



Virtual Reality Oculokinetic Perimetry Test Reproducibility and Relationship to Conventional Perimetry and OCT

Jason A. Greenfield, BA,^{1,*} Michael Deiner, PhD,^{2,*} Anwell Nguyen, BA,² Gadi Wollstein, MD,^{1,3,4}
Bertil Damato, MD, PhD,^{2,5} Benjamin T. Backus, PhD,⁶ Mengfei Wu, MS,¹ Joel S. Schuman, MD,^{1,3,4,7,8}
Yvonne Ou, MD⁹

Purpose: Vivid Vision Perimetry (VVP; Vivid Vision, Inc) is a novel method for performing in-office and home-based visual field assessment using a virtual reality platform and oculokinetic perimetry. Here we examine the reproducibility of VVP Swift and compare results with conventional standard automated perimetry (SAP) and spectral-domain (SD) OCT.

Design: Cross-sectional study.

Participants: Fourteen eyes of 7 patients with open-angle glaucoma (OAG) (average age, 64.6 years; 29% women) and 10 eyes of 5 patients with suspected glaucoma (average age, 61.8 years; 40% women) were enrolled.

Methods: Patients with OAG and suspected glaucoma were enrolled prospectively and underwent 2 VVP Swift examinations. Results were compared with 1 conventional SAP examination (Humphrey Visual Field [HVF]; Zeiss) and 1 SD OCT examination.

Main Outcome Measures: Mean sensitivity (in decibels) obtained for each eye in 2 VVP Swift test sessions and a conventional SAP examination, thickness of the retinal nerve fiber layer (RNFL) and ganglion cell complex (GCC) for the SD OCT examination, and mean test durations of the VVP Swift and SAP examinations.

Results: The mean test duration of VVP Swift in both eyes (8.5 minutes) was significantly shorter ($P < 0.001$) than SAP (12.2 minutes). The average absolute difference of the mean sensitivity between the 2 VVP Swift sessions was found to be 0.73 dB (95% confidence interval [CI], 0.40–1.06). A statistically significant association was found between average mean sensitivity measurements from the VVP and mean deviation (MD) measurements obtained by the HVF with a Pearson correlation coefficient of 0.86 (95% CI, 0.70–0.94; $P < 0.001$). Mean visual sensitivity measurements from the VVP Swift test were significantly associated with average RNFL thickness ($r = 0.66$; $P = 0.014$) and GCC thickness ($r = 0.63$; $P = 0.02$), whereas the correlation coefficients between HVF MD and RNFL and GCC were 0.86 ($P < 0.001$) and 0.83 ($P < 0.001$), respectively.

Conclusions: Our results demonstrated that the VVP Swift test can generate reproducible results and is comparable with conventional SAP. This suggests that the device can be used by clinicians to assess visual function in glaucoma. *Ophthalmology Science* 2022;2:100105 © 2021 by the American Academy of Ophthalmology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

The visual field examination is critical in identifying diseases such as glaucoma and evaluating longitudinal progression. Conventional standard automated perimetry (SAP) is the gold standard for visual field testing in glaucoma management, but it has several limitations. This technology requires that patients travel to the doctor's office to undergo multiple examinations on separate occasions, necessitates a skilled technician for its operation, and creates an uncomfortable and fatiguing experience for patients because of the constant need to suppress their foveation reflex.^{1,2} Furthermore, conventional SAP, such as the Humphrey Visual Field (HVF; Zeiss) examination, has a high inherent variability in its output data that obligates patients to return to the doctor's office for repeated testing to generate clinically reliable results.³ Recent innovations in

virtual reality systems have led to the development of novel approaches to detect visual field abnormalities. Virtual reality visual field examinations have been shown to correlate strongly with conventional SAP.^{4,5} However, the clinical relevance of this technology has yet to be established.

The Vivid Vision Perimetry (VVP; Vivid Vision, Inc) Swift test is a virtual reality-based visual field assessment that can address many challenges associated with conventional SAP.⁶ First, VVP Swift can produce a higher sample volume per patient because it can be performed at home at the patient's convenience. Second, it does not require a skilled technician for successful deployment. Third, VVP Swift uses oculokinetic perimetry methods whereby the patient moves the eye to visualize the test target

presentations. Although oculokinetic perimetry can be performed by an unsupervised patient, conventional SAP requires the patient's eye to be immobile while a trained technician monitors compliance. These novel features have the potential to significantly improve clinicians' ability to diagnose and monitor glaucoma. However, to date, the reproducibility, reliability, and structure–function relationship of this system have not been examined. The aims of this study were (1) to determine the test–retest variability of measurements obtained from the VVP Swift test, (2) to quantify agreement of the VVP test with the HVF examination, and (3) to measure the structure–function relationship between VVP and spectral-domain (SD) OCT. These studies should indicate VVP's scope for detecting and monitoring visual field abnormalities in clinical practice.

Methods

Informed consent was obtained from all participants using a consent form approved by the institutional review board for human research at the University of California, San Francisco, Medical Center (institutional review board approval no., 16-20210), and all research adhered to the tenets of the Declaration of Helsinki. Patients with suspected glaucoma and those in all stages of open-angle glaucoma (OAG) were enrolled prospectively, undergoing 2 VVP Swift examinations, 1 HVF examination, and 1 SD OCT examination (Optovue) during scheduled clinic visits spanning a period of several weeks. The enrolled participants had prior experience with HVF examinations.

Patients with glaucoma were defined as those with a diagnosis of OAG, normal-tension glaucoma, or steroid-induced glaucoma. Those included had mild, moderate, and severe OAG, as evidenced by optic nerve damage by either optic disc or retinal nerve fiber layer (RNFL) structural abnormalities, reliable and reproducible visual field abnormalities consistent with RNFL damage, or both. Visual field defects included persistent scotoma on at least 2 prior HVF tests with < 33% false-positive or false-negative results and fixation losses. Abnormal disc appearance included neuroretinal rim thinning, localized or diffuse RNFL defects, disc hemorrhages, or progressive narrowing of the neuroretinal rim with increased cupping, observed with slit-lamp biomicroscopy and a handheld lens or with SD OCT imaging. A diagnosis of suspected OAG was based on the presence of a consistently elevated intraocular pressure of > 21 mmHg, also known as ocular hypertension, or a suspicious optic nerve or RNFL in 1 or both eyes without detectable visual field defects. One patient demonstrated a plateau iris configuration with suspicious optic nerve and RNFL findings without visual field defects and was included in the suspected glaucoma group.

Additional inclusion criteria included best-corrected visual acuity of at least 20/80, age between 18 and 85 years, and an interpupillary distance of 60 to 66 mm. Patients with a history of epilepsy, active facial infections or acne or rosacea, retinal vein occlusion, wet age-related macular degeneration, proliferative diabetic retinopathy, active cold, cough, and issues with neck strain or head movements and those who could not complete the VVP training module were excluded from enrollment. In addition, patients with dilated pupils also were excluded.

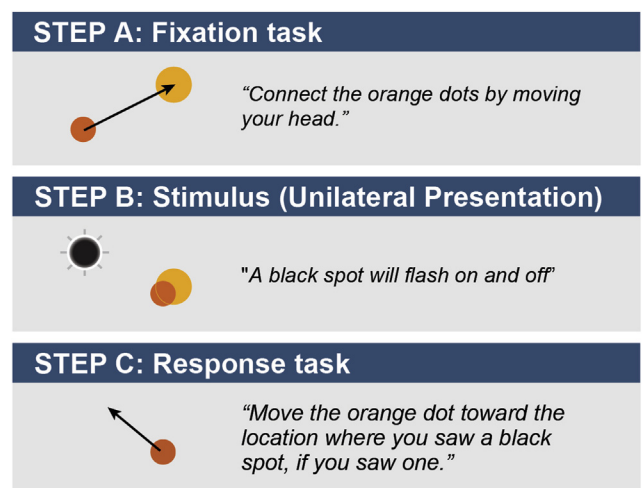
Each patient's visual fields were assessed across 54 test locations in a 24-2 pattern using the VVP device. Participants were examined with VVP Swift in 2 sessions during the same clinic visit, and the mean sensitivity (reported in decibels) was the primary global outcome measure obtained for each eye in both VVP

sessions. In addition, participants underwent 1 24-2 HVF examination (SITA Standard), and the mean sensitivity value was obtained. The cutoffs for reliability indices of HVF examinations were set at a 30% false-positive response rate, 30% false-negative response rate, and 30% fixation loss rate. An SD OCT examination was also administered to each participant, and the average ganglion cell complex (GCC) and RNFL thickness values were obtained. All participants underwent 2 VVP Swift examinations, 1 HVF examination, and 1 SD OCT examination, with completion of HVF and SD OCT examinations occurring within a mean of 15.5 and 18 weeks, respectively, from the VVP Swift examinations and 75% of participants completing all examinations within a 1-year period.

Vivid Vision Perimetry Swift testing was performed in a standardized fashion (Fig 1). Vivid Vision Perimetry Swift tests both eyes during a single session using randomly alternating left-eye and right-eye stimuli. Stimuli were decremental, small black spots of luminance 0.2 cd/m² on a white background of luminance 25 cd/m². Stimulus location was randomly selected from the remaining locations in both eyes. If the first stimulus at a given location was seen, it was not presented again. If it was missed, it was presented 2 more times. Stimuli had a duration of 300 ms and were round, with a diameter of 0.43°. Vivid Vision Perimetry Swift is platform independent, but for this study, the hardware for all tests was the Oculus Go mobile virtual reality headset (Facebook, Inc).

Data Analysis

Summary statistics for numerical variables were presented as mean ± standard deviation, whereas categorical variables were summarized by count (percentage). At a given location, stimuli were missed 0, 1, 2, or 3 times, giving an estimated sensitivity of 27, 12, 6, or 0 dB, respectively.⁷ The repeatability of mean sensitivity was assessed through the differences between the 2 VVP sessions. A Bland-Altman plot based on a mixed-effects model (adjusting for average sensitivity and intrapatient correlation) was created to examine the level of agreement between measurements. To test for the reliability of VVP Swift measurements, the mean sensitivities obtained from first and second VVP Swift sessions were averaged for each eye to compare against the mean deviation (MD) measured by the HVF examination, adjusting for the intrapatient correlation.



For all stimuli points repeat steps A-C

Figure 1. Schematic demonstrating the Vivid Vision Perimetry (VVP) Swift test, taken by participants wearing a virtual reality headset.

Table 1. Patient Demographics and Clinical Characteristics

Characteristic	Suspected Glaucoma (n = 5)	Glaucoma (n = 7)
Age (yrs)	61.8 ± 6.5	64.6 ± 11.4
Gender		
Female	2	2
Male	3	5
HVF		
Average mean deviation (dB)	1.00 ± 0.37	6.72 ± 5.41
Average pattern standard deviation (dB)	1.86 ± 0.44	7.24 ± 3.65
Average visual field index (%)	98.4 ± 1.17	82.79 ± 16.21
VVP mean sensitivity (dB)	26.14 ± 0.59	22.83 ± 2.72
SD OCT		
Mean RNFL (μm)	82.89 ± 11	70.66 ± 11.02
Mean GCC (μm)	87.09 ± 10.57	78.78 ± 9.17

GCC = ganglion cell complex; HVF = Humphrey Visual Field; RNFL = retinal nerve fiber layer; SD = spectral-domain; VVP = Vivid Vision Perimeter. Data are presented as mean ± standard deviation or no. (%).

Finally, the structure–function relationship was evaluated by comparing average mean sensitivity measurements measured by the VVP Swift test and HVF examination with average RNFL and GCC thickness values obtained by SD OCT assessment. The association between mean sensitivity and SD OCT parameters was determined using a repeated-measures correlation coefficient (r) that accounted for the inpatient correlation. Statistical analysis was performed using R software version 3.5.2 (R Foundation for Statistical Computing, Vienna, Austria). A comparison of the test duration between the VVP Swift and HVF examinations was performed using a Student’s paired t test. A P value of < 0.05 was considered statistically significant.

Results

Twelve participants met inclusion criteria and were enrolled, constituting 14 eyes of 7 participants with glaucoma (mean age, 64.6 ± 11.4 years; 29% women) and 10 eyes of 5 participants with suspected glaucoma (mean age, 61.8 ± 6.5 years; 40% women). The clinical characteristics of the study

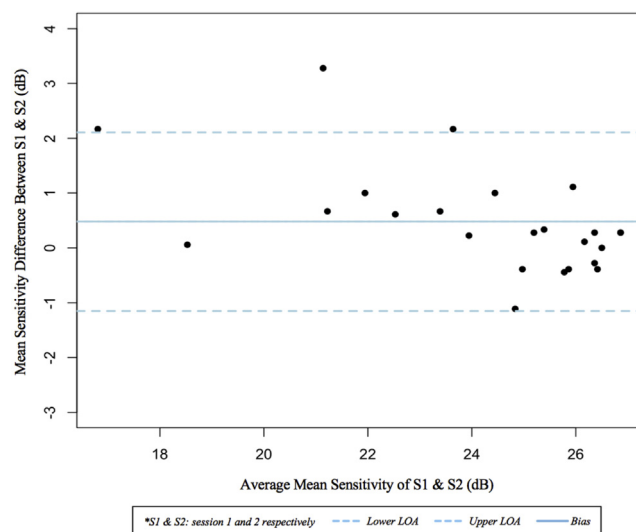


Figure 2. The reproducibility of Vivid Vision Perimetry (VVP) Swift’s mean sensitivity measurements. LOA = limit of agreement.

sample are summarized in Table 1. The mean test duration of VVP Swift examination in both eyes (8.5 minutes from start to finish) was significantly shorter ($P < 0.001$) than for the HVF examination (12.2 minutes of combined time from start to finish for each eye, not including the additional time needed to switch between eyes).

Measurements obtained from the VVP Swift examination demonstrated repeatability as seen in the Bland Altman plot (Fig 2). The average absolute difference of the mean sensitivity between the 2 VVP Swift sessions was found to be 0.73 dB (95% confidence interval [CI], 0.40–1.06). Within our sample of 24 eyes, the test–retest variability (standard deviation of the 2 VVP Swift tests) had a mean value of 0.52 dB (95% CI, 0.28–0.75 dB). For glaucomatous eyes only, the mean was 0.65 dB (95% CI, 0.26–1.04 dB). The results of 3 eyes (12.5%) fell outside the upper and lower limits of agreement (95% CI, –1.15 to 2.11 dB). The level of agreement between repeated VVP Swift measurements showed a general trend of increasing precision as mean sensitivity values increased.

A statistically significant association between average mean sensitivity measurements from the VVP Swift with MD measurements measured by the HVF examination was observed ($P < 0.001$; Fig 3). Using a linear mixed-effects model, the r coefficient for the association between the VVP Swift and HVF examination was 0.86 (95% CI, 0.70–0.94; $P < 0.001$).

A relationship between retinal structure and ocular function was identified through comparison of VVP Swift and SD OCT measurements. Mean visual sensitivity measurements were significantly associated with average RNFL thickness ($r = 0.66$; $P = 0.014$) and GCC thickness ($r = 0.63$; $P = 0.02$), whereas the correlation coefficients between Humphrey MD and RNFL and GCC were 0.86 ($P < 0.001$) and 0.83 ($P < 0.001$), respectively.

Discussion

Conventional SAP is the recognized standard for measuring visual function in glaucoma. Limitations for this technology

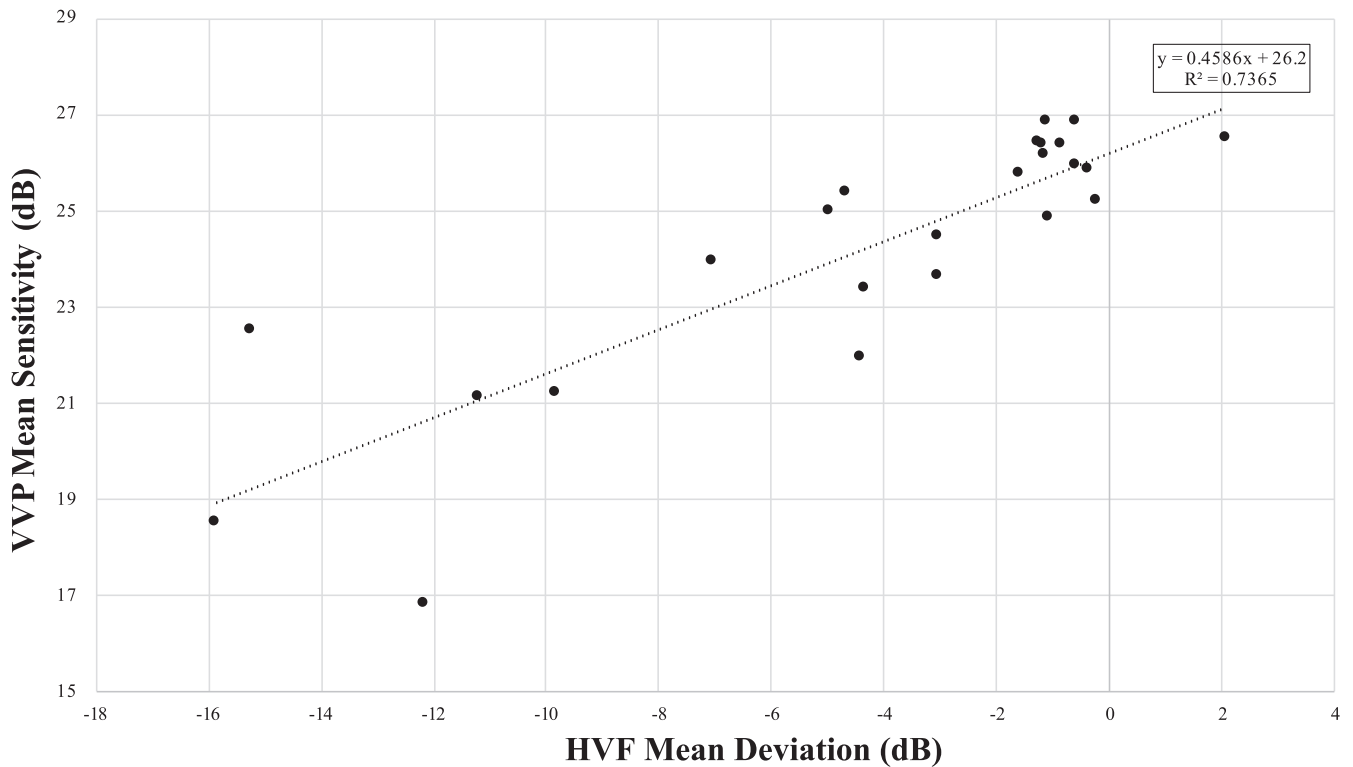


Figure 3. The association between Vivid Vision Perimetry (VVP) Swift mean sensitivity and Humphrey Visual Field (HVF) mean deviation.

are its high test–retest variability, its long examination time, and its requirement for patients to suppress their foveation reflex.⁸ Recent advances in virtual reality perimetry systems have shown the potential for this technology to quantify visual field loss in patients with glaucoma.^{4,9} In this study, we demonstrated that the VVP Swift test can perform comparably with an HVF examination in an in-office setting through examining its measurement repeatability and structure–function relationship.

Our results indicate that the VVP Swift test can generate reproducible measurements. We also found that VVP Swift examinations were significantly shorter in test duration than HVF examinations. Relative to the HVF examination, the VVP Swift test showed a test–retest variability that was comparable with reported values for moderately reliable test-takers with glaucoma (1.0 dB for conventional SAP¹⁰ vs. 0.65 dB for glaucoma eyes in this study). Previous literature findings reported poorer precision in eyes with advanced glaucoma and in regions of the visual field characterized by severe damage.¹¹ In the present study, this trend of decreasing precision with increased visual field damage was similarly observed with VVP Swift. This trend may also suggest the possibility of a practice effect, whereby patients with poorer visual sensitivity gained experience with the device after the first session and were able to perform better during the second session relative to patients with higher visual sensitivity measurements. In addition, average mean sensitivity measurements obtained with the VVP Swift demonstrated reliability through a statistically significant association

with MD measurements from the HVF. This relationship indicates agreement between global indices measured by the VVP Swift test and HVF examination.

Visual field measurements recorded by the HVF have an established relationship with retinal structure measured by SD OCT.¹² Similarly, VVP Swift measurements of the visual field were found to be associated with the structural measurements of both average RNFL and GCC thickness by SD OCT imaging. The correspondence identified in this study is comparable with the known structure–function relationship demonstrated by SAP and SD OCT.^{13,14} In addition, the association between VVP Swift and SD OCT was comparable with the association between HVF and SD OCT conducted in this study, whereby both associations demonstrated statistical significance. The structure–function results lend further support to indicate that VVP Swift is an effective clinical tool for quantifying visual function in glaucoma.

Several limitations of this study must be noted. This study focused on the VVP Swift’s ability to perform clinic-based examinations in the presence of a skilled technician. Thus, outcomes obtained from VVP Swift’s home-based assessment may differ. This study was also limited by a small and nonuniform sample size that hindered meaningful statistical analyses. The scheduling of data collected was similarly nonuniform, where the participants’ average time to undergo SD OCT and HVF examinations was within 15.5 weeks and 18 weeks of the VVP Swift examinations, respectively. In some cases, change in disease progression may have occurred, thereby causing bias. The

analysis of reproducibility between the VVP Swift and HVF examinations was hampered by nonstandardized output statistics between the 2 devices. Thus, no conversion of mean sensitivity between devices was possible, and a correlation was determined to be the best available statistical measure. Furthermore, this reproducibility assessment was relative to the current gold standard (HVF), which has measurable imprecision and bias. The possibility of a learning effect could have impacted the results of the repeatability analysis because the participants could have gained increased skill at using the VVP Swift between the 2 test sessions. Finally, data analysis was based on summary statistics and did not include localized parameters.

Footnotes and Disclosures

Originally received: November 29, 2021.

Final revision: December 15, 2021.

Accepted: December 15, 2021.

Available online: December 21, 2021. Manuscript no. D-21-00197.

¹ Department of Ophthalmology, NYU Langone Health, NYU Grossman School of Medicine, New York, New York.

² Department of Ophthalmology and Proctor Foundation, University of California, San Francisco, San Francisco, California.

³ Department of Biomedical Engineering, NYU Tandon School of Engineering, New York, New York.

⁴ Department of Neural Science, NYU School of Arts and Sciences, New York, New York.

⁵ Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom.

⁶ Vivid Vision, Inc., San Francisco, California.

⁷ Department of Electrical and Computer Engineering, NYU Tandon School of Engineering, New York, New York.

⁸ Department of Neuroscience and Physiology, NYU Langone Health, NYU Grossman School of Medicine, New York, New York.

⁹ Department of Ophthalmology, University of California, San Francisco, San Francisco, California.

*Both authors contributed equally as first authors.

Disclosure(s):

All authors have completed and submitted the ICMJE disclosures form.

The author(s) have made the following disclosure(s): B.T.B.: Employee, Intellectual property, Equity owner – Vivid Vision, Inc.; Patent – Systems and Methods for Visual Field Analysis

J.S.S.: Royalties – Zeiss

Supported in part by the National Eye Institute, National Institutes of Health, Bethesda, Maryland (grant no.: P30EY002162); the University of California, San Francisco, San Francisco, California (Research Allocation Program); and Research to Prevent Blindness, Inc., New York, New York (unrestricted grants to the Department of Ophthalmology at University of California, San Francisco, and the Department of Ophthalmology, NYU Langone Health, NYU Grossman School of Medicine). The sponsors or

In conclusion, our results demonstrate that the VVP Swift test generates reproducible results and is comparable with conventional SAP. This suggests that the device can be used by clinicians to manage patients with suspected or established glaucoma. Future studies should include examinations using newer versions of VVP Swift tests that include “gamified” interfaces and applications for at-home testing that can collect a higher volume of data than the Swift test and thereby improve precision. Future pointwise analyses should assess the location-based reliability of the VVP Swift test compared with the HVF examination, and sectoral analyses should determine if localized areas of nerve fiber thinning demonstrate similar or stronger associations with functional assessment by VVP Swift.

funding organization had no role in the design or conduct of this research. Vivid Vision, Inc., were scientific collaborators; the University of California, San Francisco, has an unfunded collaborative agreement with them.

Dr Joel Schuman, an Editor of this journal, was recused from the peer-review process of this article and had no access to information regarding its peer-review.

HUMAN SUBJECTS: Human subjects were included in this study. Informed consent was obtained from all subjects using a consent form approved by the Institutional Review Board for human research at University of California San Francisco (UCSF) Medical Center (IRB Approval number: 16-20210) and all research adhered to the tenets of the Declaration of Helsinki.

No animal subjects were included in this study.

Author Contributions:

Conception and design: Greenfield, Deiner, Damato, Backus, Schuman, Ou
Analysis and interpretation: Greenfield, Deiner, Wollstein, Wu, Schuman, Ou

Data collection: Deiner, Nguyen

Obtained funding: N/A

Overall responsibility: Greenfield, Deiner, Nguyen, Wollstein, Damato, Backus, Wu, Schuman, Ou

Abbreviations and Acronyms:

CI = confidence interval; **GCC** = ganglion cell complex; **HVF** = Humphrey Visual Field; **MD** = mean deviation; **OAG** = open-angle glaucoma; **RNFL** = retinal nerve fiber layer; **SAP** = standard automated perimetry; **SD** = spectral-domain; **VVP** = Vivid Vision Perimetry.

Keywords:

Glaucoma, OCT, Perimetry, SAP, Tele-ophthalmology, Virtual reality, Oculokinetic perimetry, Visual field.

Correspondence:

Yvonne Ou, MD, Department of Ophthalmology, University of California, San Francisco, 10 Koret Way, K141, San Francisco, CA 94143. E-mail: Yvonne.Ou@ucsf.edu.

References

- Hudson C, Wild JM, O'Neill EC. Fatigue effects during a single session of automated static threshold perimetry. *Invest Ophthalmol Vis Sci*. 1994;35(1):268–280.
- Schimiti RB, Avelino RR, Newton KJ, Costa VP, et al. Full-threshold versus Swedish Interactive Threshold Algorithm (SITA) in normal individuals undergoing automated perimetry

- for the first time. *Ophthalmology*. 2002;109(11):2084–2092. discussion 2092.
3. Wu Z, Saunders LJ, Daga FB, et al. Frequency of testing to detect visual field progression derived using a longitudinal cohort of glaucoma patients. *Ophthalmology*. 2017;124(6):786–792.
 4. Tsapakis S, Papaconstantinou D, Diagourtas A, et al. Visual field examination method using virtual reality glasses compared with the Humphrey perimeter. *Clin Ophthalmol*. 2017;11:1431–1443.
 5. Wroblewski D, Francis BA, Sadun A, et al. Testing of visual field with virtual reality goggles in manual and visual grasp modes. *Biomed Res Int*. 2014;2014:206082.
 6. Deiner MS, Damato BE, Ou Y. Implementing and monitoring at-home virtual reality oculo-kinetic perimetry during COVID-19. *Ophthalmology*. 2020;127(9):1258.
 7. Backus BT, Blaha J, Gupta M. Estimating visual sensitivity from single-intensity stimuli in visual field tests for patients with glaucoma: theory. *Invest Ophthalmol Vis Sci*. 2020;61(7):3895.
 8. Alencar LM, Medeiros FA. The role of standard automated perimetry and newer functional methods for glaucoma diagnosis and follow-up. *Indian J Ophthalmol*. 2011;59(Suppl):S53–S58.
 9. Razeghinejad R, Gonzalez-Garcia A, Myers JS, Katz LJ. Preliminary report on a novel virtual reality perimeter compared with standard automated perimetry. *J Glaucoma*. 2021;30(1):17–23.
 10. Chauhan BC, Garway-Heath DF, Goñi FJ, et al. Practical recommendations for measuring rates of visual field change in glaucoma. *Br J Ophthalmol*. 2008;92:569–573.
 11. Wall M, Doyle CK, Zamba KD, et al. The repeatability of mean defect with size III and size V standard automated perimetry. *Invest Ophthalmol Vis Sci*. 2013;54(2):1345–1351.
 12. Leite MT, Zangwill LM, Weinreb RN, et al. Structure-function relationships using the Cirrus spectral domain optical coherence tomograph and standard automated perimetry. *J Glaucoma*. 2012;21(1):49–54.
 13. Na JH, Kook MS, Lee Y, Baek S. Structure-function relationship of the macular visual field sensitivity and the ganglion cell complex thickness in glaucoma. *Invest Ophthalmol Vis Sci*. 2012;53(8):5044–5051.
 14. Nilforushan N, Nassiri N, Moghimi S, et al. Structure–function relationships between spectral-domain OCT and standard achromatic perimetry. *Invest Ophthalmol Vis Sci*. 2012;53(6):2740–2748.