[13] However, demonstrating such a between-group difference in a clinical trial seems impractical; two large studies have demonstrated that untreated glaucoma progresses on average by approximately 1 dB per year,[15,16] and neuroprotection studies are likely to be ethically limited to enroll only subjects treated with IOP-lowering medications. Using only select sub-regions or particular points in the field is less common in the reporting or planning of clinical trial results. One reason for this may be the notorious poorer reproducibility of single point sensitivity, with the noise of test-retest variability somewhat averaging out when a global index such as total-field MD is used. If this increase in variability can be mitigated or prevented, we propose that endpoints based on select regions or points in the field may be beneficial. For example, as glaucoma is often asymmetric in both vertical hemifields,[17] separate analysis of each hemifield may be more sensitive to change and offer a more diverse study population. Since progression is most often observed in already abnormal areas of the field,[18–20] an endpoint which analyses only points with decreased sensitivity may be more sensitive as well. Exclusive analysis

of diseased points may be particularly useful in a study seeking visual function improvement, since healthy points are not likely to show improvement and therefore may "dilute" the effect of a novel intervention.

The Compass (CenterVue, Padua, Italy) is a recently available perimeter which performs visual field testing while continuously tracking the fundus using a scanning laser ophthalmoscope, and placing each stimulus according to the fundus coordinates at that exact moment. This design is intended to reduce variability of test results caused by instability of fixation.

This study had several purposes. We sought to characterize potential endpoints that are based on the mean deviation of particular subsets of points and see if or to what degree they involved a decrease in test-retest repeatability compared with the total field MD. We analyzed each vertical hemifield and a parameter based on only the statistically abnormal points in the field. We also sought to characterize these endpoints after censoring points with low sensitivity. Finally, we sought to compare these endpoints as obtained with the Humphrey Field Analyzer (Zeiss, Dublin, CA) versus the Compass perimeter.

Methods

This was a retrospective analysis of a subset of data obtained during a multi-center cross-sectional study which has been reported previously.[21] All patients gave their written informed consent to participate in the study. Ethics Committee approval was obtained (International Ethics Committee of Milan, Zone A, July 22, 2015, ref: Prot. N^{o} 0019459), and the study was registered as a clinical trial (ISRCTN13800424). This study adhered to the tenets of the Declaration of Helsinki.

As part of that study, subjects were recruited with an expert clinical diagnosis of glaucoma based on RNFL spectral-domain OCT or optic nerve photographs. A subset of these subjects, all previously experienced with automated perimetry, performed 4 VF tests in random order using the 24-2 grid, 2 with HFA using the SITA standard strategy and 2 with Compass using the ZEST strategy, in randomized order. All examinations were done within a time span of 7 days. One eye was tested per subject.

Total deviation values and probabilities were extracted from both machines as appear in the single field analysis report. For the whole field, vertical hemifields and our defined subsets of field points, we calculated a simple, non-weighted MD by averaging all total deviation values, except the 2 blindspot locations. Importantly, for the whole field this should be distinguished from the "MD" reported by both perimeters which is a weighted average. The MDs for subsets of points were derived after censoring the points in the field according to several criteria. In the HFA tests, the total-field and hemifield MDs were first censored (MDc) for points with sensitivity less than 15 dB. Since sensitivity has been reported to be 1.4 dB lower, on average, using Compass compared to HVF,[21] we censored the Compass values at lower than 14 dB; to determine the relationship of the mean deviations of the two instruments in our cohort, Humphry and Compass mean deviations were placed on a scatterplot and principal curve analysis was performed using Rstudio ver. Version 1.3.1093 and the "lowess" fit of the *principal_curve* function of the R-package *princurve.* This indeed showed an approximate linear relationship

with a slope of 0.96 along the entire measured range of mean deviations (see figure, Supplemental Digital Content 1, http://links.lww.com/IJG/A545).

Using the censored test results we subsequently derived 2 outcome parameters based on only "diseased" points, i.e. those with a total deviation probability of either <5% or <2%, MDc5% and MDc2%, respectively.

All parameters were summarized as average with standard deviation. Test-retest variability was evaluated using the 95% limits of agreement (95%Loa) as described by Bland and Altman.[22] This parameter indicates the limits between which a certain test result can be expected to be repeated in a second test in 95 percent of occasions, i.e. excluding atypical outliers.

Results

Data from 30 subjects were available, aged 70.2 ± 8.1 years. Tables 1 (HFA) and 2 (Compass) present average, SD and 95%LoA for the 4 types of visual field indices – based on all field points, only points after censoring for low sensitivity, and only points after further censoring of statistical normality; each of the 4 endpoints is analyzed for the full field and separately for each hemifield. The top row in each table presents the simple, non-weighted average of all total deviation values; these were very close to the values for the weighted MD as reported by the machines that were -7.6 ± 7.2 dB for the HVF with 95%LoA of 5.6, and 7.5 ± 6.6 dB for the Compass with 95%LoA of 4.9. Whereas average MD was identical between the two perimeters at -7.5 dB, it varied by up to 1.0 when the alternate indices were compared.

It is clear that potential endpoints differ significantly from each other in aspects which are important when designing clinical trials. For example, censoring field data for points with low sensitivities resulted in a decrease in both the average deficit and SD of the baseline dataset, i.e. the "study population". Using the ratio of average deficit to SD as an index of signal to noise, we get an improvement for MDc relative to MD of 18% for HFA and 51% for Compass. The endpoints which encompass only those censored points which are also flagged as abnormal (MDc5% and MDc2%) show average deviations which increase again and are similar to those of the original all-points MD, while the SD continues to decrease. Consequently, the average-to-SD ratio increases substantially relative to the ratio for MD by 245% (MDc5%) and 347% (MDc2%) for Humphrey and 272% and 334% for Compass, suggesting a much improved diagnostic power. The 95%LoA of these endpoints were higher with Compass than with HVF. As is expected, with increasing limitation of eligible field points there is a progressive decrease in the number of these points, as shown in tables 3 and 4 for HVF and Compass, respectively. Whereas censoring points with lower sensitivities resulted in a similar number of eligible points with the 2 machines, limiting analysis to points with statistically abnormal deviations led to a substantially lower number of points using Compass compared with Humphrey. With increasing restriction by sensitivity, abnormality, and to hemifields, there was an increasing number of eyes that were excluded from analysis due to not having even a single eligible point.

Discussion

In this analysis we offer a framework for the construction of several potential endpoints based on visual field testing for use in glaucoma clinical trials. Investigators planning a trial and striving

for the primary endpoint with the best chance to show a statistically significant effect for their novel intervention can choose the endpoint best suited for their specific needs. For example, an intervention may be thought to be more effective in mildly diseased cells, and accordingly an endpoint with a lower (better) mean deviation may be preferred. In a trial looking for functional improvement, an outcome measure with a lower (worse) deviation may be preferred, as it allows more room for improvement. Generally, a smaller SD of the baseline population allows planning of a smaller sample size for the trial, making it more feasible. Smaller test-test variability, as expressed here using narrower Bland-Altman 95% LoA, allows planning of fewer visual field tests during the trial, again making it more feasible.

We censored points with low sensitivities following the publications of Gardiner et al. who originally constructed frequency-of-seeing curves in 4 field locations in 34 eyes with glaucoma and suggested that below sensitivity of 15-19 dB there is "response saturation" and so patient responses are unreliable.[23] They subsequently tested 36 glaucomatous eyes with 2 customized ZEST strategies, and showed that when stimuli brighter than 15 dB were omitted, test-retest variance was reduced by nearly 50%. [8] They then reviewed SITA standard 24-2 tests that had been done by 270 patients with glaucoma every 6 months for at least 8 visits, and calculated the linear change of MD -signal – and the SD of the residuals - noise. When pointwise sensitivities were censored at each cutoff between 15-19 dB, signal-to-noise was significantly better for the adjusted MD compared with the uncensored MD, suggesting that a censored global index is preferable for following progressive visual field change.[9] However, we did not find a similarly dominant effect on test-retest variability. The 95%LoA of the total-field MD using Compass was reduced by 23%, and only by 4% with HVF. An analysis of a larger dataset may clarify this, however we note that in the much larger cohort described by Montessano et al, 95%LoA for mean sensitivities were similar to what we found for mean deviations, 5.4 dB for Compass and 6.2 for HVF.[21]

We are not aware of a clinical study that used hemifield mean deviation as an outcome measure. In one retrospective analysis, Johnson et al suggested that abnormal glaucoma-hemifield-test clusters may be sensitive and specific markers of conversion from a normal to glaucomatous visual field. [24] We think that it may be useful to analyze the 2 hemifields separately since glaucoma is often asymmetric across the horizontal midline, glaucoma progression occurs more often in already diseased points, and since a particular intervention may have a different effect on less or more severely affected areas of the field. As expected, since there are less points in a hemifield, both SD and 95%LoA were somewhat higher for hemifield MD compared with whole-field MD. However, following censoring SD became similar in whole field and hemifield MDs.

We are unaware of a previously published characterization of a visual field endpoint based only on abnormal points. Previous publications have shown that glaucoma progression is most frequently observed in parts of the field with or around points with decrease sensitivity. Boden et al retrospectively reviewed visual fields of 70 patients followed at least semiannually for 2 years and 40 patients followed annually for up to 12 years.[20] Progression was observed most

commonly as deepening of existing scotoma, followed by deepening and expansion; not even a single eye showed repeatable new scotomas. De Moraes et al. analyzed 587 eyes of 587 patients who had a mean 11.1 visual field tests over a mean of 6.4 years. Rates of progression were calculated for each point as well as for the global mean deviation.[19] In a later analysis of these data, the authors reported that 90% of the global progression was driven by progression of the points that had been abnormal at baseline. [7] Here, we restricted the censored pointwise data to only those points with a total deviation probability <5%, and than only those at <2%. By doing this, the MD increased again and was almost as high as the original uncensored MD. Using HVF, the SD continued to decrease compared with the SD of the all-points censored MD, and remained as low with Compass. These attributes would seem to make these parameters ideal endpoints compared with the original uncensored MD. However, employing more restrictions necessarily means progressive decrease in the number of available points for analysis. In this regard it is interesting to note that while with both machines there were identical number of points when only low sensitivity was censored, when restricting according to total deviation probability there were fewer points with Compass, even half as many as with HVF for some of the parameters. Apparently there is some difference in how the HVF SITA strategy and the Compass Zest strategy define statistical abnormality. In any case, investigators may use our findings to guide the definition of a minimal number of abnormal points for enrolment of an eye in a particular study and expect this to limit possible recruits. It is important to note that our cohort was defined as glaucoma only based on imaging, and so some of the field tests did not have a classic focal defect. If potential recruits in a clinical study are required to have a visual field defect, this will likely lead to higher numbers of eligible points even after censoring and probability restriction, and a very low number of patients excluded from analysis due to too few eligible points. The paucity of eligible points with Compass is likely a main reason for the increased 95%LoA; enrolling only eyes with visual field defects and therefore more eligible points will likely lead to more favorable repeatability data.

Mean deviations were similar between HFA and Compass before censoring, for the whole field and for each hemifield. This is in line with the published comparison by Montesano et al in a much larger cohort of 499 eyes of 499 subjects with glaucomatous neuropathy. [21] Whereas these authors reported a statistically significantly lower average mean sensitivity with Compass, 20.5 ± 6.7 dB compared with HFA, 21.9 ± 6.9 dB, the more clinically useful parameter of mean deviation was nearly identical and not statistically significant, -6.55 ± 6.60 dB with Compass and -6.50 ± 6.63 dB with HFA. After we censored for low sensitivity, mean deviations were less negative with Compass compared with HFA by 0.7-0.9 dB.

Again, 95%LoA were generally lower with Compass for raw and uncensored mean deviations, implying better inter-test agreement. Conversely, limits of agreement were higher with Compass when only statistically abnormal points were included. As mentioned above, we attribute this to the lower number of eligible points with the Compass.

The present study benefits from a high-quality dataset obtained during a large prospective multi-center study, but has several limitations. The sample size is small, and so outliers can skew the

results; a larger study population may have better test-retest reproducibility manifest by narrower 95%LoA. As mentioned, enrollment was based on glaucoma diagnosed only by imaging; enrollment that includes criteria for a glaucomatous visual field defect will likely increase homogeneity of the dataset and increase the number of available points for analysis. Despite all subjects having previous experience with perimetry using HFA, they were new to using the Compass perimeter. It is possible that further experience can lead to even better reproducibility with this machine. Test-retest variability of the censored indices was based on averaged deviations in the two repeat exams, regardless of whether points were flagged as abnormal in one or both exams. While this stems from a practical clinically-oriented goal, a more fundamental analysis on a larger cohort of eyes could be performed requiring each point to be flagged in both exams; due to the large pointwise variability inherent to perimetry this methodology would likely require a much larger cohort.

In conclusion, we found that outcome measures based on mean deviation of subsets of visual field points do not necessarily involve an increase in dataset heterogeneity or test-retest variability compared with the familiar total-field mean deviation. Restricting analysis to only points with sensitivity above 14-15 dB and only points with statistically abnormally low sensitivity led to a much more homogenous dataset. Test-retest variability was better with Compass for raw and sensitivity-censored parameters, and better with HVF when only abnormal points were included. These outcome measures, applied to the whole field or separate hemifields may be useful for functional assessment in glaucoma clinical trials.

1. Levin LA, Crowe ME, Quigley HA, et al. Neuroprotection for glaucoma: Requirements for clinical translation. *Experimental Eye Research*. 2017;157:34-37. doi:10.1016/j.exer.2016.12.005
2. Chang EE, Goldberg JL. Glaucoma 2.0: Neuroprotection, neuroregeneration, neuroenhancement. *Ophthalmology*. 2012;119(5):979-986. doi:10.1016/j.ophtha.2011.11.003
3. Musch DC, Gillespie BW, Palmberg PF, Spaeth G, Niziol LM, Lichter PR. Visual field improvement in the collaborative initial glaucoma treatment study. *American Journal of Ophthalmology*. 2014;158(1). doi:10.1016/j.ajo.2014.04.003
4. Wright TM, Goharian I, Gardiner SK, Sehi M, Greenfield DS. Short-term enhancement of visual field sensitivity in glaucomatous eyes following surgical intraocular pressure reduction. *American Journal of Ophthalmology*. 2015;159(2). doi:10.1016/j.ajo.2014.11.012
5. Hsu WY, Lane HY, Lin CH. Medications used for cognitive enhancement in patients with schizophrenia, bipolar disorder, Alzheimer's disease, and Parkinson's disease. *Frontiers in Psychiatry*. 2018;9(APR). doi:10.3389/fpsyt.2018.00091
6. Wu Z, Medeiros FA. Impact of Different Visual Field Testing Paradigms on Sample Size Requirements for Glaucoma Clinical Trials. *Scientific Reports*. 2018;8(1). doi:10.1038/s41598-018-23220-w
7. de Moraes CG, Liebmann JM, Levin LA. Detection and measurement of clinically meaningful visual field progression in clinical trials for glaucoma. *Progress in Retinal and Eye Research*. 2017;56. doi:10.1016/j.preteyeres.2016.10.001
8. Gardiner SK, Mansberger SL. Effect of restricting perimetry testing algorithms to reliable sensitivities on test-retest variability. *Investigative Ophthalmology and Visual Science*. 2016;57(13):5631-5636. doi:10.1167/iovs.16-20053
9. Pathak M, Demirel S, Gardiner SK. Reducing variability of perimetric global indices from eyes with progressive glaucoma by censoring unreliable sensitivity data. *Translational Vision Science and Technology*. 2017;6(4):11. doi:10.1167/tvst.6.4.11
10. Gardiner SK, Swanson WH, Demirel S. The effect of limiting the range of perimetric sensitivities on pointwise assessment of visual field progression in glaucoma. *Investigative Ophthalmology and Visual Science*. 2016;57(1). doi:10.1167/iovs.15-18000
11. Bengtsson B, Heijl A. Lack of visual field improvement after initiation of intraocular pressure reducing treatment in the early manifest glaucoma trial. *Investigative Ophthalmology and Visual Science*. 2016;57(13). doi:10.1167/iovs.16-19389
12. Kazemian P, Lavieri MS, van Oyen MP, Andrews C, Stein JD. Personalized Prediction of Glaucoma Progression Under Different Target Intraocular Pressure Levels Using Filtered Forecasting Methods. *Ophthalmology*. 2018;125(4). doi:10.1016/j.ophtha.2017.10.033
13. Weinreb RN, Kaufman PL. The glaucoma research community and FDA look to the future: A report from the NEI/FDA CDER glaucoma clinical trial design and endpoints symposium. *Investigative Ophthalmology and Visual Science*. 2009;50(4). doi:10.1167/iovs.08-2843
14. Chauhan BC, Garway-Heath DF, Goñi FJ, et al. Practical recommendations for measuring rates of visual field change in glaucoma. *The British journal of ophthalmology*. 2008;92(4):569-573. doi:10.1136/bjo.2007.135012

15. Garway-Heath DF, Crabb DP, Bunce C, et al. Latanoprost for open-angle glaucoma (UKGTS): A randomised, multicentre, placebo-controlled trial. *The Lancet*. 2015;385(9975):1295-1304. doi:10.1016/S0140-6736(14)62111-5
16. Heijl A, Leske MC, Bengtsson B, Hyman L, Bengtsson B, Hussein M. Reduction of intraocular pressure and glaucoma progression: Results from the Early Manifest Glaucoma Trial. *Archives of Ophthalmology*. 2002;120(10). doi:10.1001/archopht.120.10.1268
17. Yousefi S, Sakai H, Murata H, et al. Asymmetric patterns of visual field defect in primary open-angle and primary angle-closure glaucoma. *Investigative Ophthalmology and Visual Science*. 2018;59(3). doi:10.1167/iovs.17-22980
18. Mikelberg FS, Drance SM. The mode of progression of visual field defects in glaucoma. *American Journal of Ophthalmology*. 1984;98(4). doi:10.1016/0002-9394(84)90128-4
19. de Moraes CG v., Juthani VJ, Liebmann JM, et al. Risk factors for visual field progression in treated glaucoma. *Archives of Ophthalmology*. 2011;129(5). doi:10.1001/archophthalmol.2011.72
20. Boden C, Blumenthal EZ, Pascual J, et al. Patterns of glaucomatous visual field progression identified by three progression criteria. *American Journal of Ophthalmology*. 2004;138(6). doi:10.1016/j.ajo.2004.07.003
21. Montesano G, Bryan SR, Crabb DP, et al. A Comparison between the Compass Fundus Perimeter and the Humphrey Field Analyzer. *Ophthalmology*. 2019;126(2):242-251. doi:10.1016/j.ophtha.2018.08.010
22. Martin Bland J, Altman DG. STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT. *The Lancet*. 1986;327(8476). doi:10.1016/S0140-6736(86)90837-8
23. Gardiner SK, Swanson WH, Goren D, Mansberger SL, Demirel S. Assessment of the reliability of standard automated perimetry in regions of glaucomatous damage. *Ophthalmology*. 2014;121(7):1359-1369. doi:10.1016/j.ophtha.2014.01.020

Table 1. Mean deviation and 95%LoA values using HFA, in dB.

	Complete field	Upper Hemifield	Lower Hemifield
MD	-7.5 ± 6.9 (5.5)	-8.4 ± 8.1 (6.6)	-6.7 ± 8.3 (6.6)
MDc	-3.6 ± 2.8 (5.3)	-4.1 ± 2.9 (5.4)	-3.4 ± 3.3 (7.0)
MDc5%	-6.4 ± 1.7 (4.6)	-6.5 ± 1.8 (4.1)	-6.4 ± 1.6 (4.6)
MDc2%	-7.3 ± 1.5 (5.6)	-7.6 ± 1.5 (5.4)	-6.9 ± 2.2 (6.0)

MD – Mean Deviation, average of total deviation of all eligible points. MDc – MD censored, average of only those points with a sensitivity 15 dB or higher. MDc5% - average of censored points with total deviation p<5%. MDc2% - average of censored points with total deviation p<2%. Presented results are average ± standard deviation, with 95%LoA in parentheses.

Table 2. Mean deviation and 95%LoA values using Compass, in dB.

	Complete field	Upper Hemifield	Lower Hemifield
MD	-7.5 ± 6.6 (4.8)	-8.6 ± 8.0 (6.5)	-6.4 ± 7.6 (5.3)
MDc	-2.9 ± 1.7 (3.7)[1]	-3.4 ± 1.6 (5.1)	-2.5 ± 2.0 (3.6)
MDc5%2	-6.3 ± 1.5 (7.1)	-6.0±1.5 (7.4)	-6.6±1.8 (5.7)
MDc2%	-7.9 ± 1.6 (7.1)	-7.8±1.9 (9.6)	-7.9±1.6 (4.9)

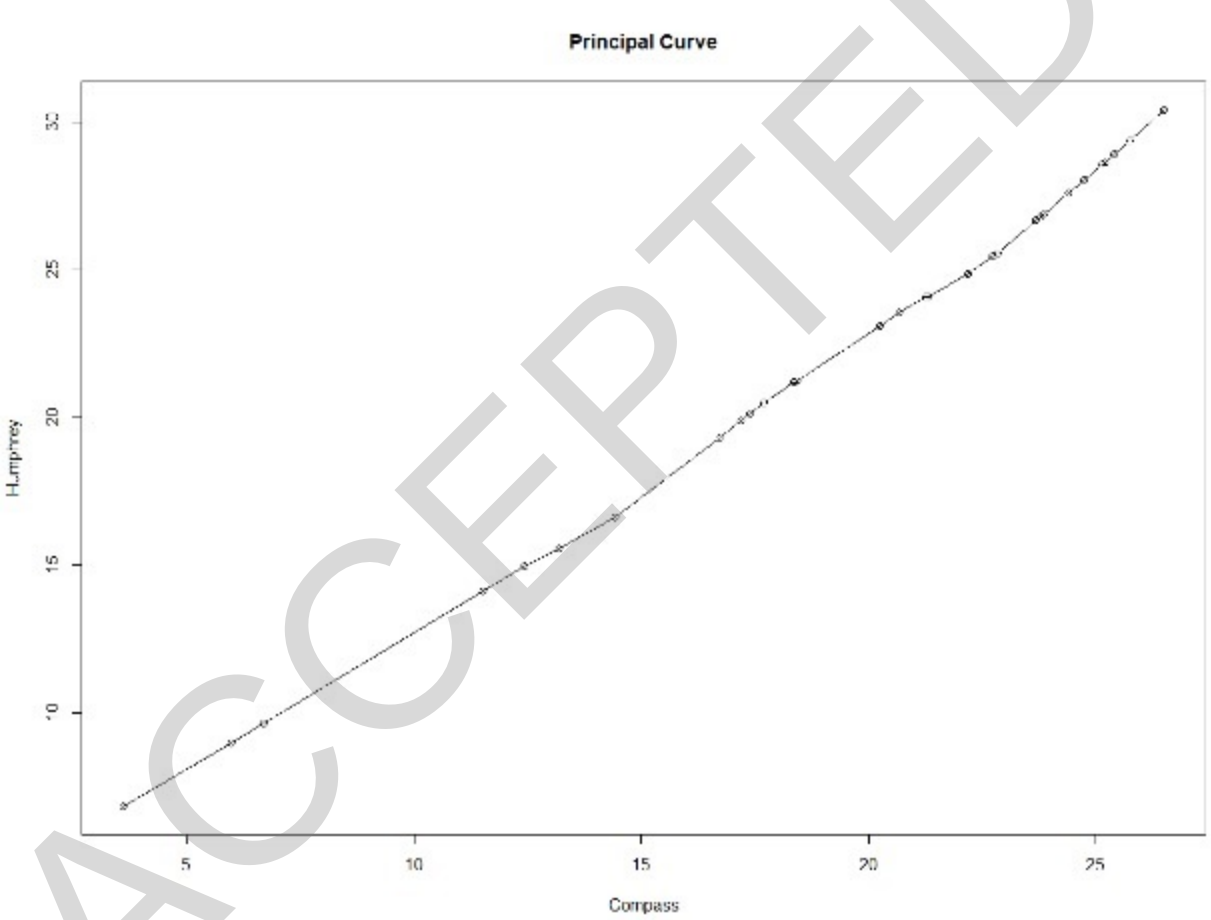
MD – Mean Deviation, average of total deviation of all eligible points. MDc – MD censored, average of only those points with a sensitivity 14 dB or higher. MDc5% - average of censored points with total deviation p<5%. MDc2% - average of censored points with total deviation p<2%. Presented results are average ± standard deviation, with 95%LoA in parentheses.

Table 3. Number of field points eligible for analysis using HVF. If there was not even a single eligible point in either or both of the exams, that eye was excluded from analysis; the number of excluded eyes appear in parenthesis.

	Complete field	Upper Hemifield	Lower Hemifield
MD	52	26	26
MDc	42 ± 12	20 ± 8	21 ± 6
MDc - only abnormal points at p<5%	20 ± 11 (3 eyes excluded)	10 ± 6 (4 eyes excluded)	20 ± 11 (3 eyes excluded)
MDc – only abnormal points at p<2%	15 ± 9 (4 eyes excluded)	8 ± 5 (7 eyes excluded)	9 ± 6 (6 eyes excluded)

Table 4. Number of field points eligible for analysis using Compass. If there was not even a single eligible point in either or both of the exams, that eye was excluded from analysis; the number of excluded eyes appear in parenthesis.

	Complete field	Upper Hemifield	Lower Hemifield
MD	52	26	26
MDc	41.2 ± 12.6	20 ± 7 (1 eye excluded)	22 ± 7
MDc - only abnormal points at p<5%	10 ± 7	7 ± 4 (3 eyes excluded)	6 ± 4 (8 eyes excluded)
MDc – only abnormal points at p<2%	7 ± 5 (3 eyes excluded)	4 ± 3 (8 eyes excluded)	4 ± 3 (11 eyes excluded)



Supplemental Figure 1. Endpoints Humphrey Compass manuscript