AMERICAN ACADEMY OF OPHTHALMOLOGY®

Esterman Visual Field Testing Using a Virtual Reality Headset in Glaucoma

Meghan Sharma, MD, MPH, Eleonore Savatovsky, MD, PhD, Laura Huertas, MPH, Robert O'Brien, PhD, Alana Grajewski, MD, Elena Bitrian, MD

Purpose: To test the use of a virtual reality visual field headset (VRVF) for implementation of the Esterman visual field (EVF) test as compared with standard automated perimetry (SAP) among people with glaucoma.

Design: Experimental design.

Subjects: Patients with mild to severe glaucoma ranging from 10 to 90 years who presented for follow-up at a glaucoma clinic in Miami, Florida were eligible.

Methods: Participants performed the EVF test on both SAP and VRVF. Five glaucoma-trained ophthalmologists were then asked to rate all anonymized SAP and RVF tests as a "pass" or "failure" based on Florida state law.

Main Outcome Measures: Point-by-point concordance between original VRVF EVF test results and SAP EVF test results was calculated using the Kappa statistic. Concordance between SAP and VRVF was secondarily assessed with a conditional logistic regression based on the pass-failure determinations by the glaucoma-trained ophthalmologists. Interrater agreement on test pass-failure determinations was also calculated. Finally, test results on SAP versus VRVF were compared based on Esterman efficiency score (EES), the number of correct points divided by the number of total points, and duration of testing.

Results: Twenty-two subjects were included in the study with ages ranging from 14 to 78 years old. Concordance between VRVF and SAP test using point-by-point analysis was poor ($\kappa = 0.332$, [95% confidence intervals {CI}: 0.157, 0.506]) and somewhat increased using pass-failure determinations from ophthalmologists ($\kappa = 0.657$, [95% CI: 0.549, 0.751]). Ophthalmologists were more likely to agree amongst themselves on pass-failure determinations for VRVF tests ($\kappa = 0.890$, [95% CI: 0.726, 0.964]) than for SAP ($\kappa = 0.590$, [95% CI: 0.372, 0.818]); however, VRVF demonstrated significantly lower EES than SAP (median EES difference: 4.5 points, $P = 0.021$).

Conclusions: This pilot study is the first to assess the implementation of the EVF test using a virtual reality headset. Based on the weak overall agreement between VRVF and SAP, the current VRVF EVF test is not an acceptable determinant of driver's licensing. However, ophthalmologists were more likely to agree amongst themselves on VRVF test reports than on SAP reports. With further testing and improvement, virtual reality may eventually become a portable and convenient method for administering the EVF test.

Financial Disclosure(s): Proprietary or commercial disclosure may be found in the Footnotes and Disclosures at the end of this article. *Ophthalmology Science 2024;4:100534*

Visual field damage has been shown to be related to impaired driving performance. Glaucoma, a disease that particularly affects peripheral vision, has been noted to be a leading reason for issues obtaining a driver's license at a higher age from an ophthalmic standpoint. The Esterman visual field (EVF) test, traditionally performed using standard automated perimetry (SAP), is a binocular, suprathreshold test consisting of 120 size III dB points and is typically performed by patients with visual field deficits such as glaucoma.[1] Outcomes are represented on a chart of the 120 points with binary outcomes as to whether the individual correctly identified each point using his or her peripheral vision.[2] As one of the first developed binocular visual field tests, the EVF test has become a standard in assessing visual field deficits that would affect driving performance due to its ability to simulate binocular vision while operating a motor vehicle. Moreover, the test emphasizes central and inferior points—areas of the visual field with more functional importance while driving.[3,4] For these reasons, motor vehicle offices typically use an ophthalmologist's interpretation of a patient's EVF test to help determine whether an individual may legally drive.[1,4]

The EVF test was first created in 1967 by Dr Benjamin Esterman as a visual field system that could quantify field function. It was originally a monocular grid of 104 points with importance given to the center, horizontal meridian, and lower visual field and a 25° radius. One year later, he presented a second monocular grid with 100 points but with a wider radius, capturing 50 degrees nasally and 80 degrees temporally. Dr Esterman presented his final visual field grid

https://doi.org/10.1016/j.xops.2024.100534
ISSN 2666-9145/24

15 years after his original grid by merging his previous monocular grids into a binocular grid with 120 tested points—this became the EVF test used today.[2,4] Within 2 years, the American Medical Association adopted the binocular Esterman disability score, defined as the percent of defects out of total points tested on the EVF test, as the standard to evaluate visual impairment, and the final Esterman binocular grid was incorporated into standard perimeters for ophthalmologists to utilize in visual field assessments.[4,5]

Interpretation of EVF test results in the assessment of driving ability is variable. In 1999, authors Anderson and Patella stated that drivers should have binocular visual fields extending $\geq 50^\circ$ both to the right and to the left of fixation in their book *Automated Static Perimetry*.[2] However, each country or state has its own rules on what degree of field loss is acceptable. In the United Kingdom and Australia, a loss of up to 3 adjoining points within the central 20° radius is allowed while only 1 missed point is allowed in the Netherlands.[4] In the United States, states may choose whether to include a visual field requirement with specifications such as "130° of uninterrupted horizontal visual field" in Florida or "120° in the horizontal meridian with at least 30° in the nasal field of 1 eye" in New Mexico.[6,7] Some states may also have visual acuity requirements or may ask questions about monocularity, diplopia, impaired night vision, or history of retinitis pigmentosa, for instance.[8]

With its widespread use in the assessment of visual fields and driving, the EVF test is commonly used for visually impaired patients, and portability with a virtual reality device may provide greater accessibility and convenience of the test. In current practice, the EVF test is performed using stationary perimetry machines located in health care facilities. Thus, individuals are often required to make an extra trip to a physician to be tested. Cost of standard perimeters and inability to accommodate multiple patients at once in current testing environments can further contribute to challenges for implementation of the EVF in ophthalmology practices. Head-mounted virtual reality headsets have been used in the assessment of a variety of ophthalmic disorders, such as ocular deviation in strabismus patients, standard perimetry in glaucoma evaluation, or superior visual field testing in patients with ptosis.[9–11] No studies have shown the use of head-mounted virtual reality headsets for the implementation of the EVF test, which could allow drivers the opportunity to perform the EVF test at public settings such as a motor vehicle office and increase user convenience. This pilot study aims to assess the use of a virtual reality visual fields (VRVF) headset for implementation of the EVF test to obtain results comparable to those from SAP among a sample of participants with glaucoma.

Methods

Patients ages 10 to 90 with mild or moderate-severe glaucoma presenting to a glaucoma clinic at the Bascom Palmer Eye Institute were eligible for the study. Patients were classified as having mild glaucoma or moderate-severe glaucoma based on the treating clinician's determination of glaucoma severity from clinical presentation and prior SAP visual field testing. New patients were excluded from the study because they did not have baseline SAP exams on chart review to determine glaucoma severity. Patients presenting for a post-operative exam were also excluded to avoid the risk of injury to the operative site with the visual field headset. All methods were approved by the University of Miami Institutional Review Board and adhere to the tenets of the Declaration of Helsinki. All participants provided written consent. Individuals below the age of 18 provided assent with consent provided by a parent or guardian. Funding was received for this study and is detailed in the disclosures section at the end of the manuscript.

Included individuals were randomized to perform the test on either VRVF or SAP first. Instructions were given by the researcher for the VRVF test and by an ophthalmic technician for the SAP binocular static Esterman test as per standard perimetry guidelines. The EVF test is performed by displaying a central point which functions as the subject's focus while intermittently displaying peripheral points for which the subject should click a controller when seen. Eyeglasses were worn during testing if the participant normally wore glasses. For the VRVF test, the subject was given a controller that was virtually connected to the headset and clicked the "start" button for a short instructional video to begin. Upon initiation of the EVF test, the VRVF device would display the test points on a monitor inside the headset based on a program engineered by Virtual Vision Health, and participants were asked to click the button when they notice a point displayed in the periphery during the test.[12] Test parameters such as duration of test, number of correct points, number of missed points, false positives, false negatives, and fixation losses for the VRVF device were uploaded to a secure, cloud-based web platform for researchers to review after completion of the test. The same test parameters were obtained from SAP tests as well. Each participant completed both the VRVF and SAP test. If an individual did not complete both the VRVF and SAP test after initial consent to participate in the pilot study, the individual was excluded from further analysis. Participants were recruited over the course of 1 month from approximately October 14, 2022 to November 16, 2022. Recruitment was stopped following this period so evaluation of the pilot device could be performed, and researchers could undergo subsequent device improvement for future studies. This pilot study discusses the findings of the initial period of data collection.

Test results were analyzed with a Kappa statistic for clustered matched-paired data that evaluated the concordance of EVF test results on SAP versus VRVF.[13] In order to calculate this, each of the 120 tested points on each EVF test was converted into "0" for a point that was not correctly identified during the test and "1" for a point that was identified correctly. If the VRVF device failed to obtain all 120 points and only obtained 119 points, which occasionally occurred with the device, the point not obtained was excluded from analysis because no comparison could be made. Since all point locations remained the same on both devices, overall point-by-point concordance between the 2 devices could be examined. Concordance between points in 6 regions of the EVF grid was also examined based on a plot created by a 2013 Korean study.[14] In the study, binocular EVF test was divided into 6 clusters: upper center 10', lower center 10', upper center 30', lower center 30', upper periphery, and lower periphery. Within each region, the correlation between the Esterman efficiency score (EES) and the score on a visual function questionnaire regarding patients' vision-specific quality of life was calculated.

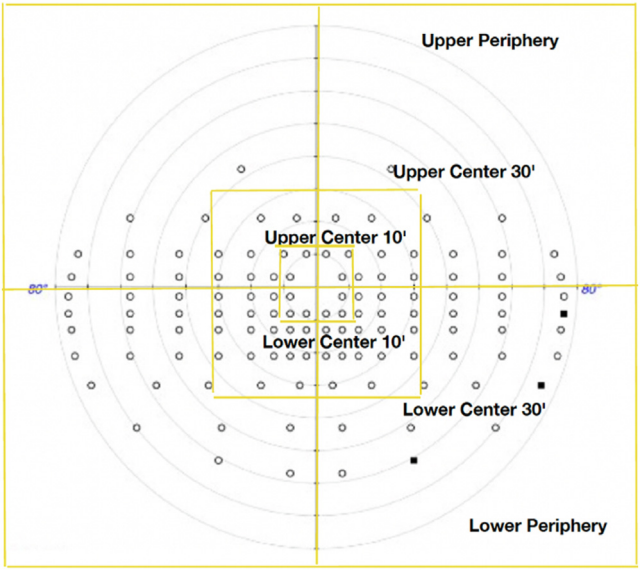

Figure 1. Standard automated perimetry Esterman visual field test divided into 6 regions. *Based on a graphic created by Lee, J.Y., Cho, H.K., Kee, C., *Assessment of the Vision-Specific Quality of Life Using Binocular Esterman Visual Field in Glaucoma Patients*. Journal of the Korean Ophthalmological Society, 2013.

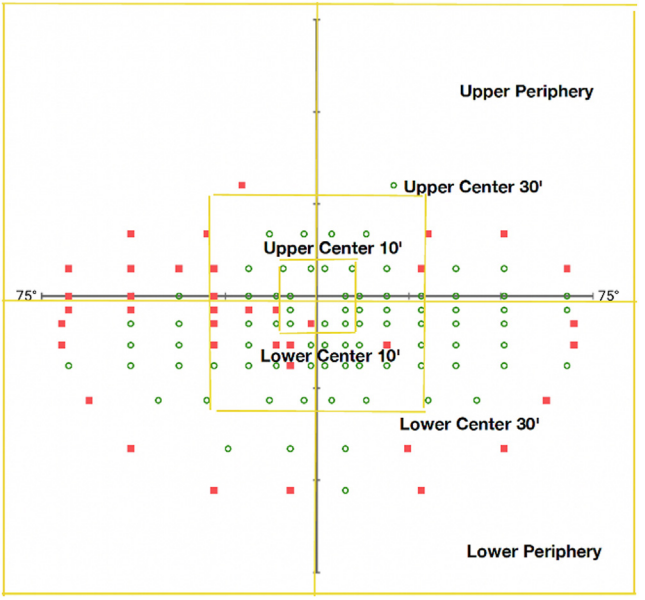

Figure 2. Virtual reality perimetry Esterman visual field test divided into 6 regions. *Based on a graphic created by Lee, J.Y., Cho, H.K., Kee, C., *Assessment of the Vision-Specific Quality of Life Using Binocular Esterman Visual Field in Glaucoma Patients*. Journal of the Korean Ophthalmological Society, 2013.

In order to calculate correlations between VRVF and SAP by region in the present study, the plot used in the 2013 study was used for analysis of both the SAP (Fig 1) and VRVF (Fig 2) tests.

Following all testing, VRVF and SAP test results were anonymized and randomized. Glaucoma-trained ophthalmologists were asked to grade all tests as "pass" or "failure" based on Florida state law requiring 130 degrees of uninterrupted horizontal visual field on the EVF test. Each test given a pass was converted to a "1" and a failure was converted to a "0" for further statistical analysis. Concordance between the 2 devices was then analyzed using the pass-failure determinations from ophthalmologists to evaluate whether ophthalmologists had the same determination of either a "pass" or a "failure" on both the SAP test and VRVF test for each subject. This was measured using a conditional logistic regression.[15] Inter-rater agreement was also calculated using Light's κ statistic to observe the agreement between different ophthalmologists on "pass" or "failure" determinations for SAP versus VRVF.[16] The bias-corrected and accelerated bootstrap was used to obtain the 95% confidence intervals (CIs) for Light's κ statistic.

All EVF tests were further analyzed based on the EES, a unit of measurement for the EVF test that is a calculation of the number of correct points divided by the total number of points displayed on the test.[4,17] The Wilcoxon signed rank test was used to compare EES and duration of testing between VRVF and SAP for patients characterized as having mild glaucoma, patients with moderate-severe glaucoma, and all patients regardless of disease severity and to calculate a κ value for each. Additionally, the Wilcoxon signed rank test was conducted to check if a difference between SAP EES and VRVF EES existed within each randomization group (VRVF test performed first or SAP test performed first) or if a difference existed between the EES of the first and second visual field test to evaluate whether the order of the visual field test performed played a role in the participant's performance. The data were analyzed using SAS version 9.4 (SAS Institute) and R version 4.3.1 with the psy, boot, and survival packages.[18–22]

Results

Of the 26 patients who initially consented to participate in the study, 4 were excluded from the final analysis as they did not complete both the VRVF and SAP tests (3 did not complete both tests) or withdrew their consent to participate (1 withdrew). Twenty-two subjects were included in the final analysis with ages ranging from 14 to 78 years old. Eleven had mild glaucoma and 11 had moderate-severe glaucoma. Figure 3 shows an example of EVF test results from 1 subject on SAP (left) and VRVF (right). The most common pass-fail determination from the 5 glaucoma specialists was determined for each VRVF and SAP test. Table 1 shows the congruency between the pass-fail determinations on SAP versus VRVF based on the most common determinations. Using SAP as the gold standard, the false positive rate was calculated at 20.0% and the false negative rate was calculated at 35.3% (Table 1).

Measurements of point-by-point concordance between VRVF and SAP test results showed fair agreement between the tests ($\kappa = 0.332$, [95% CI: 0.157, 0.506], Table 2). When analyzed by region, the highest concordance was found in the upper center 10' ($\kappa = 0.455$, [95% CI: 0.139, 0.771]) while the lowest concordance was found in the lower center 10' ($\kappa = 0.051$, [95% CI: −0.064, 0.166], Table 2). Concordance between VRVF and SAP using pass-failure determinations from ophthalmologists demonstrated better agreement ($\kappa = 0.657$, [95% CI: 0.549, 0.751], Table 2). While there was still no complete agreement between VRVF and SAP using pass-failure determinations, ophthalmologists were more likely to agree amongst themselves on pass-failure determinations for

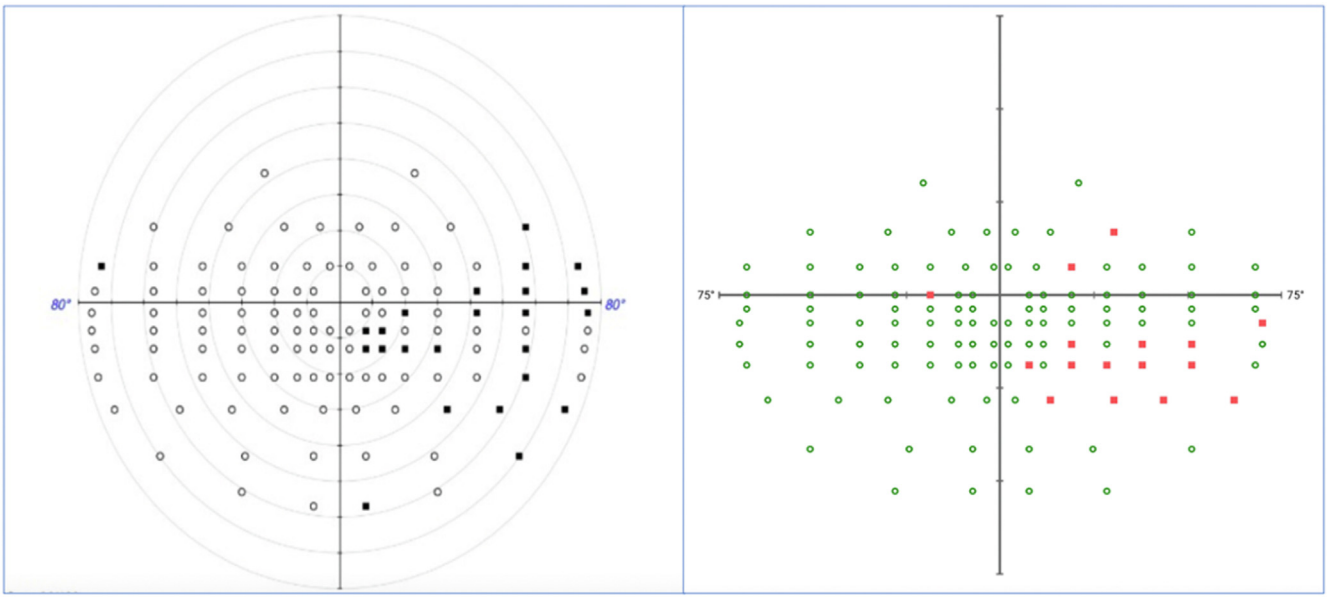

Figure 3. SAP and VRVF graphs obtained for a patient with moderate-to-severe glaucoma. For the SAP test, white points are correct points and black points are missed points. For the VRVF test, green points are correct points and red points are missed points. EES = Esterman efficiency score; SAP = standard automated perimetry; VRVF = virtual reality visual field.

VRVF tests ($\kappa = 0.890$, [95% CI: 0.726, 0.964]) than for SAP ($\kappa = 0.590$, [95% CI: 0.372, 0.818], Table 3). After results were stratified by severity, compared with SAP, there was a significant increase in interrater agreement on VRVF for mild glaucoma subjects (VRVF $\kappa = 0.917$, 95% CI: [0.661, 1.00]; SAP $\kappa = 0.467$, 95% CI: [0.258, 0.552]) and an increase in interrater agreement on VRVF for moderate-severe glaucoma subjects (VRVF $\kappa = 0.785$, 95% CI: [0.600, 1.00]; SAP $\kappa = 0.601$, 95% CI [0.302, 0.848], Table 3), but with no statistical significance.

Further analysis of test results demonstrated a significant median difference in EES with SAP capturing higher scores than VRVF (median EES difference: 4.5 points, $P = 0.021$, Table 4). This difference seemed to stem from measurements belonging to the moderate-severe glaucoma patients (median EES difference: 12 points, $P = 0.007$, Table 4), as there was no significant median difference in EES measurements from the 2 devices among mild glaucoma patients (median EES difference: 0 points, $P = 0.800$, Table 4). Notably, the significant median difference in EES taken by the SAP and VRVF devices was also present among the group randomized to VRVF first (median EES difference: 10.5 points, $P = 0.035$, data not displayed in tables) while those randomized to SAP first experienced no significant median difference in EES between the devices (median EES difference: 0.5 points, $P = 0.47$, data not displayed in tables). Yet, overall, the median difference between the first and second visual field test participants completed was not significant (median EES difference: −2.5, P=0.160, data not displayed in tables). There was also no significant median difference in duration of time between the SAP and VRVF devices in all patients (median difference in duration: −11 seconds, $P = 0.148$, Table 4) nor in mild (median difference in duration: −6 seconds, $P = 0.288$, Table 4) or moderate-severe (median difference in duration: −15 seconds, $P = 0.413$, Table 4) glaucoma patients.

Table 1. SAP vs. VRVF Headset Pass-Fail Determination Based on the Most Common Pass-Fail Determination by 5 Glaucoma Specialists for Each Participant (N = 22)

	Passed VRVF	Failed VRVF	Total
Passed SAP (Gold Standard)	11 (50.0%)	6 (27.3%)	17 (77.3%)
Failed SAP (Gold Standard)	1 (4.6%)	4 (18.2%)	5 (22.7%)
Total	12 (54.6%)	10 (45.5%)	22 (100.0%)
	False Positive Rate	**20.0%**	
	False Negative Rate	**35.3%**	

SAP = standard automated perimetry; VRVF = virtual reality visual field.

Table 2. Concordances Between Virtual Reality Visual Field Headset and Standard Automated Perimetry: By Overall Points, by Points in 6 Regions, and by Pass-Fail Determinations

	Kappa Statistic	95% Confidence Interval
Overall points	0.332	[0.157, 0.506]
Points by region: upper periphery	0.394	[0.214, 0.574]
Points by region: upper center 30'	0.282	[0.104, 0.459]
Points by region: upper center 10'	0.455	[0.139, 0.771]
Points by region: lower center 10'	0.051	[−0.064, 0.166]
Points by region: lower center 30'	0.250	[0.006, 0.494]
Points by region: lower periphery	0.330	[0.117, 0.543]
Pass-fail determinations	0.657	[0.549, 0.751]

Discussion

This pilot study is the first that assesses the implementation of the EVF test using a virtual reality headset and did not find strong concordance between SAP and VRVF results based on point-by-point analysis for overall points, point-by-point analysis for each of the 6 regions, or pass-failure determinations from ophthalmologists. Previous research on virtual reality devices has found variable results regarding the agreement between virtual reality perimetry and standard perimetry assessments. A 2023 study by Terracciano et al found that a virtual reality device displayed results that were in good correlation with a gold standard perimeter in healthy subjects. However, unlike this pilot study, the Terracciano et al study tested healthy subjects and also assessed agreement with Pearson's correlation coefficient rather than with κ values.[23] On the contrary, other studies have found discrepancies between the 2 testing modalities. In a 2021 study from Switzerland, a virtual reality perimeter was found to slightly underestimate visual field defects in glaucoma subjects.[10] In 2022, a study assessing another virtual reality perimeter concluded that the tested device misclassified 28% of moderate glaucoma cases as mild, 17% of moderate glaucoma cases as severe, and 20% of severe cases as moderate.[24] The present study found that VRVF significantly overestimated visual field defects with a significantly lower median EES on VRVF in glaucoma subjects, particularly moderate-severe glaucoma patients, when compared to SAP tests. This suggests that the 2 tests are not yet comparable.

The discrepancy between the 2 devices may be explained by several factors. One limitation of the study was that no test reports were excluded due to false negatives, false positives, and fixation losses, which may have created differences in VRVF and SAP results. The study was also limited to a small sample from a single institution. Participants may not have understood the instructions on 1 of the 2 devices—particularly on the VRVF device. Patients who are unfamiliar with virtual reality technology may have had less of an understanding of the VRVF platform, leading patients to perform worse on VRVF. Researchers attempted to minimize this limitation by randomizing order of VRVF or SAP testing. Even though subjects did not have considerably different first and second visual field tests, the group randomized to VRVF first had a significantly lower EES score on the VRVF headset than on SAP. This may be explained by the fact that users taking the VRVF test first experienced a greater learning curve of trying to learn both the visual field test and virtual reality test at the same time. Contrarily, cases that began with SAP could become acquainted with the visual field test prior to familiarizing themselves with the virtual reality device. Many patients may have also had previous familiarity with SAP visual field testing from previous office visits. While this may explain some of the differences in SAP and VRVF, it does not explain why only moderate-severe glaucoma cases had significantly lower VRVF EES than SAP EES. A limitation with the pass-failure determinations by ophthalmologists is that test type (SAP or VRVF) could not be blinded because the results are displayed in different formats on the 2 devices. Finally, this study only examines glaucomatous individuals. It is possible that results between the SAP and VRVF devices may more closely align when testing healthy individuals, particularly after seeing closer associations between the 2 devices with mild glaucoma patients as compared with severe glaucoma patients, and a future study may include healthy individuals in the study population to test this.

While the EVF test is commonly used in the assessment of driving ability, there is significant debate regarding whether the EVF test should be used at all. As the first known method to quantify binocular visual field loss, the EVF test showed significant promise in assessing how visual field loss could impact activities of daily living such as driving. Yet, Dr Benjamin Esterman did not develop the EVF test with the purpose of testing driving ability. The lack of standardization in test point positions and duration of testing poses a challenge to ensuring consistent tests. Researchers in Norway developed a binocular suprathreshold perimetry algorithm for group 1 driving licenses (car and motorcycle) in Europe to address the need for a uniform traffic perimetry algorithm and help standardize practices in enforcing visual field regulations. With its carefully chosen test pattern and definition of a positive test result, the program demonstrated adherence to European visual field requirements for group 1 drivers; however, use of the test was limited to Europe.[25] Another issue with using the EVF test is that studies have shown that some individuals who fail visual field criteria may still demonstrate safe driving behaviors or can pass other driving assessments.[26] Thus, some argue that a positive test result, or a failing EVF test, should be an indisputable assessment of driving performance, as denying an individual's driving ability could have a negative impact on a person's life and wellbeing.[25] An overestimation of visual field deficits may increase the number of people who

Table 3. Interrater Reliability Among 5 Raters for all Subjects, Mild Glaucoma Subjects, and Moderate-Severe Glaucoma Subjects on SAP vs. VRVF Headset

	Light's Kappa Statistic	95% Confidence Interval
All subjects (n = 22)		
SAP	0.590	[0.372, 0.818]
VRVF	0.890	[0.726, 0.964]
Mild glaucoma (n = 11)		
SAP	0.467	[0.258, 0.552]
VRVF	0.917	[0.661, 1.00]
Moderate-severe glaucoma (n = 11)		
SAP	0.601	[0.302, 0.848]
VRVF	0.785	[0.600, 1.00]

SAP = standard automated perimetry; VRVF = virtual reality visual field.

may not be able to obtain a license or have their license revoked. Thus, is it important to perform further research on the VRVF device in order to obtain increased concordance between the VRVF and SAP devices.

Several other traffic perimetry algorithms have been tested to develop an alternative to the EVF test to better assess driving ability beyond standard visual field testing. Prior research demonstrates that visual field test results from the standard visual field tests such as the Goldmann and Esterman tests may not be predictive of driving performance for participants with and without visual field loss when compared with an on-road driving assessment.[26] This is partially because the EVF test cannot account for certain techniques that visually impaired individuals can use to compensate for their impairment, such as turning their heads to the direction of better sight or moving their eyes. These actions cannot be performed during the EVF test in which the individual must look at a solitary point throughout the duration of testing without turning his or her head. It is possible that glaucoma patients, the population of this study, utilize compensatory mechanisms when driving that they could not display during the EVF tests. Thus, recent studies have tested the use of pupil tracking devices in traffic perimetry algorithms to analyze how participants view traffic scenes on a virtual screen without restriction of head or eye movements.[27] A study from the United Kingdom specifically looked at the eye movements of glaucomatous patients when viewing driving scenes in a hazard perception test and noted significant differences in eye movements among patients with glaucoma when compared with age-matched controls.[28] While novel traffic perimetry algorithms such as these may show future utility, none have been widely accepted. For this reason, the EVF test remains the standard of driving assessment.[4] However, based on the weak overall agreement between VRVF and SAP, the current VRVF EVF test is not currently an acceptable determinant of driver's licensing.

With significant debate on the use of the EVF test, future studies may assess whether a VRVF device could be an

Table 4. Quantitative Analysis of EES and Duration of Testing for all Subjects, Mild Glaucoma Subjects, and Moderate-Severe Glaucoma Subjects on SAP vs. VRVF

	Variable	Median (IQR)	Mean (SD)	Median Difference	P Value*
All subjects	SAP EES	92.5 (77.0–99.0)	85.4 (18.4)	4.5	0.021†
	VRVF EES	82.5 (56.0–97.0)	74.1 (24.8)		
All subjects	SAP duration (in sec)	262.5 (249.0–317.0)	289.8 (62.5)	−11	0.148
	VRVF duration (in sec)	293.5 (254.0–350.0)	302.0 (47.4)		
Mild glaucoma	SAP EES	95.0 (78.0–100.0)	90.8 (10.9)	0	0.800
	VRVF EES	93.0 (82.0–99.0)	86.7 (21.5)		
Moderate–severe glaucoma	SAP EES	88.0 (75.0–99.0)	80.0 (23.1)	12	0.007‡
	VRVF EES	59.0 (49.0–83.0)	61.5 (21.9)		
Mild glaucoma	SAP duration (in sec)	252.0 (240.0–294.0)	265.0 (31.4)	−6	0.288
	VRVF duration (in sec)	270.0 (246.0–294.0)	272.7 (35.6)		
Moderate-severe glaucoma	SAP duration (in sec)	284.0 (260.0–366.0)	314.5 (76.6)	−15	0.413
	VRVF duration (in sec)	343.0 (293.0–357.0)	331.2 (39.6)		

EES = Esterman efficiency score; EVF = Esterman visual field; IQR = interquartile range; SAP = standard automated perimetry; SD = standard deviation; VRVF = virtual reality visual field headset.
Median difference: SAP measurements − VRVF measurements.
*Wilcoxon Signed rank test.
†<0.05.
‡<0.01.

acceptable screening tool for visual field deficits that may impact driving performance. This study did not find that the VRVF device was comparable to SAP for the EVF test. Although this VRVF study overestimated visual field loss when compared with SAP, physicians were more likely to agree amongst themselves on VRVF reports. However, this study is only a pilot study, and the results of this study are not yet generalizable. Further testing must be performed to obtain better concordance between the 2 devices. Also, some of the limitations with this study should be addressed. A limitation that the study team noticed was that many patients may have had difficulty with using virtual reality technology, and this may have led to the overall low EES scores observed on the VRVF device. Future studies should ensure that all participants understand VRVF technology by, for instance, undergoing a standardized trial of the device prior to undergoing the test for the study. In order for this device to eventually be easily used in a real-world setting, increased instruction on VRVF may be necessary. While the device already plays an instructional video prior to beginning the test, the instructional video may need to be amended to explain the use of the device more thoroughly, and it may need to include a thorough sample test as well. A future study may also examine test-retest reliability after each participant completes the VRVF test twice and the SAP test twice to see whether scores are stable between the first and second test on each respective device or if users improve on the second try with increased practice and familiarity. If significant differences still exist between the 2 devices after ensuring proper VRVF training, changes to the VRVF software may be necessary prior to retesting the device.

The EVF remains the standard for driving assessment across many countries, including the United States. Virtual reality visual field devices may offer a portable and virtual alternative to SAP, increasing user convenience. This pilot study is the first, to our knowledge, to assess the implementation of the EVF test using a virtual reality headset but did not find strong concordance between the 2 devices. The present device showed only fair concordance between SAP and VRVF results among tested individuals after point-by-point analysis (0.332 [95% CI: 0.157, 0.506]) and somewhat better concordance after pass-failure analysis (0.657 [95% CI: 0.549, 0.751]), as VRVF tended to overestimate visual field defects (median EES difference: 4.5 points, $P = 0.021$). Nevertheless, ophthalmologists were more likely to agree amongst themselves on VRVF test reports ($\kappa = 0.890$, [95% CI: 0.726, 0.964]) than on SAP reports. While results on the 2 devices differed, the pilot study showed that glaucoma patients may successfully complete EVF testing with a virtual reality device and shows promise for implementing the EVF test using virtual reality devices in the future. With further device testing and evaluation, virtual reality device use may become an acceptable, portable, and more convenient method to screen for individuals who may require further visual field evaluation by an ophthalmologist.

Footnotes and Disclosures

Originally received: September 29, 2023.
Final revision: March 27, 2024.
Accepted: April 8, 2024.
Available online: April 16, 2024. Manuscript no. XOPS-D-23-00236.

Bascom Palmer Eye Institute at the University of Miami Miller School of Medicine, 900 NW 17 Street, Miami, Florida 33136.

Presented at the World Glaucoma Conference 2023, Association for Research in Vision & Ophthalmology 2023.

Disclosure(s):

All authors have completed and submitted the ICMJE disclosures form.

The authors made the following disclosures:

M. S.: Support - University of Miami Dean's Research Excellence Award in Medicine scholarship.

A. G.: Support - Virtual Vision Health (one VRVF Device purchased); Patents - Total Vision/Virtual Vision Health (co-inventor); Leadership - Vision Health Device (Medical Director); Shares - Virtual Vision Health; Receipt of equipment - Virtual Vision Health.

L. H.: Support - Bascom Palmer Eye Institute at the University of Miami Miller School of Medicine.

R. O. B.: Support - Research to Prevent Blindness (unrestricted grant [GR004596-1], the Department of Ophthalmology, University of Miami Miller School of Medicine); Consultant - Character Biosciences, Inc (nonpaid consulting [to date]).

HUMAN SUBJECTS: Human subjects were included in this study. All methods were approved by the University of Miami Institutional Review Board and adhere to the tenets of the Declaration of Helsinki. All participants provided written consent. Individuals below the age of 18 provided assent with consent provided by a parent or guardian.

No animal subjects were included in this study.

Author Contributions:

Conception and design: Sharma, Savatovsky, Grajewski, Bitrian

Analysis and interpretation: Sharma, Savatovsky, Huertas, O'Brien

Data collection: Sharma, Savatovsky

Obtained funding: Sharma, Grajewski

Overall responsibility: Sharma, Savatovsky, Huertas, O'Brien, Grajewski, Bitrian

Abbreviations and Acronyms:
CI = confidence interval; **EES** = Esterman efficiency score; **EVF** = Esterman visual field; **SAP** = standard automated perimetry; **VRVF** = virtual reality visual field headset.

Keywords:
virtual reality, visual field headset, virtual reality visual fields, Esterman visual field, driving assessment.

Correspondence:
Eleonore Savatovsky, MD, PhD, The Samuel and Ethel International Pediatric Glaucoma Center, Bascom Palmer Eye Institute, 900 NW 17th Street, Miami, FL 33136. E-mail: esavatovsky@med.miami.edu.

References

1. Faraji Y, Tan-Burghouwt MT, Bredewoud RA, et al. Predictive value of the esterman visual field test on the outcome of the on-road driving test. *Transl Vis Sci Technol*. 2022;11(3):20.
2. Heijl A, Patella VM, Bengtsson B. *Excellent Perimetry. 5th ed*. Dublin, CA: Carl Zeiss Meditec, Inc; 2021.
3. Jampel HD, Friedman DS, Quigley J, et al. Correlation of the binocular visual field with patient assessment of vision. *Invest Ophthalmol Vis Sci*. 2002;43(4):1059−1067.
4. Bro T. Benjamin Esterman (1906-1994) and the binocular visual field scoring grid that became a world standard for assessing driver eligibility. *Acta Ophthalmol*. 2022;100(7):828−833.
5. Crabb DP, Viswanathan AC. Integrated visual fields: a new approach to measuring the binocular field of view and visual disability. *Graefes Arch Clin Exp Ophthalmol*. 2005;243(3):210−216.
6. General Information. Medical or visual problems. https://www.flhsmv.gov/driver-licenses-id-cards/general-information/medical-visual-problems/. Accessed May 1, 2023.
7. Vision report. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwiawtfQsNbcAhVo2oMKHTkqD38QFjAAegQIABAC&url=http%3A%2F%2Frealfile.tax.newmexico.gov%2Fmvd10124.pdf&usg=AOvVaw176SYcvCbc9kvYqrZtV13t. Accessed May 1, 2023.
8. Steinkuller PG. Legal vision requirements for drivers in the United States. *Virtual Mentor*. 2010;12(12):938−940.
9. Yeh PH, Liu CH, Sun MH, et al. To measure the amount of ocular deviation in strabismus patients with an eye-tracking virtual reality headset. *BMC Ophthalmol*. 2021;21(1):246.
10. Stapelfeldt J, Kucur SS, Huber N, et al. Virtual reality-based and conventional visual field examination comparison in healthy and glaucoma patients. *Transl Vis Sci Technol*. 2021;10(12):10.
11. Patel AJ, Lee WW, Ziff M, et al. Superior visual field testing using virtual reality with and without eye tracking for functional upper eyelid surgery evaluation: a pilot study. *Ophthalmic Plast Reconstr Surg*. 2023;39:381−385.
12. Virtual vision. https://virtualvision.health/about-us/; 2023. Accessed May 1, 2023.
13. Yang Z, Zhou M. Kappa statistic for clustered matched-pair data. *Stat Med*. 2014;33(15):2612−2633.
14. Lee JY, Cho HK, Kee C. Assessment of the vision-specific quality of life using binocular esterman visual field in glaucoma patients. *J Korean Ophthalmol Soc*. 2013;54:1567−1572.
15. Therneau T, Atkinson E. The concordance statistic. A package for survival analysis in R, vignettes. https://cran.r-project.org/web/packages/survival/vignettes/concordance.pdf; 2022. Accessed July 28, 2023.
16. Conger AJ. Integration and generalisation of Kappas for multiple raters. *Psychol Bull*. 1980;88:322−328.
17. Stanojcic N, Wilkins M, Bunce C, et al. Visual fields in patients with multifocal intraocular lens implants and monovision: an exploratory study. *Eye (Lond)*. 2010;24(11):1645−1651.
18. Falissard B. Psy: various procedures used in psychometrics. 2022: R package version 1.2. https://CRAN.R-project.org/package=psy. Accessed August 1, 2023.
19. Canty A, Ripley B. boot: Bootstrap R (S-Plus) Functions, R package version 1.3-28.1. https://cran.r-project.org/web/packages/boot/. Accessed August 1, 2023.
20. Davison AC, Hinkley DV. *Bootstrap Methods and Their Applications*. Cambridge: Cambridge University Press; 1997.
21. Therneau T. A package for survival analysis in R. R package version 3.5-5. https://CRAN.R-project.org/package=survival; 2023. Accessed August 1, 2023.
22. Therneau TM, Grambsch PM. *Modeling Survival Data: Extending the Cox Model*. New York: Springer; 2000.
23. Terracciano R, Mascolo A, Venturo L, et al. Kinetic perimetry on virtual reality headset. *IEEE Trans Biomed Circuits Syst*. 2023;17(3):413−419.
24. Shetty V, Sankhe P, Haldipurkar SS, et al. Diagnostic performance of the PalmScan VF2000 virtual reality visual field analyzer for identification and classification of glaucoma. *J Ophthalmic Vis Res*. 2022;17(1):33−41.
25. Jorstad OK, Jonsdottir TE, Zysset S, et al. A traffic perimetry test that adheres to the European visual field requirements for group 2 drivers. *Acta Ophthalmol*. 2021;99(7):e1253−e1254.
26. Silveira S, Jolly N, Heard R, et al. Current licensing authority standards for peripheral visual field and safe on-road senior aged automobile driving performance. *Clin Exp Ophthalmol*. 2007;35(7):612−620.
27. Faraji Y, van Rijn JW, van Nispen RMA, et al. A toolkit for wide-screen dynamic area of interest measurements using the Pupil Labs Core Eye Tracker. *Behav Res Methods*. 2022;55:3820−3830.
28. Crabb DP, Smith ND, Rauscher FG, et al. Exploring eye movements in patients with glaucoma when viewing a driving scene. *PLoS One*. 2010;5(3):e9710.