Evaluation of Virtual Reality-Based Diagnostic Devices in Tertiary Care Ophthalmology

Catherine Dang

Thesis submitted to the University of Ottawa

in partial fulfillment of the requirements for the degree of

Master of Science in Neuroscience

Department of Cellular and Molecular Medicine

Faculty of Medicine

University of Ottawa

ABSTRACT

The integration of virtual reality (VR) technologies into clinical ophthalmic practice has accelerated in recent years. VR-based diagnostic platforms have the potential to provide comparable diagnostic performance relative to conventional ophthalmic testing while offering advantages in portability, cost-effectiveness, standardized test administration, and accessibility.

This comparative cross-sectional study evaluated the validity and diagnostic utility of two commercially available VR headsets integrated with two novel diagnostic applications, (i) Order of Magnitude (OM) and (ii) Retinalogik ©, for visual field and pupillary assessment relative to the current standards of care at the University of Ottawa Eye Institute. Primary visual field outcome measures included test duration, mean deviation (MD), and pattern standard deviation (PSD). Agreement between the VR-based systems and the reference standards was evaluated using Pearson correlation and Bland–Altman analysis, while pointwise threshold agreement was assessed through heatmap visualizations and spatial analyses. Pupillary assessment focused on the detection and categorical grading of a relative afferent pupillary defect (RAPD), with diagnostic concordance evaluated through a contingency table analysis. The VR-based diagnostic platforms demonstrated diagnostic performance comparable to the standard of care testing methods, supporting their diagnostic validity relative to conventional clinical methods.

Overall, this study demonstrated the feasibility and diagnostic validity of integrating VR-based diagnostic platforms into tertiary eye care settings, while highlighting their potential to expand access through remote and decentralized clinical workflows.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

CN III	Cranial nerve III
dB	Decibel
GHT	Glaucoma hemifield test
HFA	Humphrey field analyzer
HMD	Head-mounted display
LGN	Lateral geniculate nucleus
LOA	Limits of agreement
LogMAR	Logarithm of the minimum angle of resolution
LOESS	Locally estimated scatterplot smoothing
MD	Mean deviation
min	Minute
mmHg	Millimeters of mercury
NAION	Non-arteritic anterior ischemic optic neuropathy
ND	Neutral density
OM	Order of Magnitude
PERRLA	Pupils equal, round, reactive to light and accommodation
PLR	Pupillary light reflex
PSD	Pattern standard deviation
RAPD	Relative afferent pupillary defect
RNFL	Retinal nerve fiber layer
RVF	Retinalogik visual field
SAP	Standard automated perimetry

SD	Standard deviation
SITA	Swedish Interactive Thresholding Algorithm
3D	Three-dimensional
TD	Total deviation
VF	Visual field
VR	Virtual reality

ACKNOWLEDGEMENTS

I would like to express my utmost gratitude to my supervisors, Dr. Rustum Karanjia and Dr. Catherine Tsilfidis, for providing me with the opportunity to undertake this research project. Your expertise, mentorship, and invaluable guidance have been instrumental in shaping both this work and my development as a researcher and individual, making my time under your supervision an exceptionally rewarding learning experience.

I am also deeply grateful to Hong-An Nguyen and Dr. Soumya Podury for their mentorship, insight, and guidance throughout every stage of this research. Your insight, dedication to scientific inquiry, and generosity in sharing your knowledge have been truly inspiring. I greatly appreciate your patience and encouragement, which have significantly contributed to the completion of this work.

I am thankful to my thesis advisory committee member, Dr. Pierre Mattar, for your valuable and constructive feedback that has strengthened both the direction and quality of this thesis. My sincere appreciation also extends to my colleagues and friends, whose encouragement, collaboration, and companionship made this journey both enjoyable and memorable.

Finally, I wish to express my deepest gratitude to my family for their unwavering love, patience, and unconditional support. Their constant encouragement and belief in me have sustained me throughout this journey, and this achievement would not have been possible without them.

CHAPTER 1. INTRODUCTION

1.1. Overview of the Afferent Visual Pathway

The visual system is composed of a highly organized and specialized network responsible for the detection, transmission, and interpretation of visual sensory information. The intricacy of visual processing begins when light enters the eye and is focused on the retina at the macula (Gupta et al., 2026). Photoreceptors embedded in the retinal tissue layers convert light energy to electricals signals, which are transmitted via retinal ganglion cell axons forming the retinal nerve fiber layer. These nerve fibers converge to form the optic nerve, carrying the visual signals posteriorly towards the optic chiasm as shown in Figure 1.1. At the optic chiasm, the nasal retinal nerve fibers decussate to the contralateral side, while temporal fibers remain ipsilateral, marking the start of the optic tract. Continuing posteriorly, the optic tracts convey these signals to the lateral geniculate nucleus (LGN) of the thalamus, where they synapse onto relay neurons. From the LGN, the optic radiations transmit visual signals to the primary visual cortex in the occipital lobe, where initial cortical processing of visual input occurs.

The integrity of each component of the visual pathway is imperative for developing and maintaining normal visual function. Any structural anomaly that perturbs the system may result in visual field (VF) defects or abnormal pupillary function (e.g., relative afferent pupillary defects) (Kaeser & Kawasaki, 2010; Wall, 2021). Thus, the nature of visual field testing helps localize the area of pathology.

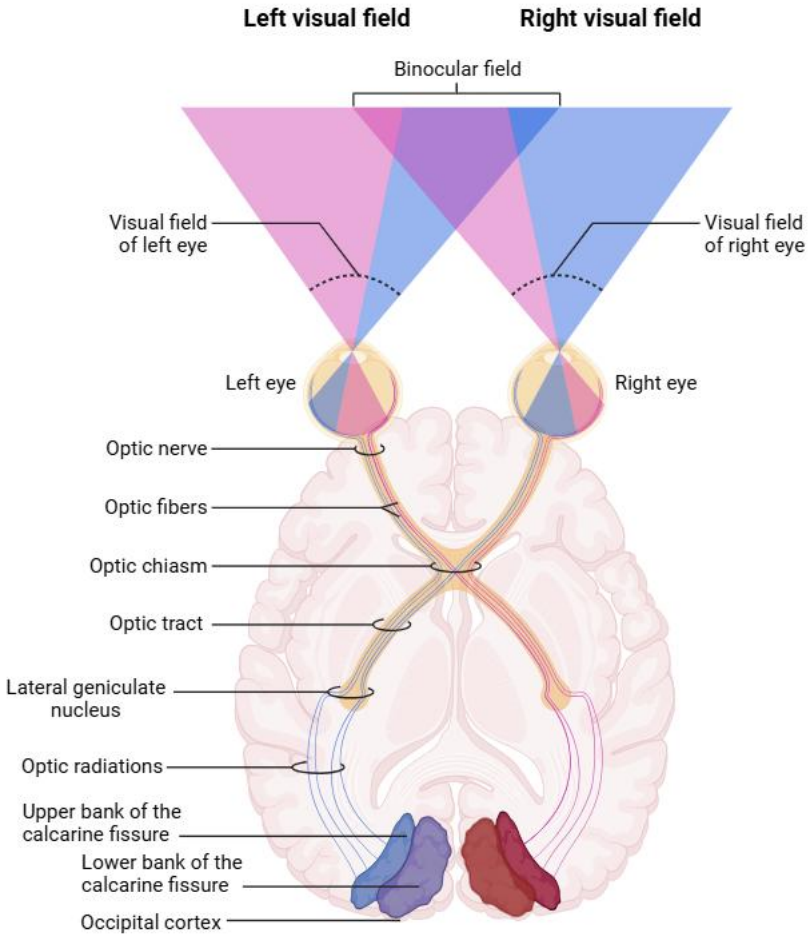


Figure 1.1. Schematic illustration of the afferent visual pathway. The purple and blue colors track the segregation of the left and right visual hemifields, respectively, with each pathway comprising of nerve fibers originating from both the nasal and temporal retina. Light from each visual hemifield is projected onto the nasal retina of one eye and the temporal retina of the fellow eye. Temporal retinal fibers remain ipsilateral, whereas nasal retinal fibers decussate at the optic chiasm, resulting in visual information from the left visual hemifield being processed in the right cerebral hemisphere and information from the right visual hemifield being processed in the left cerebral hemisphere. These fibers mark the beginning of the optic tracts as they synapse into the lateral geniculate nucleus (LGN), before continuing via the optic radiations to the primary visual cortex of the occipital lobe. Adapted from "Human Visual System", by BioRender.com (2026). Retrieved from https://app.biorender.com/biorender-templates.

1.1.1. Afferent Visual Pathway Lesions and Clinical Manifestations

Ocular and systemic diseases can produce abnormalities throughout the afferent visual pathway, resulting in visual field loss which, depending on the pattern of field loss, can localize the area of the structural abnormality. Retinal disorders, including age-related macular degeneration, diabetic retinopathy, and retinal vascular occlusions, typically produce localized monocular scotomas corresponding to the affected retinal regions (Kanski & Bowling, 2016). Similarly, optic nerve disorders, such as optic neuritis, ischemic and compressive optic neuropathies, and glaucoma, can produce monocular visual field defects (Miller et al., 2021). In glaucoma, progressive retinal ganglion cell loss and retinal nerve fiber layer (RNFL) thinning give rise to characteristic visual field defects that evolve with disease severity.

Visual field defects may manifest as generalized constriction, localized scotomas, altitudinal, or hemianopia defects, depending on the site of pathology (Figure 1.2 and 1.3). In glaucoma, early visual field defects typically include paracentral scotomas, enlargement of the physiological blind spot (Seidel scotoma), and arcuate (Bjerrum) scotomas that follow the anatomical trajectory of the retinal nerve fiber layer (RNFL) axonal bundles. As retinal ganglion cell (RGC) damage progresses, these defects enlarge and coalesce. Preferential involvement of either the superior or inferior arcuate nerve fiber bundles produces the characteristic nasal step, an abrupt reduction in retinal sensitivity across the horizontal meridian of the nasal visual field. Continued RGC loss results in progressive peripheral visual field constriction, with central vision often preserved until late stages as the papillomacular bundle is relatively resistant to early glaucomatous damage. In very advanced disease stages, only a small central island of vision may remain. Furthermore, generalized visual field constriction may occur in retinal degenerations and advanced glaucoma, whereas altitudinal defects are more characteristic of anterior ischemic optic

neuropathy, retinal vascular occlusions, and advanced glaucoma. Cecocentral scotomas, which involve both the fixation point and physiological blind spot, are typically associated with toxic, nutritional, hereditary, and demyelinating optic neuropathies and are only rarely observed in glaucoma.

While chiasmal lesions classically produce bitemporal hemianopia through disruption of the decussating nasal retinal fibers, primarily due to pituitary adenomas or other suprasellar masses, they can also very rarely induce a binasal hemianopia by involvement of the uncrossed temporal retinal fibers (Ruddy et al., 2024). Due to its rare occurrence, documented etiologies are diverse, spanning both intraocular and intracranial pathologies including bilateral internal carotid artery aneurysms, advanced glaucoma, intracranial masses, neurosyphilis, hydrocephalus, and elevated intracranial pressure (Bryan et al., 2014; Hamann et al., 2015; Ruddy et al., 2024). Lesions posterior to the optic chiasm produce contralateral homonymous visual field defects, reflecting the convergence of visual information from the ipsilateral temporal retina and contralateral nasal retina after decussation (Miller et al., 2021). Lesions of the optic tract and lateral geniculate nucleus typically cause homonymous hemianopia, whereas temporal lobe lesions involving Meyer's loop produce contralateral superior quadrantanopia, and parietal lobe lesions involving the superior optic radiations produce contralateral inferior quadrantanopia. Occipital lobe lesions classically result in complete contralateral homonymous hemianopia with macular sparing, attributable to the dual vascular supply of the occipital pole and the disproportionate cortical representation of central vision (Miller et al., 2021). Furthermore, the congruency of homonymous visual field defects generally increases with progressively more posterior lesions, providing an important anatomical correlate for lesion localization.

Consequently, the pattern of visual field loss provides valuable information for pathology localization and contributes to the diagnosis of both ocular and neurological disease. Visual field testing therefore serves as a fundamental functional assessment of the afferent visual pathway and remains an essential tool for the detection, localization, and monitoring of visual system pathology.

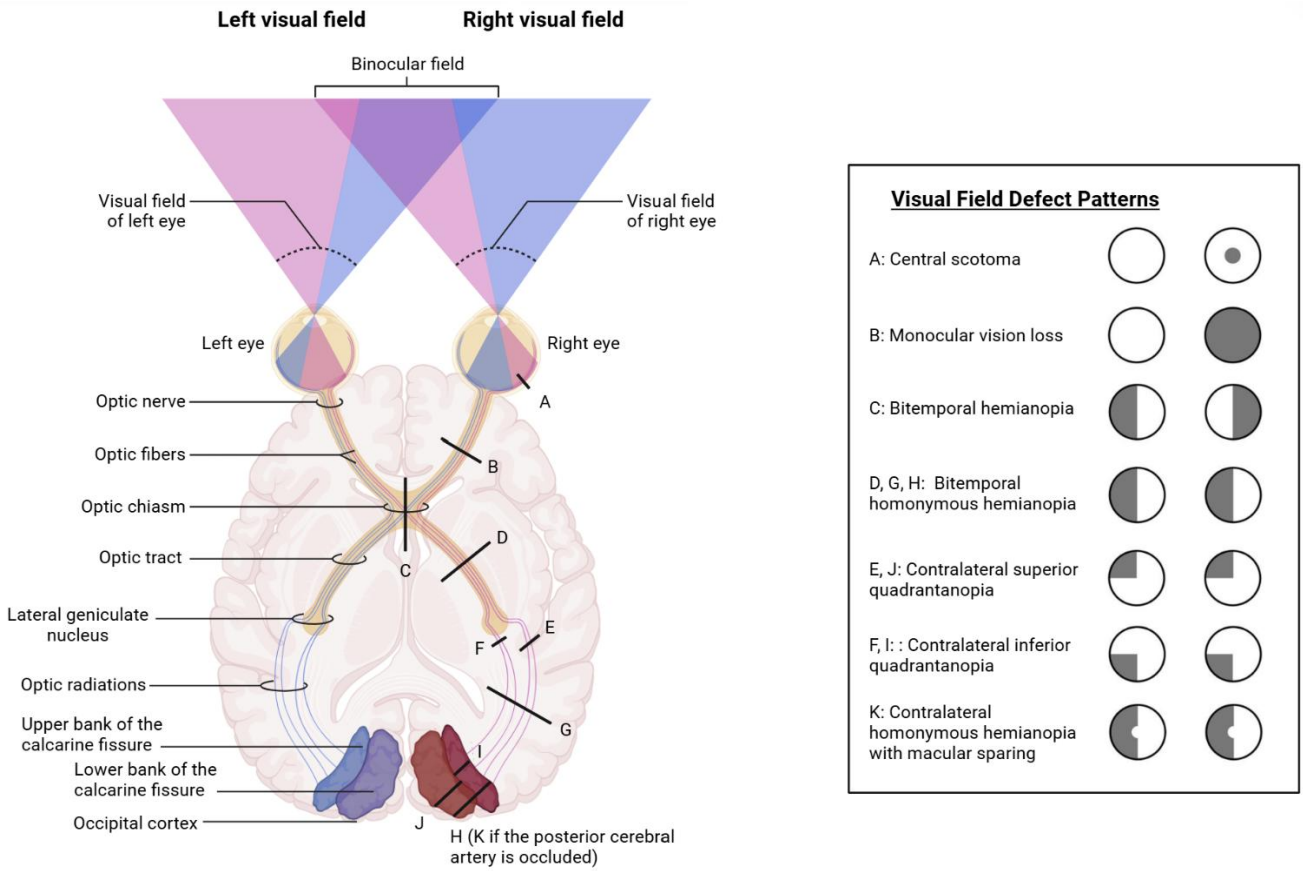


Figure 1.2. Visual field defects associated with structural lesions in the afferent visual pathway. Schematic representation of the anatomical organization of the afferent visual pathway, illustrating signal transmission from the retina through the optic nerve, optic chiasm, optic tracts, lateral geniculate nucleus, optic radiations, and the primary visual cortex. Adjacent to the pathway diagram is a chart illustrating characteristic visual field defects associated with specific lesion sites (labeled A–K). Adapted from "Human Visual System", by BioRender.com (2026). Retrieved from https://app.biorender.com/biorender-templates.

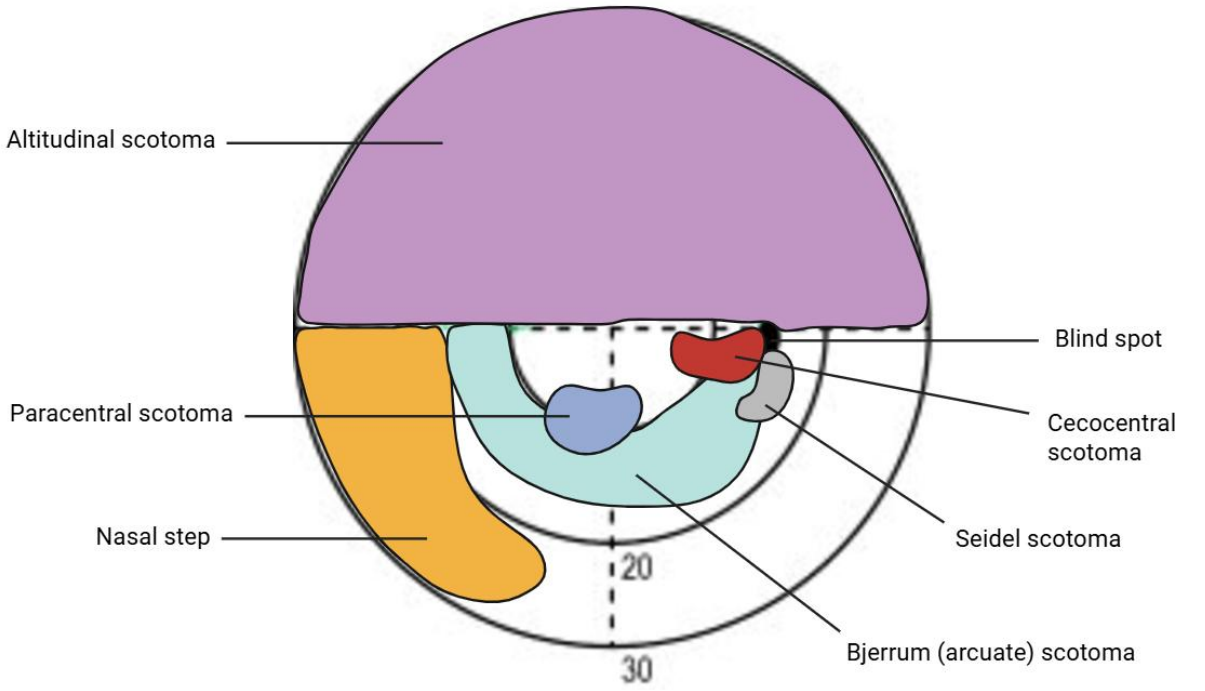


Figure 1.3. Schematic illustration the right eye visual field with common glaucomatous visual field defects. Characteristic retinal nerve fiber bundle loss patterns are illustrated, including the Bjerrum (arcuate) scotoma (teal), cecocentral scotoma (red), seidel scotoma (grey), paracentral scotoma (blue), nasal step (yellow), and altitudinal scotoma (purple). Created with BioRender.com.

1.2. The Field of Vision

The field of vision is the area of space seen while maintaining a steady fixation of gaze at a central fixation point. A normal monocular visual field extends approximately 60 degrees superiorly, 75 degrees inferiorly, 100 degrees temporally, and 60 degrees nasally (Roy et al., 1990) (Figure 1.4). The central visual field corresponds to the macular region in the retina, while peripheral visual fields are mediated by increasingly eccentric retinal areas. Within the monocular visual field, a physiological blind spot exists approximately 15° temporally and 1.5° below the horizontal meridian (Choplin & Edwards, 1998). This region corresponds anatomically to the optic disc, where the optic nerve exits the retina and where photoreceptors are absent.

Conceptually, the visual field is best represented as a three-dimensional "hill of vision", where the peak represents maximal retinal sensitivity at the fovea, and the slope gradually declines toward the peripheral retina (Katz & Sommer, 1986; Lewis et al., 2025) (Figure 1.5). Within this framework, the physiological blind spot appears as a vertical "crater", reflecting absolute loss of sensitivity at the optic disc. Pathological lesions along the visual pathway effectively "erode" portions of this hill; for instance, glaucomatous damage produces characteristic localized depressions in sensitivity, resulting in visual field defects such as arcuate scotomas and nasal steps that reflect retinal nerve fiber layer damage.

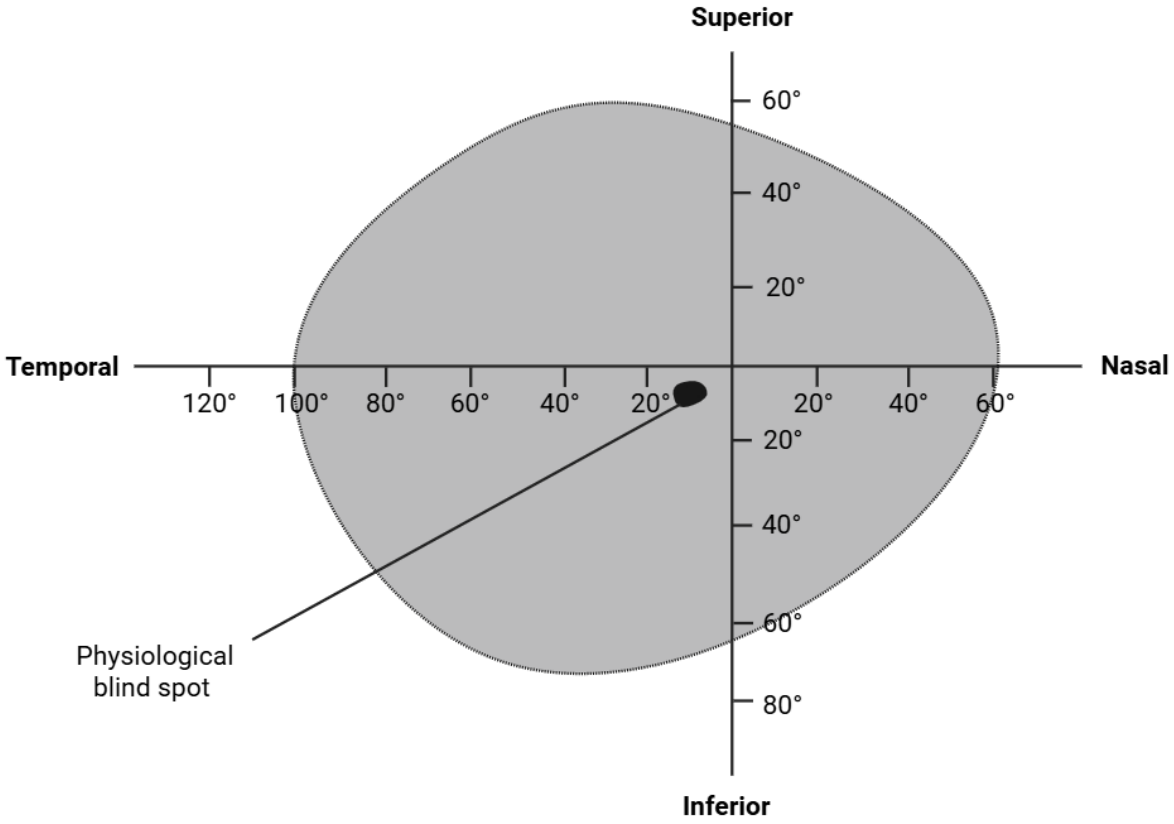


Figure 1.4. Boundaries of the normal monocular visual field. A schematic diagram illustrating a normal monocular visual field of the left eye. The diagram demonstrates the typical boundary extents (60 degrees superiorly, 75 degrees inferiorly, 100 degrees temporally, and 60 degrees nasally) and the approximate spatial location of the physiological blind spot (15° temporally and 1.5° below the horizontal meridian) relative to the central fixation point at (0, 0). Created with BioRender.com.

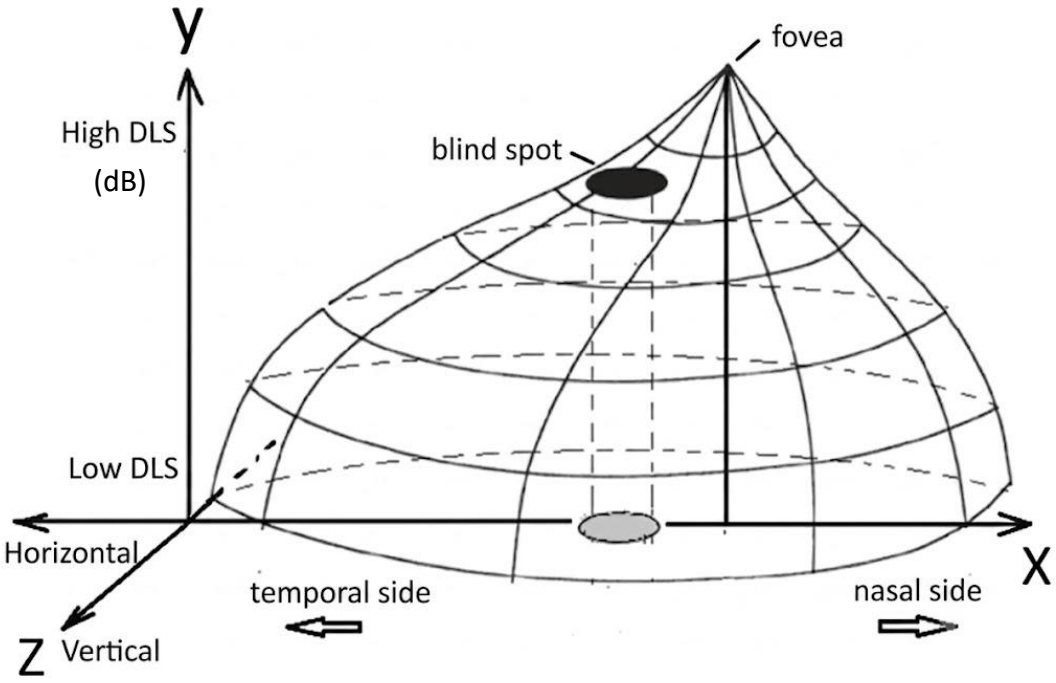


Figure 1.5. Schematic representation of the hill of vision on a Cartesian plane system. Three-dimensional representation of visual field sensitivity as a hill of vision. The vertical axis (*y*-axis) denotes differential retinal light sensitivity (dB), which peaks sharply at the foveal fixation point and declines progressively toward the peripheral boundaries. The physiological blind spot is depicted as an absolute pit extending to the baseline (*x*-axis) within the temporal slope. Modified from Choplin & Edwards, (1998); Chandrinos, (2017). DLS: Differential light sensitivity; dB: Decibel.

1.2.1. Standard Automated Perimetry

Quantifying topographic variations in retinal sensitivity requires standardized clinical assessments. Current modalities include confrontation visual field testing for rapid bedside screening, Goldmann perimetry for manual kinetic assessment, and standard automated perimetry (SAP) (Figure 1.6A–C). The most widely used SAP instrument, the Humphrey® Field Analyzer (HFA; Carl Zeiss Meditec, Dublin, CA, USA), evaluates differential light sensitivity at predefined test locations by presenting stationary white-light stimuli of varying luminance against a uniformly illuminated white background. Threshold sensitivity at each location is then determined to quantify functional visual field loss (Ruia & Tripathy, 2025).

Several testing patterns are available on the HFA, each designed to assess the central region of the visual field. Commonly used testing strategies include the 30-2, 24-2 and 10-2 test patterns. The 30-2 program evaluates the central 30° of the visual field using 76 test locations and is particularly valuable for detecting visual field defects associated with neuro-ophthalmic disorders (Choplin & Edwards, 1998; Ruia & Tripathy, 2025). The 24-2 program assesses the central 24° of the visual field using 54 test locations; it essentially reduces the number of superior, inferior, and temporal test points from the 30-2 grid, while preserving nasal test points critical for early glaucoma detection. Finally, the 10-2 program evaluates the central 10° of the visual field using 68 closely spaced test locations and is commonly used to assess macular function and advanced glaucomatous damage.

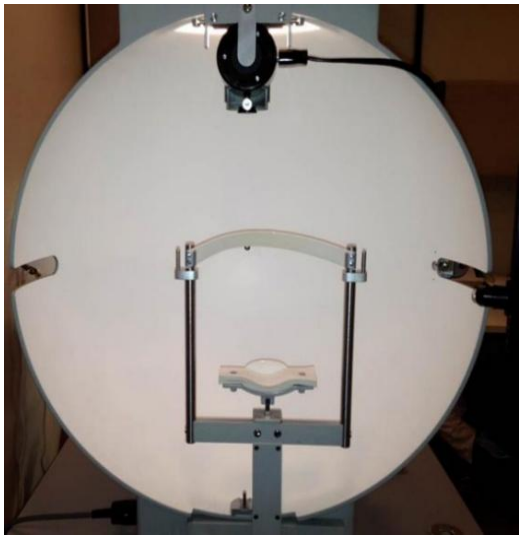
The Swedish Interactive Thresholding Algorithm (SITA) standard is a widely adopted testing strategy to estimate visual field sensitivity efficiently and reliably (J. Tan et al., 2025). It employs prior knowledge of expected threshold values, adaptive stimulus presentation, and real-time response modeling to reduce the number of presentations required at each test location while

maintaining measurement accuracy. By adjusting stimulus intensity based on patient responses, SITA standard shortens overall test duration compared with full-threshold methods and improves patient comfort without substantially compromising diagnostic precision. Sensitivity thresholds are reported in decibels (dB), representing the attenuation of stimulus intensity subjectively seen by the patient relative to the instrument's maximum luminance. The resulting threshold sensitivity measurements provide a quantitative assessment of visual function.

Clinicians routinely evaluate reliability metrics, including fixation losses, false-positive responses, and false-negative responses, as well as global maps and indices such as mean deviation (MD), pattern standard deviation (PSD), total deviation (TD) maps, and the glaucoma hemifield test (GHT) (Ruia & Tripathy, 2025). A representative sample of the HFA printouts illustrating these parameters is shown in Figure 1.6D. MD represents the average difference between a patient's measured visual field sensitivity and age-corrected normative values, providing an overall measure of diffuse visual field loss (Wall, 2021). More negative MD values indicate greater generalized depression of visual field sensitivity. In contrast, PSD quantifies localized irregularities in the visual field by measuring the variability of sensitivity values relative to the expected hill of visual field (Choplin & Edwards, 1998). Elevated PSD values are indicative of focal visual field defects. The TD map displays pointwise sensitivity relative to age-matched normative data, highlighting both diffuse and localized defects, while the GHT compares corresponding regions of the superior and inferior hemifields to identify glaucomatous patterns of asymmetry. Together, these parameters facilitate the identification, localization, characterization, and longitudinal monitoring of visual field abnormalities.

A

B



C

D

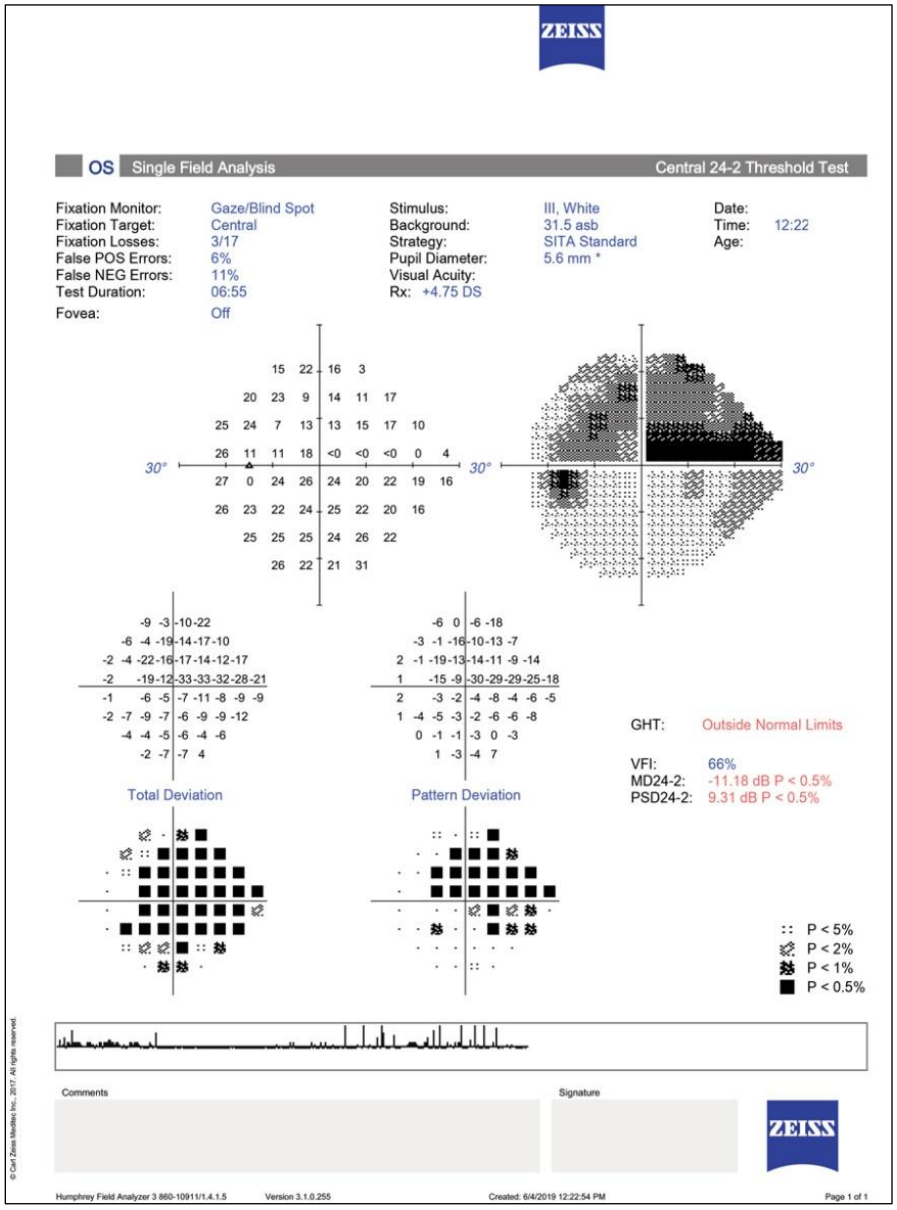


Figure 1.6. Clinical methods for visual field assessment. (A) Confrontational visual field testing is a traditional bedside method for gross evaluation of visual fields. Photo credits: Broadway & Kyari, (2019). (B) Goldmann perimeter used for manual kinetic perimetry, in which moving stimuli of varying size and intensity are presented to map visual field isopters. (C) Humphrey Field Analyzer (HFA) used for standardized static automated threshold perimetry, the current clinical gold standard for visual field assessment. (D) Sample HFA 24-2 threshold test printout demonstrating left eye visual field sensitivity, including test reliability indices (fixation losses, false-positive and false-negative responses), grayscale and probability maps, and global indices, including mean deviation (MD) and pattern standard deviation (PSD), used for clinical evaluation of visual field defects.

1.2.2. Factors Affecting Visual Field Testing Performance

The clinical utility of perimetry depends on the reliability and reproducibility of its measurements, which are highly sensitive to multiple confounding factors. These can be broadly grouped into stimulus characteristics, environmental and optical test conditions, and patient-dependent factors. Stimulus properties such as size, duration, luminance, chromaticity, and contrast directly influence retinal threshold detection by altering the signal-to-noise ratio of the visual response, with inadequate calibration potentially producing artifactual defects or obscuring true scotomas. Environmental and optical conditions, including vertex distance and background illumination, affect retinal adaptation state and can introduce systematic variability across testing sessions if not properly controlled. Furthermore, as a psychophysical task requiring sustained attention, test performance is heavily influenced by patient-related factors including cognitive impairment, fatigue, and learning effects, as well as ergonomic factors, such as prolonged chin-rest discomfort, neck strain, and physical fatigue, which can accelerate attentional decline and degrade response reliability (Montolio et al., 2012a; K.-L. Tan et al., 2017). Recognition and management of these factors are essential for obtaining accurate, reproducible, and clinically meaningful visual field assessments.

1.3. The Pupillary Light Reflex Pathway

Complementing the visual field assessment, pupillary light reflexes (PLR) provide an additional functional measure of afferent and efferent pathway integrity. The pupil, a dynamic aperture controlled by the iris, regulates the amount of light entering the eye. Beyond this physiological role, pupillary responses offer a clinically accessible window into the functional integrity of the afferent visual pathway and associated central autonomic circuits.

The afferent component of the PLR originates in the retina, where light stimulation is transduced by photoreceptors and transmitted via retinal ganglion cells through the optic nerve, optic chiasm, and optic tract to the pretectal nuclei of the midbrain (Figure 1.7). From the pretectal region, bilateral projections are sent to the Edinger–Westphal nuclei, ensuring a consensual response in both pupils (Kandel et al., 2021; Purves et al., 2001). Preganglionic parasympathetic fibers then travel within the oculomotor nerve (cranial nerve III) and synapse in the ciliary ganglion. Postganglionic fibers subsequently reach the iris sphincter pupillae muscle via the short ciliary nerves, mediating pupillary constriction.

The process of pupillary dilation is governed by a three-neuron sympathetic pathway (Dalley et al., 2023; Splittgerber & Snell, 2019). First-order neurons arise in the hypothalamus and descend ipsilaterally through the brainstem to synapse in the intermediolateral cell column of the spinal cord at the C8–T2 levels, known as the ciliospinal center of Budge. Second-order preganglionic neurons exit the spinal cord, ascend through the cervical sympathetic chain, and synapse in the superior cervical ganglion. Third-order postganglionic fibers then follow the internal carotid artery plexus, traverse the cavernous sinus, and enter the orbit via the superior orbital fissure. These fibers ultimately innervate the iris dilator pupillae muscle via the long ciliary nerves, facilitating pupillary dilation.

Together, the parasympathetic and sympathetic pathways provide finely tuned autonomic control of pupil size. Given their anatomical proximity to key structures of the visual and autonomic nervous systems, pupillary responses serve as a clinically valuable indicator of dysfunction along the afferent visual pathway, brainstem, and sympathetic chain.

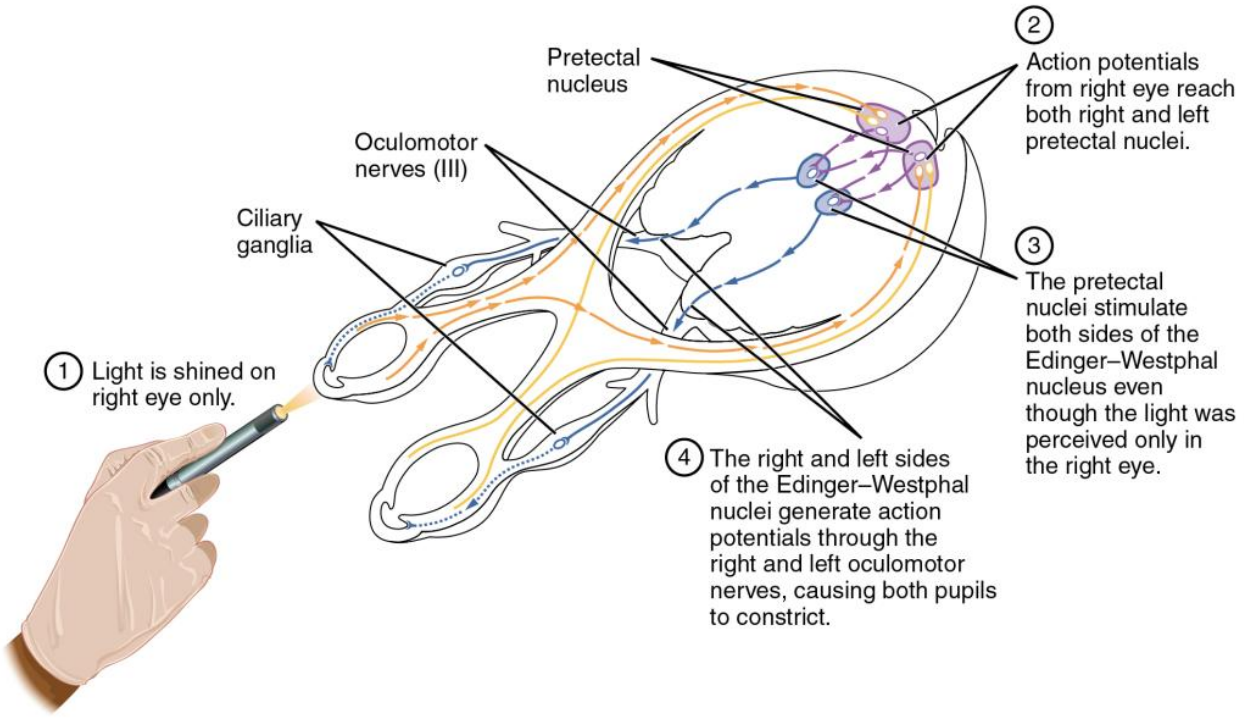


Figure 1.7. Cross-sectional schematic of the pupillary light reflex pathway. Afferent (orange arrows) and efferent (blue arrows) components of the pathway. Light stimulation of the retina initiates afferent signals that travel via the optic nerve, partially decussate at the optic chiasm, and propagate through both optic tracts to the pretectal nuclei in the midbrain. Interneurons project bilaterally from the pretectal nuclei to synapse with the Edinger–Westphal nuclei. Parasympathetic efferent fibers then travel via the oculomotor nerves (CN III) to the ciliary ganglia, stimulating the iris sphincter muscles to induce bilateral pupillary constriction. This dual decussation—at both the optic chiasm and the pretectal-Edinger–Westphal connections—mediates the symmetry between direct and consensual pupillary light responses. Modified from Hutchins et al., 2024 with BioRender.com.

1.3.1. Relative Afferent Pupillary Defects

A key clinical application of the pupillary light reflex integrity is the detection of a relative afferent pupillary defect (RAPD). A RAPD serves as a sensitive diagnostic indicator of asymmetric dysfunction within the afferent visual pathway (Bell et al., 1993). It is clinically assessed using the swinging flashlight test, in which alternating illumination of each eye reveals a paradoxical dilation of the affected pupil when light is directed toward it. This phenomenon occurs due to a reduced afferent input relative to the contralateral eye and consequent reduction in efferent parasympathetic drive required to maintain pupillary constriction.

RAPDs are clinically graded on a four-point severity scale (Grades 1+ to 4+) based on the degree of asymmetry in the pupillary light response observed during clinical examination (Table 1.1) (Yotharak, 2012). Grade 1+ is defined as an initial weak pupillary constriction followed by redilation. Grade 2+ is characterized by an initial stall in pupillary response followed by dilation. Grade 3+ is defined as immediate pupillary dilation upon illumination of the affected eye. Grade 4+ represents a severe RAPD, in which the affected eye exhibits amaurotic pupillary responses with no effective afferent input and sustained dilation when stimulated.

A RAPD is not observed in symmetric bilateral disease, as equal impairment of both afferent pathways does not produce an interocular difference in the pupillary light response. It is therefore most commonly associated with unilateral or asymmetric lesions of the afferent visual pathway, including optic neuritis, ischemic optic neuropathy, compressive optic nerve lesions, advanced glaucoma, and severe unilateral retinal disease. Thus, RAPD provides a sensitive functional marker of anterior visual pathway integrity and complements structural and perimetric assessment in lesion localization.

1.3.2. Standard Clinical Pupillary Assessment

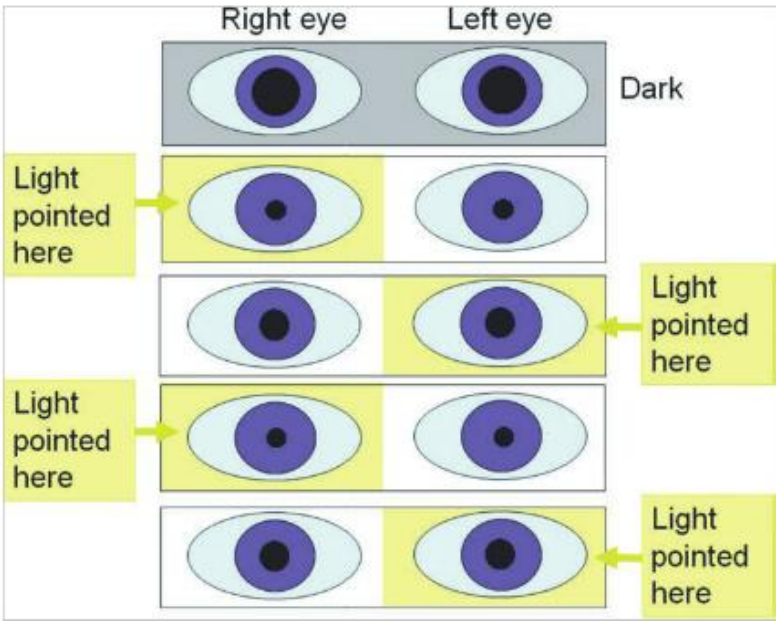
Standard clinical assessment of the pupils is guided by the acronym PERRLA, which denotes that "pupils are equal, round, and reactive to light and accommodation" (Buratto et al., 1990). This examination evaluates four key aspects of pupillary function using a penlight and ruler: (1) pupil size; (2) pupil shape and symmetry; (3) reactivity to light under varying illumination conditions; and (4) the accommodative response. Following the PERRLA examination, the presence of a RAPD is assessed using the swinging flashlight test, as illustrated in Figure 1.8A–B. Together, these components provide a rapid clinical overview of the integrity of both the afferent and efferent pupillary pathways.

As a supplementary diagnostic tool, neutral density (ND) filters may be used during pupillary examination to quantify the severity of a RAPD (Yotharak, 2012). By attenuating the intensity of light entering the unaffected eye in calibrated increments, ND filters equalize the afferent input between the two eyes, allowing the magnitude of the RAPD to be objectively graded. This approach improves the detection and monitoring of subtle asymmetries in afferent visual pathway function that may not be readily apparent during routine clinical examination.

Table 1.1. The clinical grading of a RAPD and pupil observations during the swinging light test.

RAPD Grade	Pupil Observations
Subtle/trace	Both pupils may transiently enlarge during the interval between illumination of the normal and affected eye. Upon illumination of the affected eye, a brief constriction may occur before the pupils begin to dilate.
1+	The pupils exhibit an initial weak constriction when the light is directed to the affected eye and subsequently dilates after some delay, indicating a mild RAPD.
2+	The pupils exhibit an initial stall upon illumination of the affected eye, followed by redilation, indicating a moderate RAPD.
3+	The pupils dilate immediately (no initial constriction or stall) upon illumination of the affected eye, indicating a marked RAPD.
4+ or Amaurotic	Observed in an eye with no light perception. The pupils constrict only during illumination of the fellow eye. When the light is redirected to the amaurotic eye, the pupil rapidly dilates despite continued illumination, reflecting the absence of an afferent response.

A



B

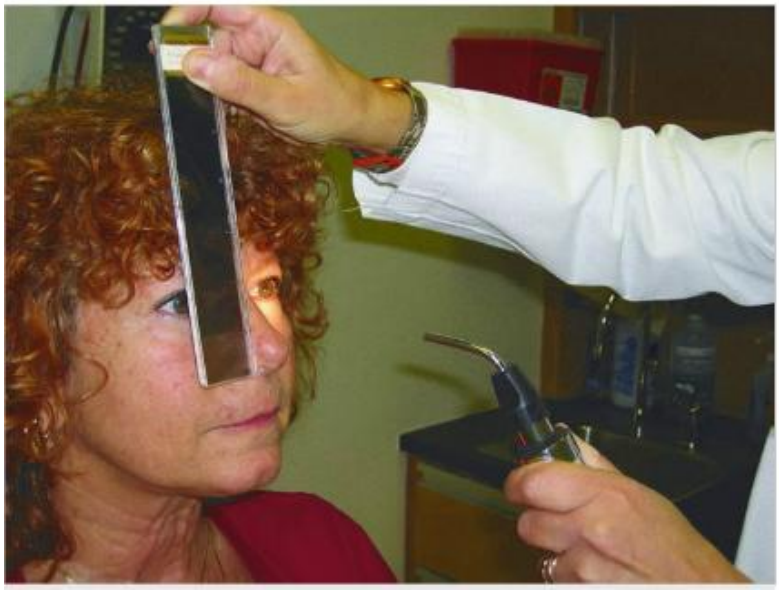


Figure 1.8. Clinical assessment of a RAPD. (A) An illustration of the swinging flashlight test demonstrating a left RAPD. Bilateral pupillary constriction occurs when the right eye is illuminated, whereas illumination of the left eye produces relative pupillary dilation due to reduced afferent input. (B) The application of a neutral density filter to grade the severity of a RAPD. Photo credits: Biousse & Newman, (2016).

1.4. Current Limitations of Conventional Methods

Despite its status as the clinical gold standard, conventional automated perimetry has several inherent limitations that constrain its accessibility and broader clinical applicability. Current standard automated perimetry systems typically lack portability, necessitate controlled darkened testing environments and trained personnel, and can be challenging for patients with limited mobility or those who cannot be adequately positioned or comfortably accommodated at the machine (Lu et al., 2024). Furthermore, testing is relatively time-intensive and subject to patient-related variability, including fatigue, fluctuating attention, and learning effects, which may reduce test–retest reliability and complicate the longitudinal interpretation of disease progression. Similarly, conventional pupillometry, including manual assessment with a standard pupil gauge, remains widely used in routine clinical practice due to its inherent simplicity and speed. However, this approach is fundamentally subjective and limited in measurement precision and reproducibility. Manual estimation is frequently confounded by inter-observer variability, fluctuations in ambient lighting conditions, and physiological factors such as age, pharmacologic agents, and baseline pupillary diameter. These limitations hinder the detection of subtle pupillary dysfunction. As a result, there has been growing interest in VR-based approaches to visual function assessment, including integrated perimetry and pupillometry.

1.5. Virtual Reality

Virtual reality (VR) is a technology that enables users to explore and interact with computer-generated, three-dimensional (3D) multimedia sensory environments in real time (Farrell & MacDougall, 2023). By facilitating an immersive, first-person experience, VR allows for precise temporal and spatial modulation of the visual stimuli. In a clinical context, this high level of immersion minimizes external distraction, standardizes testing environments, and enables

reproducible delivery of visual stimuli. VR can be delivered through a variety of platforms, including computer or mobile displays, projection-based immersive systems, and head-mounted displays (HMDs). HMDs, in particular, can provide a fully immersive, individualized experience by presenting stereoscopic visual stimuli directly to the user while isolating them from the external environment.

1.5.1. Virtual Reality and Its Implementation in Ophthalmology

Head-mounted VR devices are driving a paradigm shift in ophthalmic diagnostics by providing portable, patient-centered alternatives to conventional clinic-based assessments. A narrative review by Ahuja et al. (2025) highlighted the expanding role of VR in ophthalmology, encompassing medical education, surgical simulation, pediatric visual rehabilitation, and clinical diagnostics. Within diagnostic applications, VR has shown promise for visual field assessment applications, offering standardized testing conditions in an immersive and portable platform.

Despite these advances, the clinical adoption of VR-based perimetry and pupillometry remain largely investigational. Validation studies have demonstrated encouraging agreement with conventional automated perimetry; however, concerns remain regarding diagnostic accuracy, reliability indices (including fixation losses and false positive and false negative responses), and performance across diverse patient populations (Montelongo et al., 2021; Park et al., 2026). In addition, implementation costs, ongoing software subscriptions, and integration into existing clinical workflows continue to influence the feasibility of widespread adoption. Collectively, further technical validation and economic evaluation studies are required prior to routine incorporation into ophthalmic practice.

1.5.2 Review of Current Virtual Reality–Based Visual Field Devices

Systematic reviews have reported encouraging diagnostic performance for VR-based visual field testing, suggesting that these technologies may serve as viable alternatives or adjuncts to conventional standard automated perimetry (Selvan et al., 2024). In this comprehensive review of 64 studies evaluating 36 VR headset-based perimetry devices, the authors highlighted considerable heterogeneity in hardware platforms, testing protocols, thresholding algorithms, and outcome measures, limiting direct comparisons between studies and emphasizing the need for greater standardization before widespread clinical implementation. Similarly, the authors also concluded that while VR perimetry demonstrates promising diagnostic accuracy and offers advantages such as portability and improved accessibility, the evidence remains limited by small study populations, methodological variability, and a lack of robust validation across diverse ocular pathologies. Although several commercially available platforms, including Heru and Olleyes (VisuALL), have adopted established SAP protocols such as the 24-2 and 30-2 testing strategies, much of the available evidence is derived from conference abstracts and preliminary reports rather than peer-reviewed clinical studies (Razeghinejad et al., 2020; Tharp et al., 2023). Consequently, further high-quality validation studies comparing VR-based platforms with established gold-standard perimetry are required before routine clinical adoption can be recommended.

1.5.3 Review of Current Automated and VR-Based Pupillometers

Previously, automated infrared pupillometers have been developed to provide objective, quantitative, and reproducible measurements of PLR dynamics. However, most commercially available automated pupillometers employ monocular designs that assess each eye sequentially, limiting the simultaneous evaluation of direct and consensual pupillary responses and introducing potential temporal variability between measurements (Sander et al., 2025).

More recently, VR head-mounted displays equipped with integrated binocular eye-tracking technology have emerged as a novel approach to automated pupillometry (Negi et al., 2024a; Park et al., 2026b). By providing simultaneous bilateral pupil tracking within a controlled, light-occluded environment, these systems have the potential to improve the standardization and objectivity of pupillary assessment. Despite these theoretical advantages, evidence supporting the clinical performance of VR-based pupillometry remains limited and is largely derived from pilot studies and proof-of-concept investigations. Consequently, further validation against established clinical assessment methods is required to determine the diagnostic accuracy, reliability, and clinical utility of these emerging technologies across a range of ocular and neuro-ophthalmic conditions.

1.5.4. Order of Magnitude

Order of Magnitude (OM) Labs (Hyderabad, India), has developed VR-based platforms designed for concurrent automated perimetry and pupillometric assessment. The visual field-testing module replicates SAP paradigms within a portable HMD integrated with Bluetooth-enabled handheld response controllers, specifically the Pico Neo 2 Eye® VR headset (Pico Immersive Pte. Ltd., Singapore). This perimetry system utilizes a two-level full-thresholding algorithm to quantify retinal sensitivity, with stimulus presentation and background luminance calibrated directly via the HMD's internal display.

Complementing this visual field module, OM has also developed Pupil X, a video-based software application engineered to objectively quantify relative afferent pupillary defects (RAPDs). The Pupil X application also operates natively on the Pico Neo 2 VR headset, which incorporates integrated infrared eye-tracking technology developed by Tobii AB® (Stockholm,

Sweden). This configuration enables automated digital capture of binocular pupillary dynamics within a highly standardized, lighting-controlled testing environment.

Both the OM visual field platform and the Pupil X application are currently undergoing clinical validation. Although these technologies offer the potential to integrate visual field testing and pupillary assessment within a single portable VR platform, evidence supporting their diagnostic performance and clinical utility remains limited. Consequently, further validation against established gold-standard methods is required before their routine clinical implementation can be considered.

1.5.5. Retinalogik Inc. ©

Retinalogik Inc. (Calgary, Canada) has developed a portable, software-as-a-service platform for functional ophthalmic assessment using commercially available VR head-mounted displays integrated with cloud-based data management and analytics. The platform supports multiple visual assessment protocols, including visual acuity, contrast sensitivity, colour vision, and visual field testing; however, this study focused exclusively on its visual field examination module.

The Retinalogik visual field (RVF) platform adapts standard automated perimetry principles to a portable VR-based format, enabling visual field assessment outside conventional clinic-based testing environments. Although the platform is commercially available, published evidence evaluating its performance relative to the standard HFA remains limited. Consequently, further validation is required to establish its diagnostic agreement, reliability, and clinical utility. Such evidence is necessary to determine the potential role of VR-based perimetry in routine ophthalmic practice, including applications in teleophthalmology and community-based eye care.

1.6. Rationale, Aims, and Hypothesis

Standard automated perimetry and pupillometry are fundamental clinical modalities for the diagnosis and monitoring of ophthalmic disorders. However, conventional methods have substantial limitations. Standard automated perimetry, including the HFA, requires specialized clinical infrastructure and incurs substantial equipment and maintenance costs, whereas manual pupillary assessment is inherently subjective and susceptible to inter-examiner variability.

In contrast, emerging VR-based head-mounted platforms offer a portable, cost-effective, and scalable alternative that may enable decentralized ophthalmic assessment and remote screening. Despite this, there remains a paucity of high-quality evidence validating their diagnostic accuracy and clinical equivalence to established gold-standard methods. Rigorous evaluation within controlled tertiary care settings is therefore warranted to define their clinical utility and establish performance benchmarks across heterogeneous patient populations.

The primary objective of this study was to assess the diagnostic validity and clinical utility of selected VR-based ophthalmic assessment platforms: (a) Order of Magnitude (OM) visual field, (b) Pupil X pupillometry, and (c) Retinalogik visual field (RVF), relative to conventional clinical standards. Specifically, the study quantified the level of agreement between VR-derived functional measures and those obtained using standard automated perimetry (HFA) and manual pupillary examination in both healthy individuals and patients with confirmed ocular pathology.

It was hypothesized that the VR platforms will demonstrate comparable performance to the gold-standard HFA and manual pupillary assessments, with similar sensitivity and specificity for the detection of visual field and pupillary abnormalities in a tertiary care hospital setting. To address the study hypothesis, the following specific aims were established:

Aim 1: To evaluate the clinical validity of the OM visual field, Pupil X, and RVF platforms against gold-standard visual field and pupillary assessments in healthy controls and patients with ocular disease.

Aim 2: To assess the feasibility and clinical utility of VR testing platforms in a high-volume tertiary eye care setting.

1.7. Significance of the research

The comparative validation of VR-based perimetry and pupillometry against conventional diagnostic standards has important clinical and operational implications. Traditional visual field assessments, such as the HFA, are limited by restricted accessibility for patients with limited mobility, lack of portability, and high acquisition and maintenance costs, reducing their cost-effectiveness in non-specialized or decentralized care settings. Similarly, current standalone pupillometry methods are often constrained by inter-rater variability, lack of standardized testing environments, and susceptibility to ambient light artefacts, which can adversely affect measurement reliability and reproducibility. Conversely, portable VR-based platforms can offer a scalable and more accessible alternative that may reduce environmental variability and facilitate the integration of multiple functional assessments within a single automated testing session.

This thesis builds upon existing research on VR-based ophthalmic devices by providing a direct clinical evaluation of their validity, diagnostic performance, and feasibility within a high-volume tertiary eye care setting. While previous studies have largely focused on individual VR platforms or proof-of-concept applications, these findings can provide a more comprehensive assessment of the clinical applicability and real-world implementation of VR-based ophthalmic testing, while also identifying areas requiring further optimization and validation. Validation of these portable systems may also support the transition toward more decentralized models of

ophthalmic care, with potential applications in teleophthalmology, rural healthcare settings, and underserved populations.

CHAPTER 2. MATERIALS AND METHODS

2.1. Humphrey Field Analyzer Specifications

This study used the Humphrey Field Analyzer 3 (HFA; Carl Zeiss Meditec, Dublin, CA, USA) as the clinical reference standard for static automated perimetry. The HFA utilizes a projection-bowl perimeter to deliver a calibrated background luminance of 31.5 apostilbs (10 cd/m²) adhering to standard photopic testing conditions. Visual stimuli are presented using standard Goldmann Size III target configurations (0.43° angular subtense) flashed for white-on-white (achromatic) perimetry.

Testing parameters were configured to standard clinical grids, specifically the 30–2, 24–2, and 10–2 test patterns. To optimize test efficiency without compromising diagnostic reproducibility, the HFA system employs the Swedish Interactive Threshold Algorithm (SITA Standard) for threshold estimation. Retinal sensitivity thresholds were determined point-by-point across the visual field using a staircase thresholding strategy. Primary output parameters were extracted via the Zeiss FORUM® software platform for comparative analysis including greyscale maps, sensitivity maps, mean deviation (MD), and pattern standard deviation (PSD).

2.2. Order of Magnitude Platform Specifications

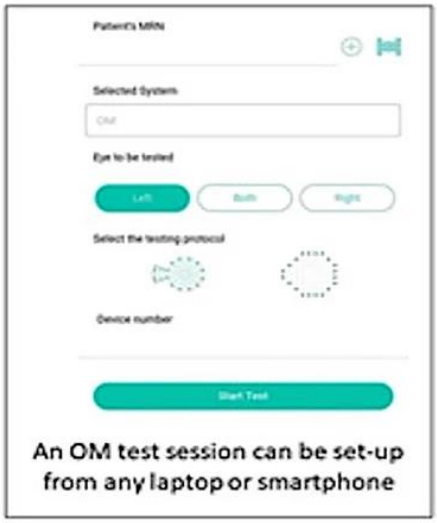
The OM visual field-testing protocol was integrated into a commercially available Pico Neo VR headset equipped with integrated eye tracking (Tobii AB®, Stockholm, Sweden) and Bluetooth handheld controllers (Figure 2.1). The system was calibrated to a background luminance of 1.3 cd/m² with a stimulus intensity range of 0–50 dB and employed a standard 24–2 test pattern comprising 54 locations identical to the HFA. Testing used a Goldmann size III stimulus (0.43°) within a two-step full-threshold strategy, with initial high- and low-intensity stimuli set at 20 dB and 7.5 dB, respectively. The system additionally generates standard global indices, including MD,

and PSD, with pointwise responses classified as normal, relative defect, or absolute defect based on sensitivity profiles.

The Pupil X pupillometry testing protocol was integrated and modularized into the same Pico Neo 2 VR headset and employs a video-based assessment for the detection of relative afferent pupillary defect (RAPD). The stimulus illumination was calibrated to 125 lux, which has been reported as an optimal lighting condition for RAPD detection in patients with unilateral optic neuropathy (Negi et al., 2024b).

A

B



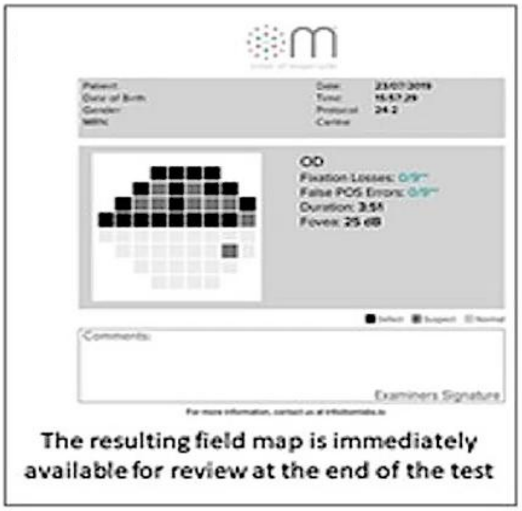


Figure 2.1. OM virtual reality perimetry and Pupil X platform. (A) The figure illustrates the Pico Neo 2 VR headset with OM platform integration and (B) the associated portal interface for test setup and results display.

2.3. Retinalogik Platform Specifications

The Retinalogik visual field (RVF) platform is comprised of two components: (1) a proprietary Retinalogik VR application preloaded onto a Pico Neo VR head-mounted display, which administers visual field examinations with Bluetooth handheld controllers; and (2) a secure, browser-based portal enabling data management, test review, and clinical analytics (Figure 2.2). Together, these components facilitate the delivery, storage, and interpretation of functional visual assessments within a portable and scalable VR-based framework.

To ensure diagnostic comparability with the gold-standard HFA, the RVF platform uses a SITA-standard testing protocol to determine sensitivity thresholds with background luminance calibrated to 1.3 cd/m², equivalent to the HFA background illumination. The system also employs cloud-based processing to derive standard visual field indices, including MD, and PSD while also generating a proprietary metric, the Retinalogik Field Index (RLFI).

A

B

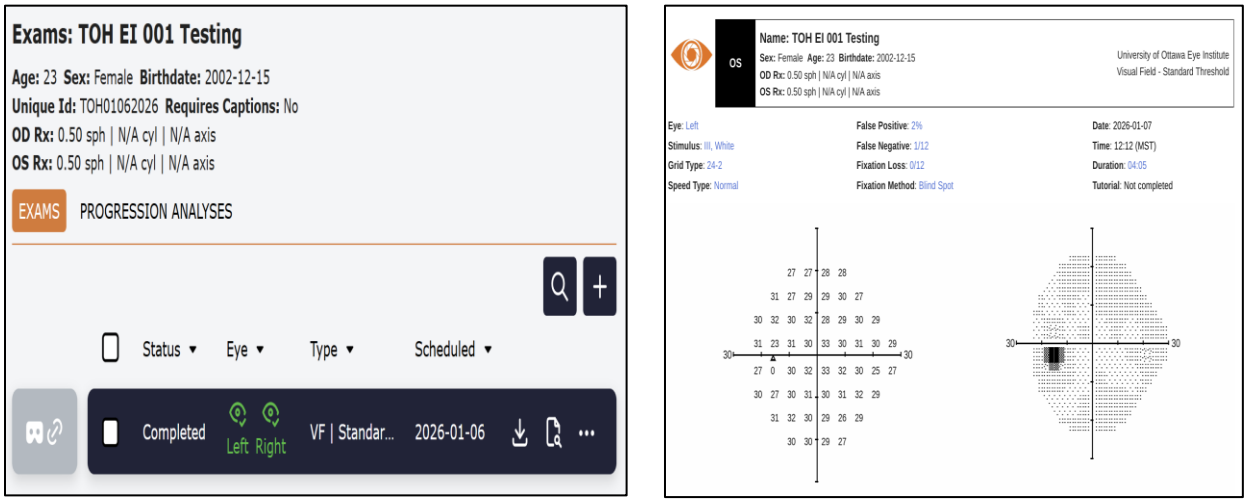


Figure 2.2. RVF virtual reality perimetry platform. (A) The figure illustrates the Pico Neo 2 VR headset with RVF integration and (B) the associated portal interface for test setup and results display.

2.4. Study Design

This prospective, cross-sectional validation study evaluated the validity and clinical utility of three emerging virtual reality-based diagnostic platforms against established clinical reference standards. The visual field platforms, OM and RVF, and the Pupil X pupillometry platform were evaluated according to the study workflow outlined in Figure 2.3. Participants were recruited via convenience sampling from the University of Ottawa Eye Institute and comprised both healthy controls and patients with confirmed ocular pathology. Healthy controls underwent a comprehensive ophthalmic examination by an experienced ophthalmologist to confirm the absence of ocular disease and establish baseline normative metrics. The study adhered to the Declaration of Helsinki and was approved by the Ottawa Health Science Network Research Ethics Board (OHSN-REB). Written informed consent was obtained from all participants prior to enrollment.

Following eligibility screening and written informed consent, participants underwent a standardized testing protocol:

- **Visual Field Testing:** The reference standard assessments were performed and consisted of standard automated perimetry using the Humphrey Field Analyzer. Participants subsequently underwent assessment using the VR devices, either the OM or the RVF visual field platform.
- **Pupil Testing:** A manual clinical pupillary examination was performed prior to using the Pupil X platform either during the same study visit as the visual field testing or at a separate visit, depending on participant availability and study logistics.

Where operationally feasible, the order of the VR-based assessments was randomized to minimize potential fatigue, learning, and order effects. All assessments were conducted according to standardized testing protocols to ensure consistency across participants. Conventional

examinations were conducted by certified ophthalmic technologists (visual field) or by experienced clinicians (pupil exam) using predefined operating procedures, while all VR-based testing was conducted by CD. Device setup and calibration, participant positioning, test instructions, and examination conditions were standardized throughout the study to minimize operator-dependent variability and maximize the reproducibility of study measurements.

The diagnostic performance of each VR platform was evaluated by assessing the agreement with the corresponding reference standard using established measures of diagnostic accuracy and agreement. These analyses were performed to determine the diagnostic validity, clinical performance, and potential utility of each platform in a tertiary ophthalmology setting.

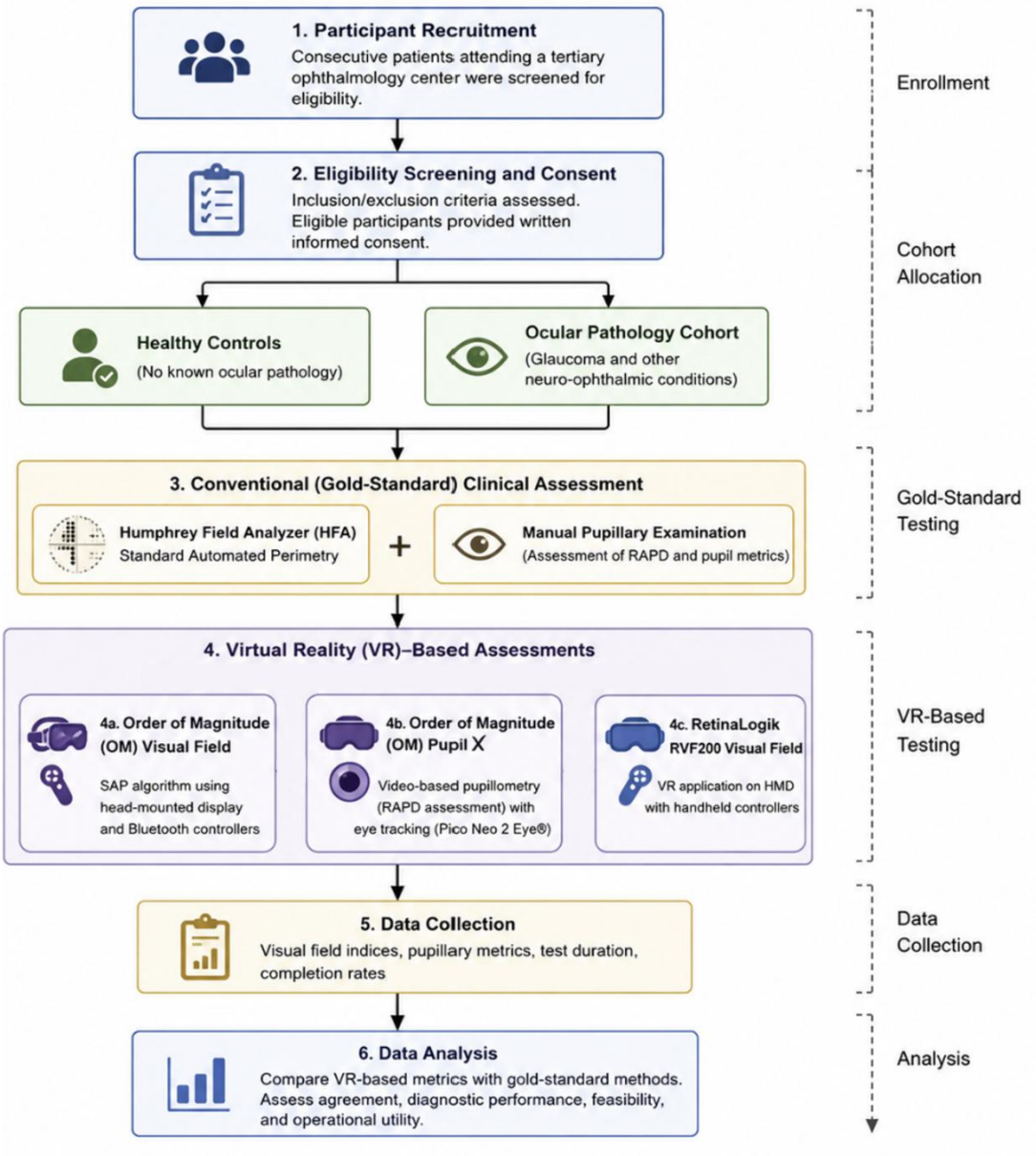


Figure 2.3. Study design and participant recruitment workflow. Participants were screened and consented, followed by baseline conventional assessment (HFA perimetry and manual pupillometry) and VR-based testing (OM visual field, Pupil X, and Retinalogik visual field). VR outcomes were compared with reference standard measures for validation and agreement analysis. Created with BioRender.com.

2.4.1. Inclusion and Exclusion Criteria

To ensure a representative sample across a broad spectrum of visual function, the study population comprised both healthy controls and patients with confirmed ocular pathologies. Eligible participants were stratified based on specific systemic and ophthalmic criteria. Table 2.1

establishes a comprehensive inclusion and exclusion criteria applied across each participant group and testing modality cohort.

Table 2.1. Inclusion and exclusion criteria for all research participant groups.

Criteria	Healthy Control Group	Patient Group
Inclusion	▪ Age ≥18 years ▪ Ability to provide written informed consent ▪ Best-corrected visual acuity sufficient to complete visual field testing ▪ Normal ophthalmic examination with no history of ocular or neuro-ophthalmic disease ▪ Able to undergo VR-based testing and conventional clinical assessments	▪ Age ≥ 18 years ▪ Ability to provide written informed consent ▪ Able to undergo VR testing and conventional clinical assessments
Exclusion	▪ History of ocular or neuro-ophthalmic disease ▪ Previous ocular surgery or trauma affecting visual function (unless clinically insignificant) ▪ Cognitive impairment preventing reliable testing ▪ Inability to maintain fixation or comply with testing instructions ▪ Medical or physical conditions preventing safe use of the VR headset	▪ Cognitive impairment preventing reliable testing ▪ Medical or physical conditions preventing safe use of the VR headset

2.4.2. Visual Field Testing Protocol

Visual field (VF) testing was performed using both standard automated perimetry and VR-based platforms as part of a structured comparative protocol. Standard of care VF assessment was conducted using the HFA with a 30-2, 24-2, or 10-2 testing strategy, as ordered by the treating physician based on clinical indication. Testing was performed under controlled ambient lighting conditions in accordance with routine clinical procedures and completed by trained ophthalmic technicians.

Following standard automated perimetry, VR-based visual field testing was performed using either the OM visual field platform or the Retinalogik VR system. Prior to testing, the head-mounted display was adjusted to achieve an appropriate interpupillary distance, secure fit, and optimal visual alignment. Participants were comfortably positioned and instructed to maintain steady central fixation on the fixation target while minimizing head and body movement throughout the examination. Visual stimuli were presented monocularly according to the device-specific thresholding algorithm, and participants indicated stimulus detection using the handheld controllers. Testing was performed under standardized conditions consistent with each manufacturer's recommended protocol to optimize fixation stability and response reliability.

Testing was conducted in a randomized order between modalities to minimize potential order effects. To reduce fatigue-related bias, a minimum rest period of one hour was provided between consecutive visual field examinations. All visual field assessments (HFA and VR testing) were completed within one month to minimize the likelihood of clinically significant changes in visual function occurring between tests and to ensure comparability of visual field measurements across modalities.

2.4.3. Pupillometry Protocol

Pupillary assessment was performed using both standard clinical examination and VR pupillometry as part of the structured comparative protocol. Standard of care assessment was conducted manually using a pupil gauge under standard light conditions for the evaluation of pupil size, shape, symmetry, and near responses. For pupillary reactivity with a penlight, such as pupil constriction, dilation, and RAPD assessment, standard dark conditions were applied in accordance with routine clinical practice.

After participants were appropriately placed in a room were appropriately positioned to ensure stable head alignment and optimal device fit, the paradigm with the VR pupillometer (Pupil X platform) started with an initial 5 seconds-long period of complete darkness for both pupils to reach a steady baseline state, followed by a 1-second-long light pulse to both eyes simultaneously and 3-seconds of darkness. This was followed by three cycles of independent light stimulation of the two eyes, with each cycle consisting of a 1-second-long pulse of 125 lux to one eye followed by 3-seconds of darkness. This temporal sequence of light simulation followed the recommendations of the standards for assessment of RAPD by Kelbsch et al (2019). Measurements were acquired under standard mesopic testing conditions to minimize the chances of light entering from the sides of the VR headset.

2.5. Outcome Measures

The primary outcome measures for visual field assessment included global and localized indices derived from both standard automated perimetry and VR-based platforms. Specific measures to compare the machines included test duration, mean deviation (MD), and pattern standard deviation (PSD). Additional performance outcomes included test reliability indices such as fixation losses, false-positive rates, and false-negative rates. Pupillary assessment outcomes included baseline pupil diameter, the presence or absence of a RAPD, and RAPD severity graded on a standardized scale from 1+ to 4+. Collectively, these outcome measures allowed for comprehensive assessment of concordance, diagnostic performance, and test reliability between VR-based testing and standard of care assessments.

2.6. Sample Size Considerations

A total target recruitment of approximately 500 participants was established to provide adequate statistical power for validation across the VR-based VF and pupillometry platforms, with

approximately equal numbers of participants targeted for each platform and representation from both healthy controls and patients with ocular pathology. Sample size calculations were performed using R Statistical Software (version 4.6.1; R Foundation for Statistical Computing, Vienna, Austria) based on conventional assumptions of a two-sided significance level of $\alpha = 0.05$ and statistical power of 80–90%. This sample size was expected to provide sufficient power to detect moderate associations (correlation coefficient $r \approx 0.30–0.40$) while allowing robust evaluation of agreement between devices using correlation and Bland–Altman analyses. Furthermore, the proposed sample size was anticipated to provide adequate representation of a range of visual field defects and pupillary abnormalities, supporting subgroup analyses and improving the precision of estimates for diagnostic performance and inter-device agreement.

2.7. Statistical Analysis Methods

Statistical analyses were performed using R Statistical Software (version 4.6.1; R Foundation for Statistical Computing, Vienna, Austria) and Microsoft Excel (Microsoft Corporation, Redmond, WA, USA). Microsoft Excel was also used for data management and preliminary organization prior to statistical analysis. Continuous variables were summarized using means and standard deviations, while categorical variables were presented as frequencies and percentages. Comparisons between standard-of-care and VR-based measurements were performed using paired statistical tests. In cases where the data exhibited mild skewness, approximate normality was assumed based on the Pearson skewness coefficient. Given the sufficiently large sample size, the Central Limit Theorem was considered applicable, supporting the use of the paired-samples t-test. Statistical significance was defined as a p-value of < 0.05 and asterisks represent significant values as follows: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

Correlation between modalities was evaluated using Pearson correlation coefficients. Agreement between VR-based and standard of care measurements for visual field indices, including MD and PSD, was assessed using Bland–Altman analysis, with calculation of mean bias and 95% limits of agreement.

For pupillary outcomes, inter-eye symmetry and inter-modality differences were analyzed using similar paired comparative approaches. Diagnostic performance and classification agreement for the RAPD measurements were assessed using sensitivity, specificity, and agreement statistics such as Cohen's kappa.

2.7.1. Delta Heatmaps for Sensitivity Map Findings

To evaluate the point-by-point spatial alignment and localized sensitivity variance of the VR perimetry platforms, OM and RVF, against the gold-standard HFA, differential delta heat maps were constructed across the sensitivity grids. All selected cases of perimetric sensitivity maps were co-registered and normalized onto a uniform 30° Cartesian grid, with the central fixation established as the coordinate origin (0,0). Localized sensitivity variance was quantified as a decibel delta (Δ dB), calculated by subtracting the HFA point-specific threshold value from the corresponding VR platform sensitivity point output. All statistical computations were performed using R programming.

2.7.2. Spatial Profiling for Greyscale Map Analysis

To assess spatial resolution and defect depth fidelity between the VR perimeters and the HFA, continuous cross-sectional profiles of localized visual field defects from the grayscale maps were extracted. The selected cases of sensitivity maps were co-registered and normalized to a ±30° Cartesian grid centered on fixation (0,0). Sensitivity loss was expressed as normalized defect depth (0% = normal, 100% = absolute scotoma) as a function of horizontal and vertical eccentricity

position. Data were binned into 2° spatial intervals, with descriptive statistics (mean, standard deviation, and standard error) calculated for each bin. Spatial trends and boundary profiles were modelled using locally estimated scatterplot smoothing (LOESS) (α = 0.22–0.25). Inter-device spatial discrepancies were defined as regions where 95% confidence intervals did not overlap. All analyses were performed in R.

CHAPTER 3. RESULTS

3.1. Visual Field Data Subset – Baseline Characteristics of the Study Population

A total of 251 participants (88 males and 163 females), comprising healthy individuals and patients with confirmed ocular pathology, were included in the study. There were 20 participants who were excluded from the final analysis due to incomplete study completion or unrecoverable data. Participant ages ranged from 20 to 86 years. Baseline demographic and clinical characteristics were summarized separately for the two visual field VR device groups: OM cohort and RVF cohort (Table 3.1).

Order of Magnitude Cohort

The OM cohort included 78 participants (147 eyes), comprising 33 healthy controls and 45 patients diagnosed with ocular pathology. Among the patient subgroup, glaucoma was the most prevalent diagnosis (n = 21 patients), including 20 patients diagnosed with open angle, angle closure, pseudoexfoliation, neovascular, pigment dispersion syndrome, and mixed mechanisms. The remaining diagnoses included pituitary adenoma, meningioma, or craniopharyngioma (n = 5), idiopathic intracranial hypertension (n = 3), optic neuritis (n = 3), neurosarcoidosis (n = 3), and a heterogeneous group of genetic, strabismus, disc, high myopia, and retinal pathologies (n = 12).

Baseline demographic characteristics were comparable between the healthy and patient groups, with no significant differences in age (48.6 ± 21.1 vs. 48.9 ± 19.8 years; $p = 0.95$) or sex distribution (71% vs. 62% female; $p = 0.66$), respectively. Best-corrected distance visual acuity (BCVA) differed significantly between groups (0.05 ± 0.04 vs. 0.47 ± 0.32 logMAR; $p < 0.001$), while intraocular pressure (IOP) was similar (12.3 ± 1.9 vs 12.2 ± 2.8 mmHg; $p = 0.82$). Slit-lamp examination and fundus findings were within normal limits in the healthy cohort, whereas findings in the patient group were variable and consistent with the underlying clinical diagnoses.

Retinalogik Cohort

The RVF cohort included 202 participants (404 eyes), comprising 30 healthy controls (60 eyes) and 172 patients (344 eyes). The clinical diagnoses within the patient subgroup were heterogeneous. Neuro-ophthalmic conditions were most prevalent, including idiopathic intracranial hypertension (n = 52), optic neuropathy (n = 21), optic neuritis (n = 10), pituitary, pineal, or meningioma tumors (n = 10), non-arteritic anterior ischemic optic neuropathy (n = 7), and giant cell arteritis (n = 2). Glaucoma accounted for 15 patients, followed by retinal pathologies (n = 11), optic disc anomalies or edema (n = 9), and multiple sclerosis (n = 7). The remaining diagnoses formed a heterogeneous group of neuro-inflammatory, cranial nerve palsy, vascular, and Plaquenil toxicity conditions (n = 28).

Baseline demographic characteristics were comparable between the healthy and patient groups, with no significant differences in age (47.7 ± 15.0 vs 51.9 ± 19.1 years; $p = 0.18$) or sex distribution (68% vs 66% female; $p = 1.00$), respectively. Mean BCVA differed significantly between groups (0.02 ± 0.04 vs 0.35 ± 0.81 logMAR; $p < 0.001$), while IOP was similar (11.3 ± 1.9 vs 10.2 ± 2.8 mmHg; $p = 0.49$). Slit-lamp examination and fundus findings were within normal limits in the healthy cohort, whereas findings in the patient group varied according to the underlying ocular pathology.

Table 3.1. Baseline demographics and clinical characteristics of OM and RVF cohorts stratified by disease status.

Device	Parameter	Healthy	Patient	*p*-value
OM	Participants, N	33	45	—
	Eyes, N	66	81	—
	Sex (% female)	71	62	0.66[1]
	Mean age (years; mean ± SD)	48.6 ± 21.1	48.9 ± 19.8	0.95
	Best-corrected distance visual acuity (logMAR; mean ± SD)	0.05 ± 0.04	0.47 ± 0.32	< 0.001***
	Intraocular pressure (mmHg; mean ± SD)	12.3 ± 1.9	12.2 ± 2.8	0.82
	Clinical findings	Within normal limits	Variable	—
RVF	Participants, N	30	172	—
	Eyes (N)	60	344	—
	Sex (% female)	68	66	1.00[1]
	Mean age (years; mean ± SD)	47.7 ± 15.0	51.7 ± 19.1	0.18
	Best-corrected distance visual acuity (logMAR; mean ± SD)	0.02 ± 0.04	0.35 ± 0.81	<0.001***
	Intraocular pressure (mmHg; mean ± SD)	11.3 ± 1.9	10.2 ± 2.8	0.49
	Clinical findings	Within normal limits	Variable	—

All *p*-values were calculated using a two-tailed standard t-test. [1]Chi-squared test; *Statistically significant values, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

3.2. Comparison of Operating Characteristics Across Machines

The operating characteristics of the VR perimetry devices were compared to the clinical standards to assess their practical implementation in a clinical setting. Both VR-based devices were portable and required less physical space than the standard HFA. Testing was performed in standard examination rooms without dedicated infrastructure or controlled lighting conditions. In contrast, the HFA required fixed installation and a dedicated testing environment.

Automated instructions and integrated eye-tracking functionality were present in both VR systems, reducing operator input during test administration. Occasional headset repositioning and adjustment were required to maintain device alignment during testing. The HFA required monocular testing with an eye patch and appropriate use of trial lenses, calculated according to each participant's age and refractive error. In comparison, VR-based visual field testing did not require trial lenses, as participants were permitted to use their habitual refractive correction, including glasses or contact lenses, when applicable. Further technical and operational characteristics are summarized in Table 3.2. Differences were observed across devices in testing strategy, stimulus presentation method, fixation monitoring approach, and requirements for refractive correction and monocular testing. Variations were also noted in environmental testing conditions and setup requirements.

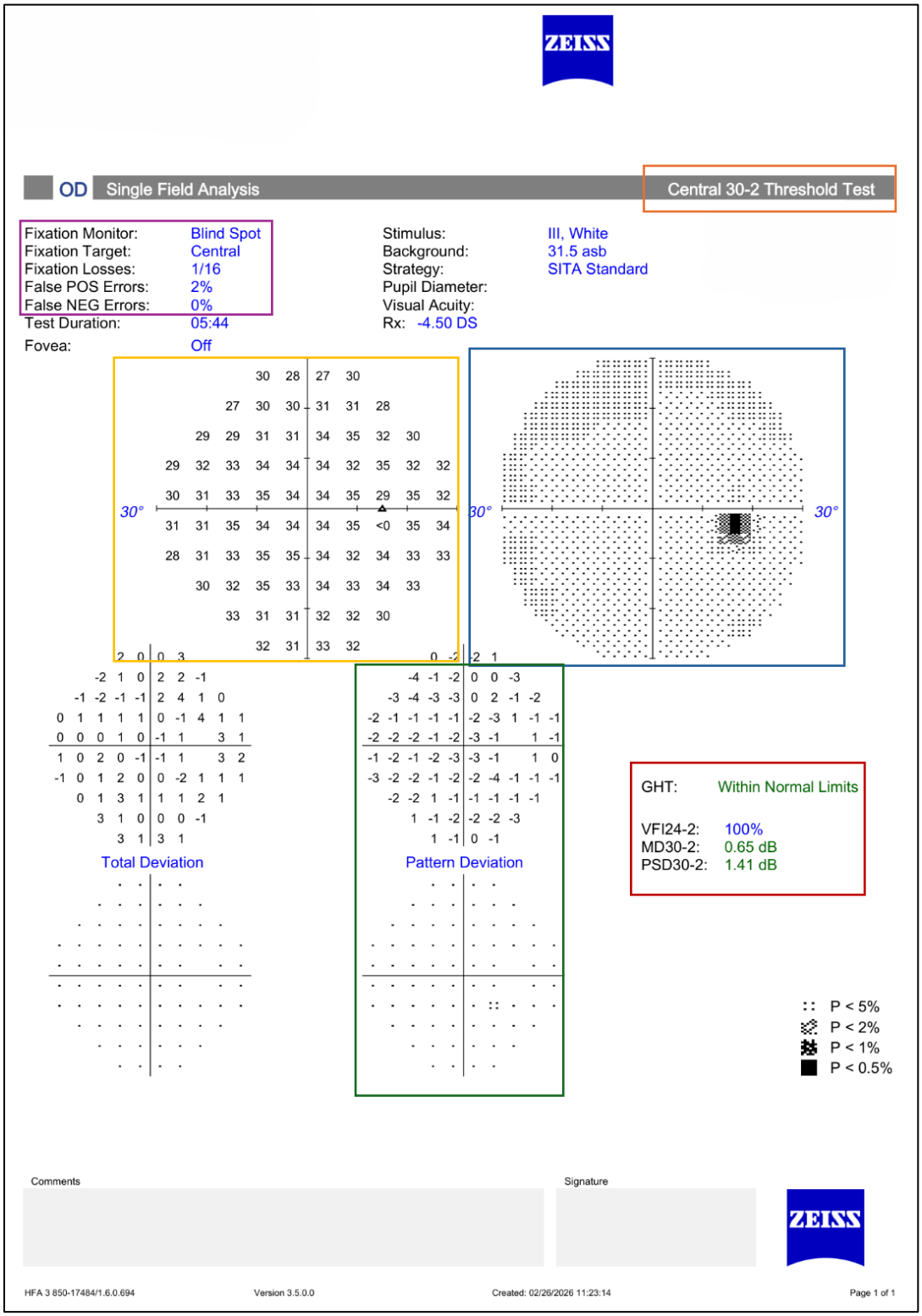
Evaluation of the diagnostic printouts revealed shared layout organization and key differences in resolution, metric reporting, and spatial representation across platforms (Figure 3.1A–C). All three systems demonstrated a consistent overall layout, with patient reliability indices (e.g., fixation losses, false positives, and false negatives) positioned at the top section, and primary sensitivity maps with global indices presented in the middle and bottom regions of the printouts. The standard HFA provided a high-resolution, fully integrated layout comprising numerical grids

and corresponding pattern standard deviation (PSD) probability maps, anchored by a central raw dB sensitivity map and global indices including glaucoma hemifield test (GHT), visual field index (VFI), mean deviation (MD), and PSD. The OM VR system retained this multi-map and metric structure but exhibited reduced spatial resolution in probability maps with coarser pixelation. In contrast, the RVF platform adopted a simplified layout with similar maps and metrics as the HFA, omitting total deviation (TD) and PSD numerical grids and presenting only gradient probability maps, with reduced emphasis on explicit numerical global maps.

Table 3.2. Comparative analysis of visual field operating and technical parameters from the HFA and OM, and RVF.

Parameters	HFA	OM VR	RVF VR
Manufacturer	Zeiss	PICO x OM labs	PICO x RVF
Test Type	30-2, 24-2 & 10-2 SITA Standard	24-2 Full Threshold	30-2, 24-2 & 10-2 SITA Standard Threshold
Background illumination (cd/m^2)	10.0	1.3	1.3
Stimulus source	Projection	Display screen	Display screen
Stimulus size and colour	III, White	III, White	III, White
Interstimulus time	Random	Random	Random
Fixation control	Manual and automatic tracking	Manual and automatic tracking	Manual and automatic tracking
Refractive correction	Trial lens holder	Corrective wear	Corrective wear
Fellow eye patched	Yes	No	No

A


ZEISS
OD Single Field Analysis
Central 30-2 Threshold Test
Fixation Monitor: Blind Spot
Fixation Target: Central
Fixation Losses: 1/16
False POS Errors: 2%
False NEG Errors: 0%
Test Duration: 05:44
Fovea: Off
Stimulus: III, White
Background: 31.5 asb
Strategy: SITA Standard
Pupil Diameter:
Visual Acuity:
Rx: -4.50 DS
30°
30°
30°
Total Deviation
Pattern Deviation
GHT: Within Normal Limits
VFI24-2: 100%
MD30-2: 0.65 dB
PSD30-2: 1.41 dB
P < 5%
P < 2%
P < 1%
P < 0.5%
Comments
Signature
ZEISS
HFA 3 850-17484/1.6.0.694
Version 3.5.0.0
Created: 02/26/2026 11:23:14
Page 1 of 1

B

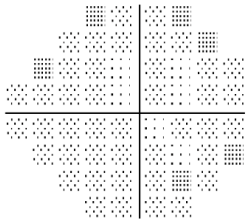
order of magnitude

OD | 24-2-Full

Fixation Monitor :	Gaze Monitor and Blindspot Monitor	Stimulus :	III, white	Date :	May 14, 2025
Fixation Target :	Central	Background :	1.3 cd/m^2	Time :	09:36 a.m.
Fixation Losses :	22.22% (9)	Strategy :	Custom-made Staircase	Age :	57 years
False Positive :	57.14% (14)	Pupil Diameter :	3		
False Negative :	55.56% (9)	Visual Acuity :	0		
Test Duration :	0 h : 07 m: 43 s	Rx :	0		
Fovea Threshold :	40 dB				

				27	30	29	26		
			30	30	29	30	30	27	
		28	31	30	33	31	33	31	30
31	30	32	30	33		32	33	32	31
29	29	29	30	32		34	30	31	30
	29	30	31	31		32	33	31	28
		29	30	32		31	28	30	
			30	31		30	31		

Numerical Sensitivity Plot



Grayscale Map

			-3	0	-2	-4			
		0	-1	-4	-1	-1	-3		
	-4	0	-2	1	-3	-1	-1	0	
0	0	-1	-3	-1	-1	-1		0	
-2	-2	-3	-3	-2	0	-4		-2	
	-3	-3	-1	-3	-1	-1	-1	-2	
		-4	-2	1	0	-4	-1		
			-1	0	0	0			

Total Deviation

			-3	0	-2	-4			
		0	-1	-4	-1	-1	-3		
	-4	0	-2	1	-3	-1	-1	0	
0	0	-1	-3	-1	-1	-1		0	
-2	-2	-3	-3	-2	0	-4	1	-2	
	-3	-3	-1	-3	-1	-1	-1	-2	
		-4	-2	1	0	-4	-1		
			-1	0	0	0			

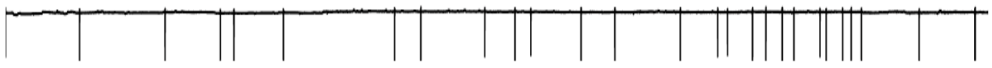
Pattern Deviation

GHT: Within normal limits

VFI 24-2: 100 %
MD 24-2: -1.50
PSD 24-2: 1.3

Test Reliability

:: P< 5%,
P < 2%,
P < 1% ,
P < 0.5%



C

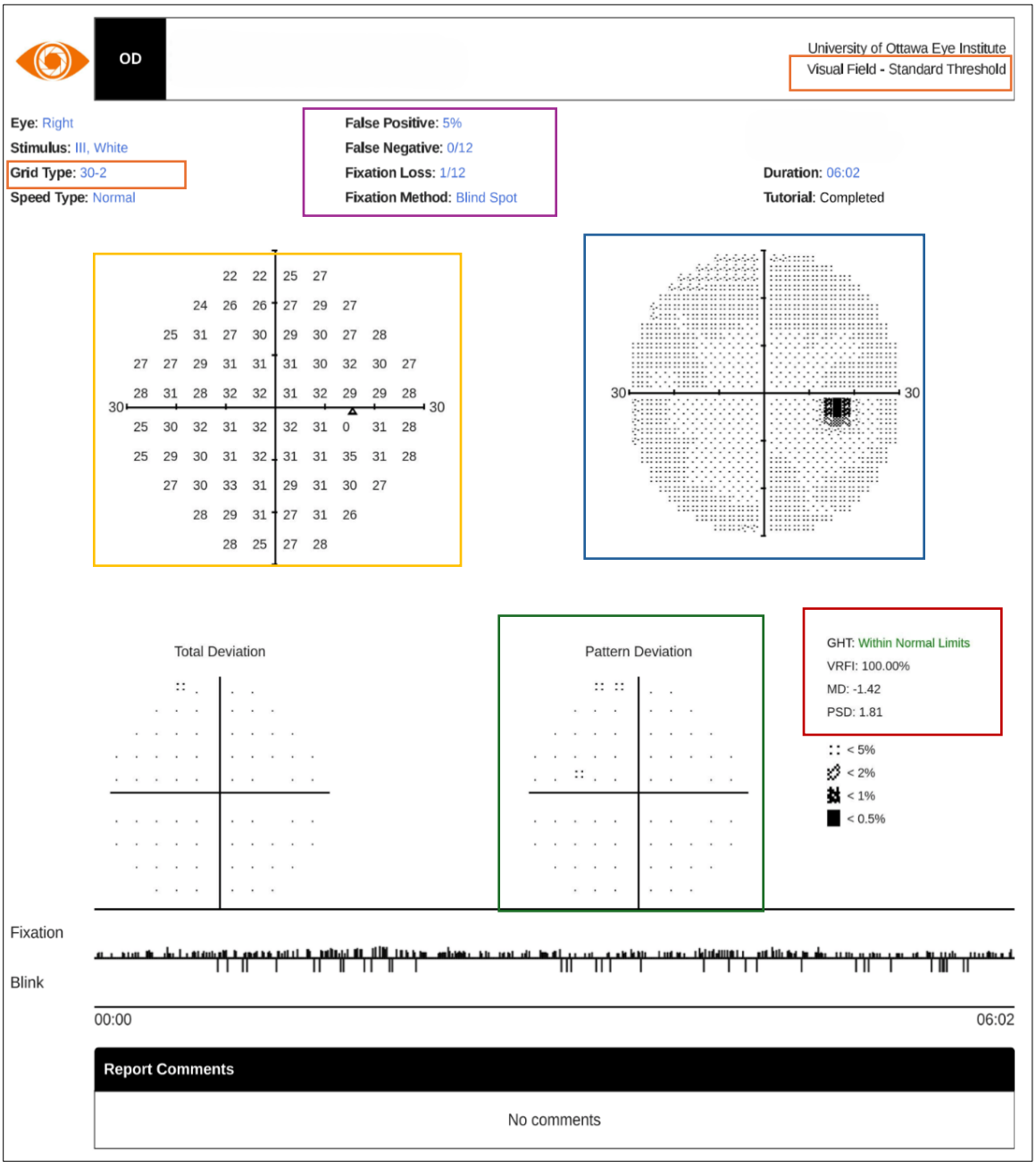
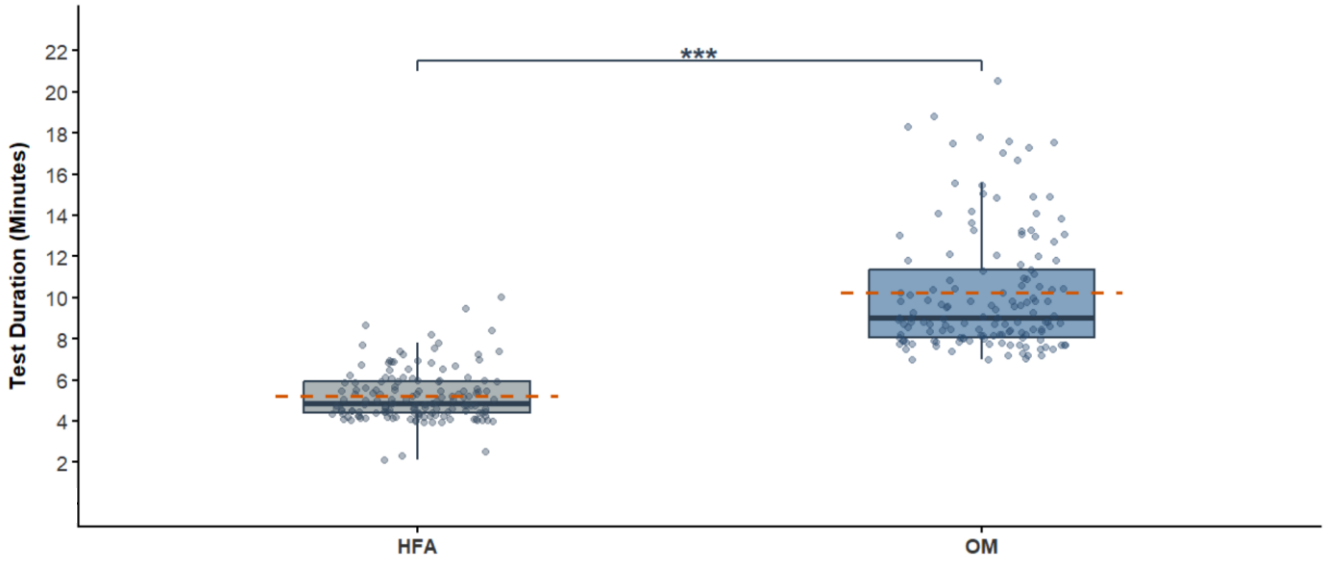


Figure 3.1. Representative visual field test printouts from normal participants across study devices. (A) Humphrey Field Analyzer; (B) Order of Magnitude; and (C) Retinalogik. Colour coding indicates key report components. Orange: Testing strategy; Purple: reliability indices; yellow: sensitivity maps; blue: grayscale maps; green: pattern deviation maps; and red: global performance metrics.

3.2.1. Test Duration of the Machines

Mean test duration differed significantly between the perimetry devices. For the 24-2 testing strategy, the OM device had a significantly longer mean test duration than the HFA (10.18 ± 0.12 min vs. 5.23 ± 0.05 min, respectively; n = 174 eyes, p < 0.001) (Figure 3.2A). In contrast, the RVF device had a significantly shorter mean test duration than the HFA (4.90 ± 0.44 min vs. 5.72 ± 1.18 min, respectively; n = 40 eyes, $p = 0.01$) (Figure 3.2B). Similarly, the RVF device also required significantly less time to complete the 30-2 testing strategy than the Humphrey Field Analyzer (6.12 ± 0.80 min vs. 7.17 ± 1.67 min, respectively; n = 364 eyes, $p < 0.001$) (Figure 3.2C).

A



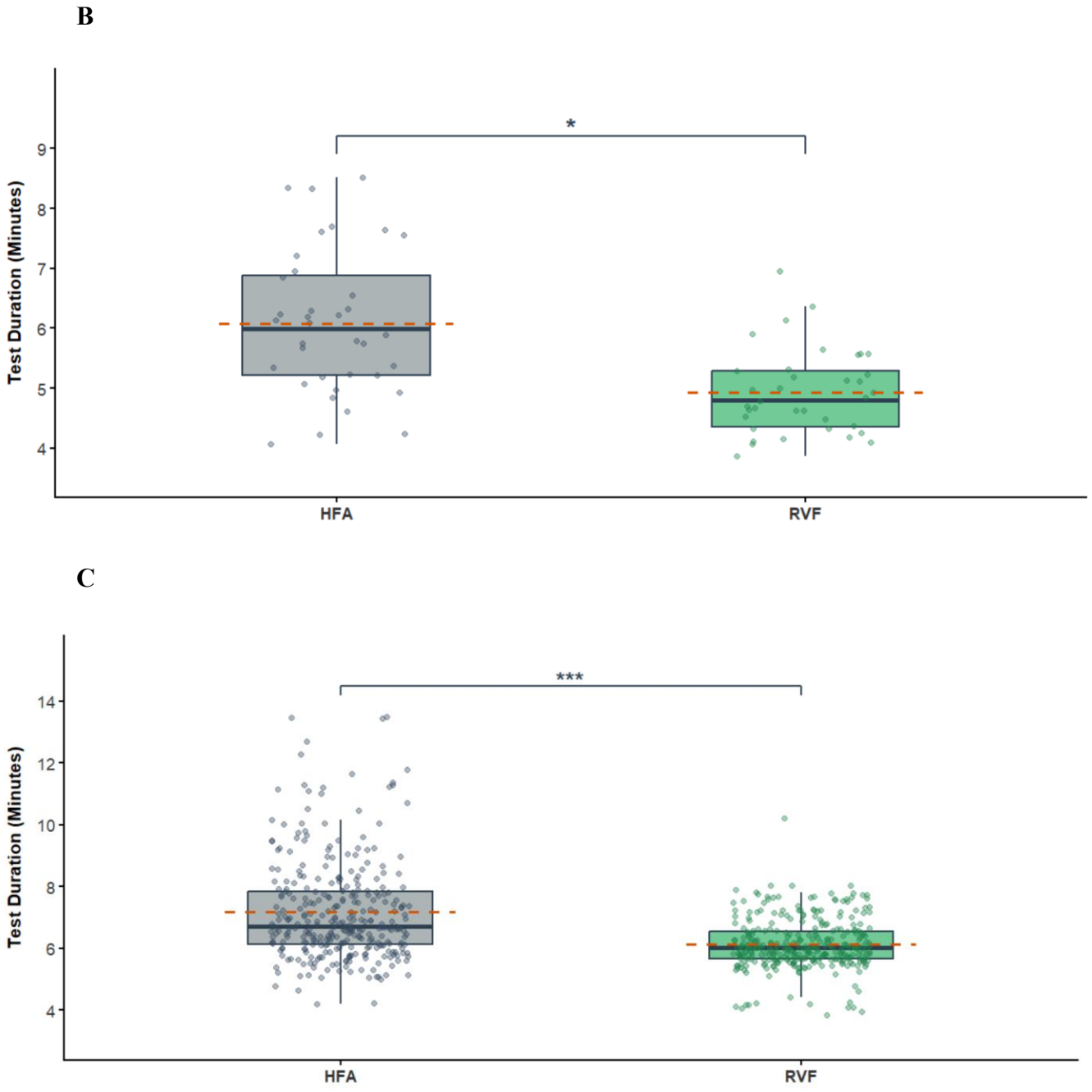


Figure 3.2. Box-and-whisker plots comparing visual field test durations per eye between the HFA and VR devices (OM and RVF) for 24-2 and 30-2 testing strategies. Distribution of visual field test durations (minutes) for (A) OM versus HFA using 24-2 testing strategy (n = 147, p < 0.001); (B) RVF versus HFA for the 24-2 testing strategy (n = 40, p = 0.01); and (C) RVF versus HFA for the 30-2 testing strategy (n = 364, p <0.001). The solid black line within each box represents the median test duration, while the dashed red line indicates the mean test duration. Individual data points are overlaid to illustrate the distribution and sample density of observations. Statistically significant values, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

3.3. Correlational Analysis

Correlational analyses were performed to evaluate the relationship between the VR-based platforms and the standard HFA for key perimetric indices. Correlational analyses revealed varying degrees of agreement between the VR platforms and the HFA. For mean deviation (MD), the OM VR system showed a moderate correlation with the HFA (Pearson's r = 0.74), whereas the RVF VR device demonstrated a strong correlation (r = 0.91) (Figure. 3.3A–B). For pattern standard deviation (PSD), the OM VR system yielded a moderate correlation coefficient of (r = 0.77), while the RVF VR device demonstrated a relatively strong correlation of (r = 0.85) (Figure 3.4A–B).

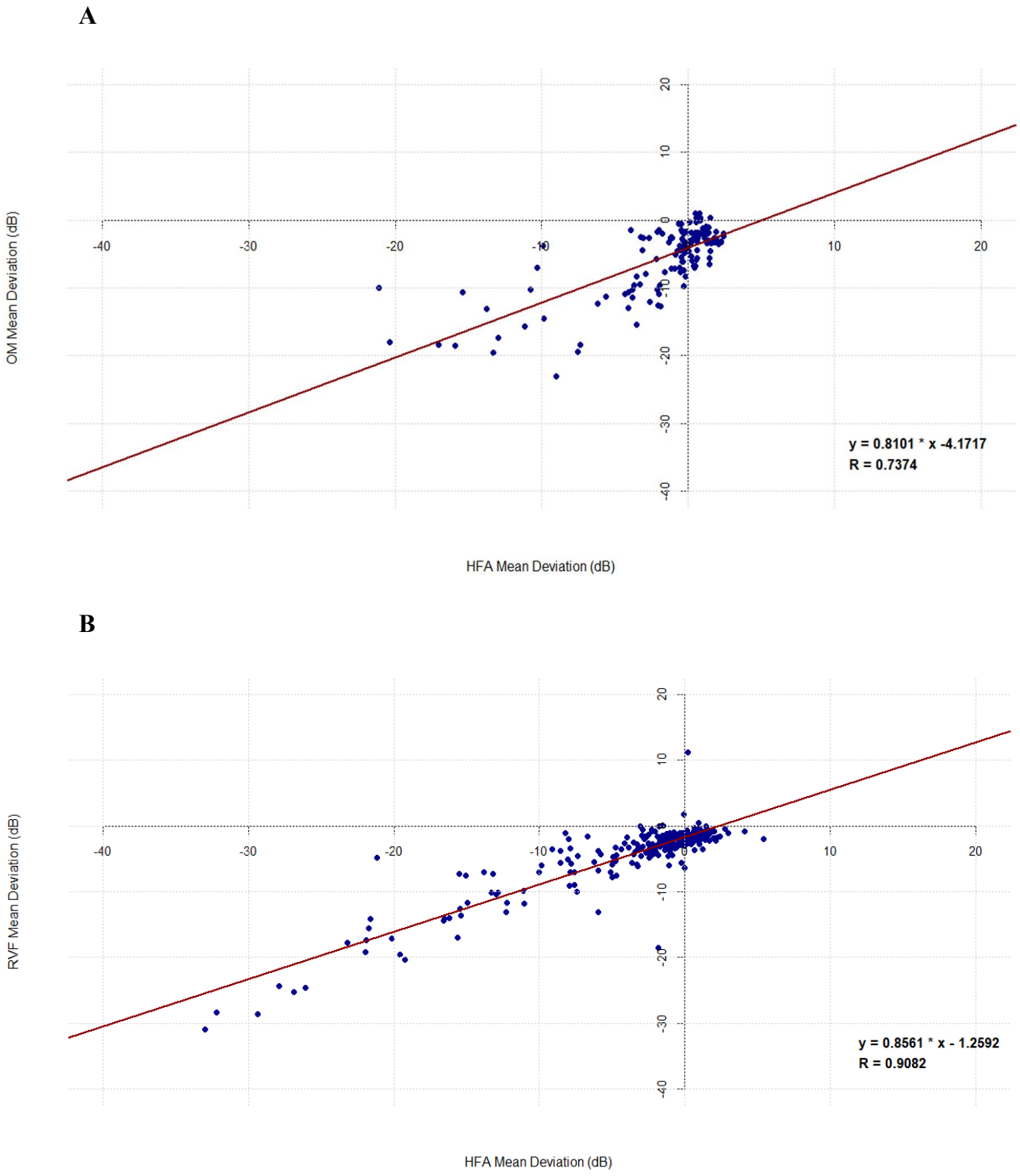


Figure 3.3. Linear regression analysis comparing mean deviation (MD) values per eye measured by the VR devices and the HFA. (A) OM device versus HFA (r = 0.74). (B) RVF device versus HFA (r = 0.91).

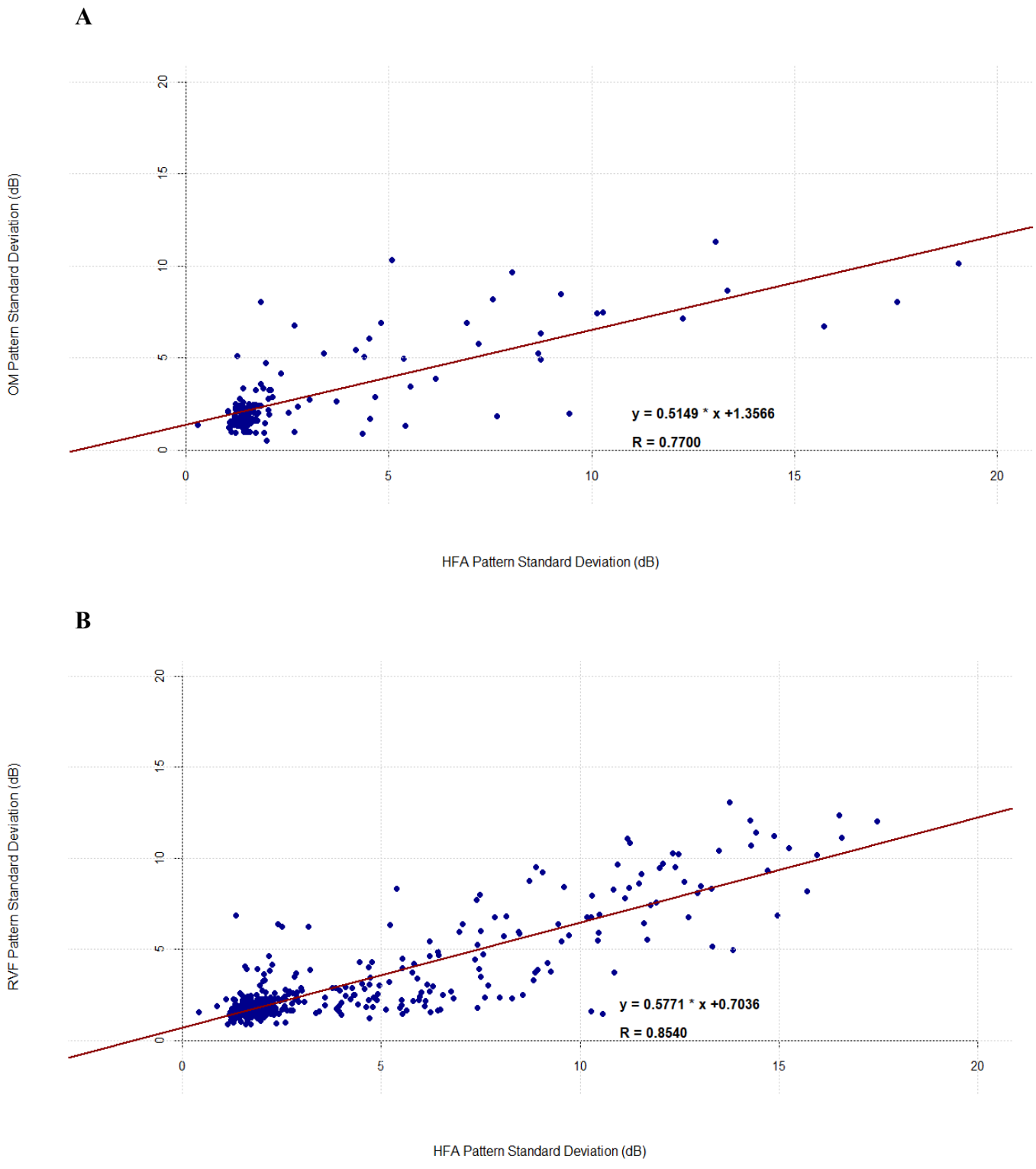


Figure 3.4. Linear regression analysis comparing pattern standard deviation (PSD) measurements per eye obtained from the VR devices and the HFA. (A) OM VR system and HFA (r = 0.77, n =147) and (B) RVF VR system and HFA (r = 0.85, n = 404).

3.4. Agreement Analysis

To assess the agreement between the novel devices (OM and RVF) and the gold-standard HFA, Bland–Altman analyses were performed for both mean deviation (MD) and pattern standard deviation (PSD). For the OM device, the Bland–Altman plots revealed a systematic underestimation of MD compared to the HFA, alongside wider variations in agreement. The mean bias was -3.85 dB, indicating that the OM device generally produced lower MD values than the HFA (Figure 3.5A). The 95% limits of agreement were relatively wide, ranging from -13.58 to 5.88 dB. The OM device also demonstrated a negligible mean bias of -0.14 dB for PSD values, though the 95% limits of agreement remained relatively broad, spanning from -5.87 to 5.58 dB (Figure 3.5B).

Compared with the OM device, the RVF device exhibited a smaller mean bias and narrower 95% limits of agreement for MD measurements (Figure 3.6A). The mean bias for MD was −1.12 dB, with 95% limits of agreement ranging from −4.76 to 2.52 dB (Figure 3.6A). For PSD measurements, the mean bias was −0.72 dB, with 95% limits of agreement ranging from −3.51 to 2.08 dB (Figure 3.6B). Overall, the RVF device exhibited smaller mean bias and narrower limits of agreement for MD compared with the OM device, whereas both devices demonstrated relatively small mean bias for PSD, although the RVF device showed narrower limits of agreement.

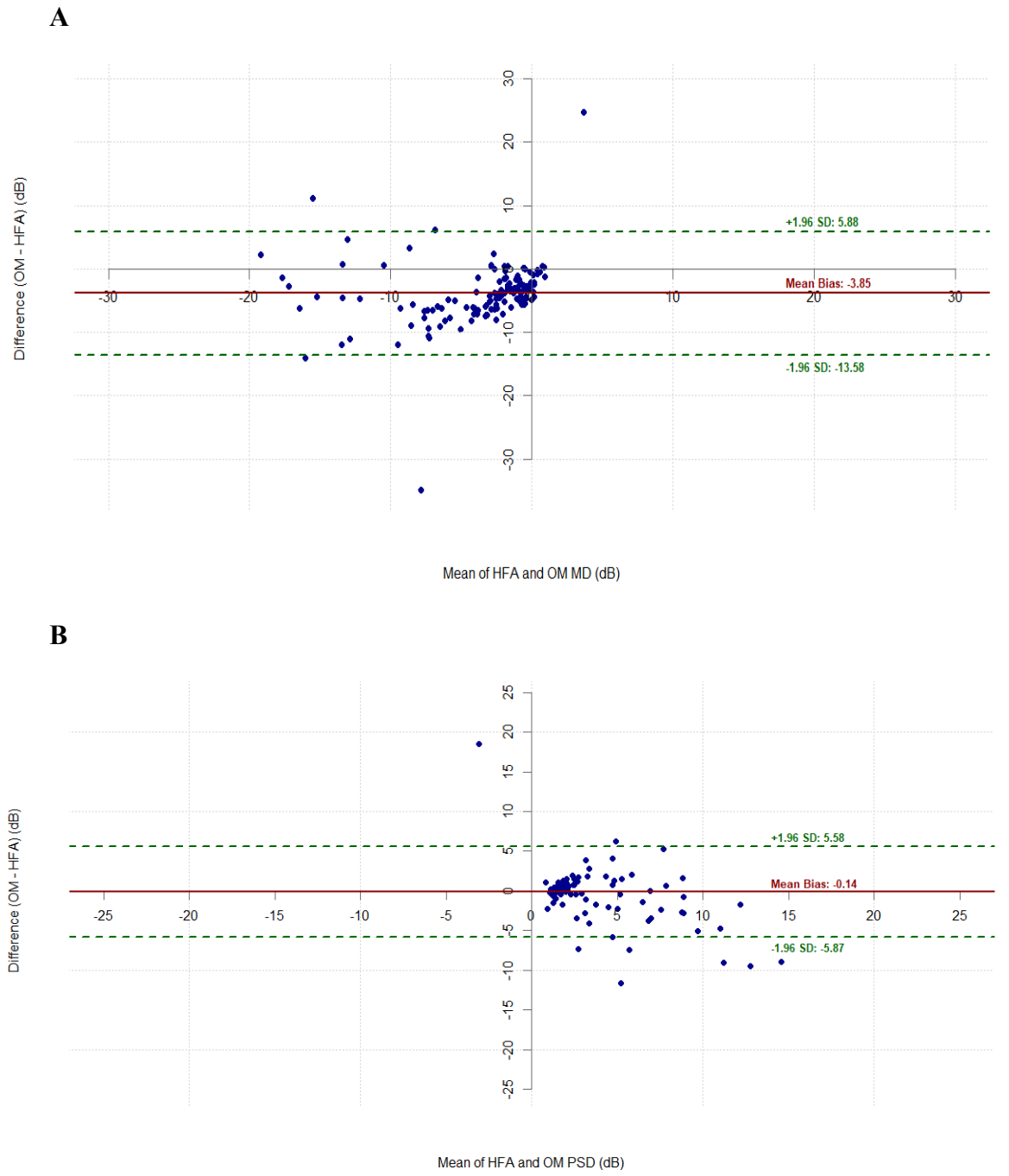


Figure 3.5. Bland–Altman analysis comparing visual field indices per eye measured by the OM device and the HFA. (A) Mean deviation (MD): The mean bias (solid red line) was −3.85 dB (n = 147), indicating that the OM device measured lower MD values than the HFA. The 95% limits of agreement (dashed green lines), calculated as the mean bias ±1.96 standard deviations, ranged from −13.58 to 5.88 dB. (B) Pattern standard deviation (PSD): The mean bias was −0.14 dB, indicating that the OM device measured slightly lower PSD values than the HFA. The 95% limits of agreement ranged from −5.87 to 5.58 dB. Each point represents a paired measurement across participants.

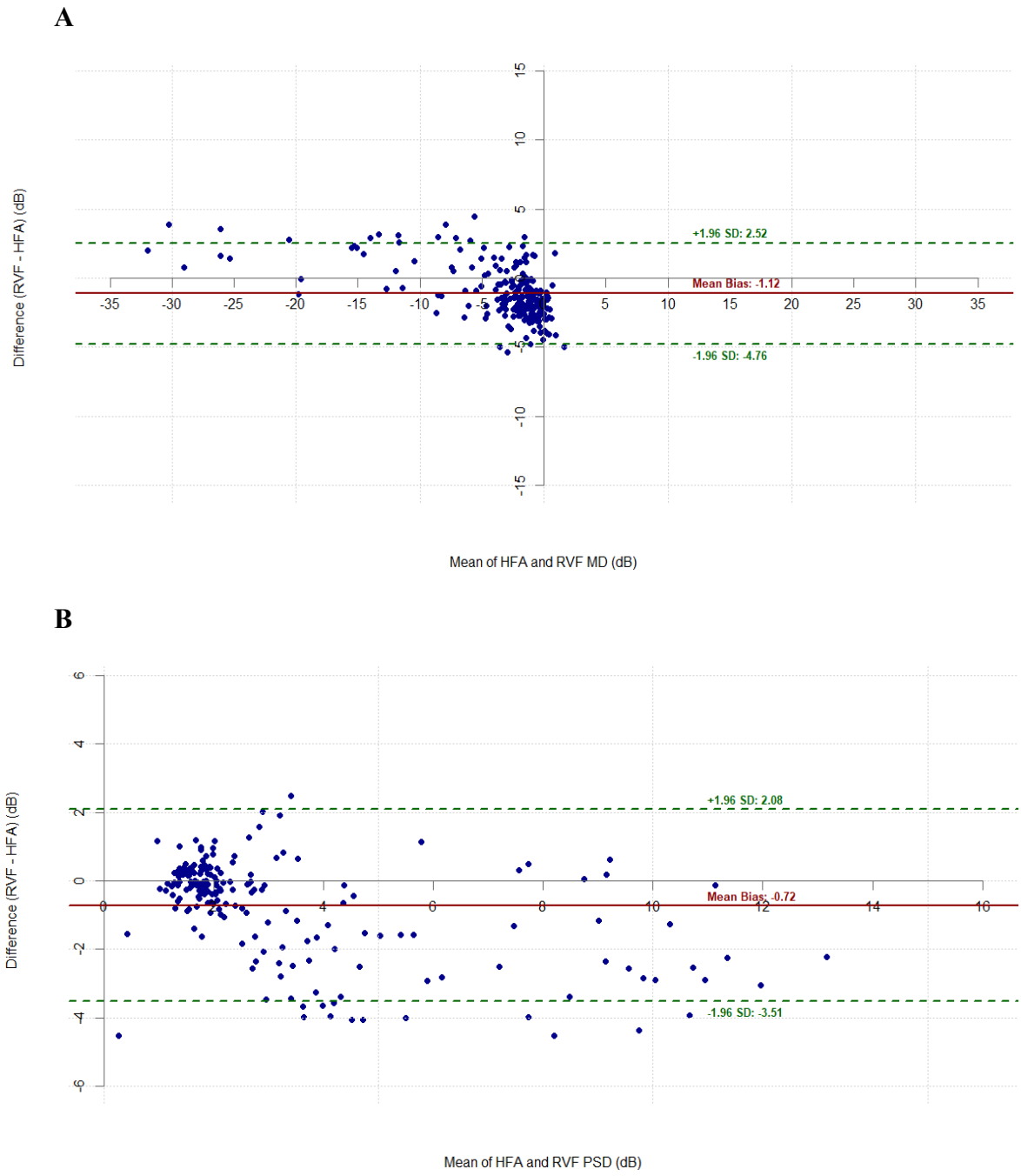


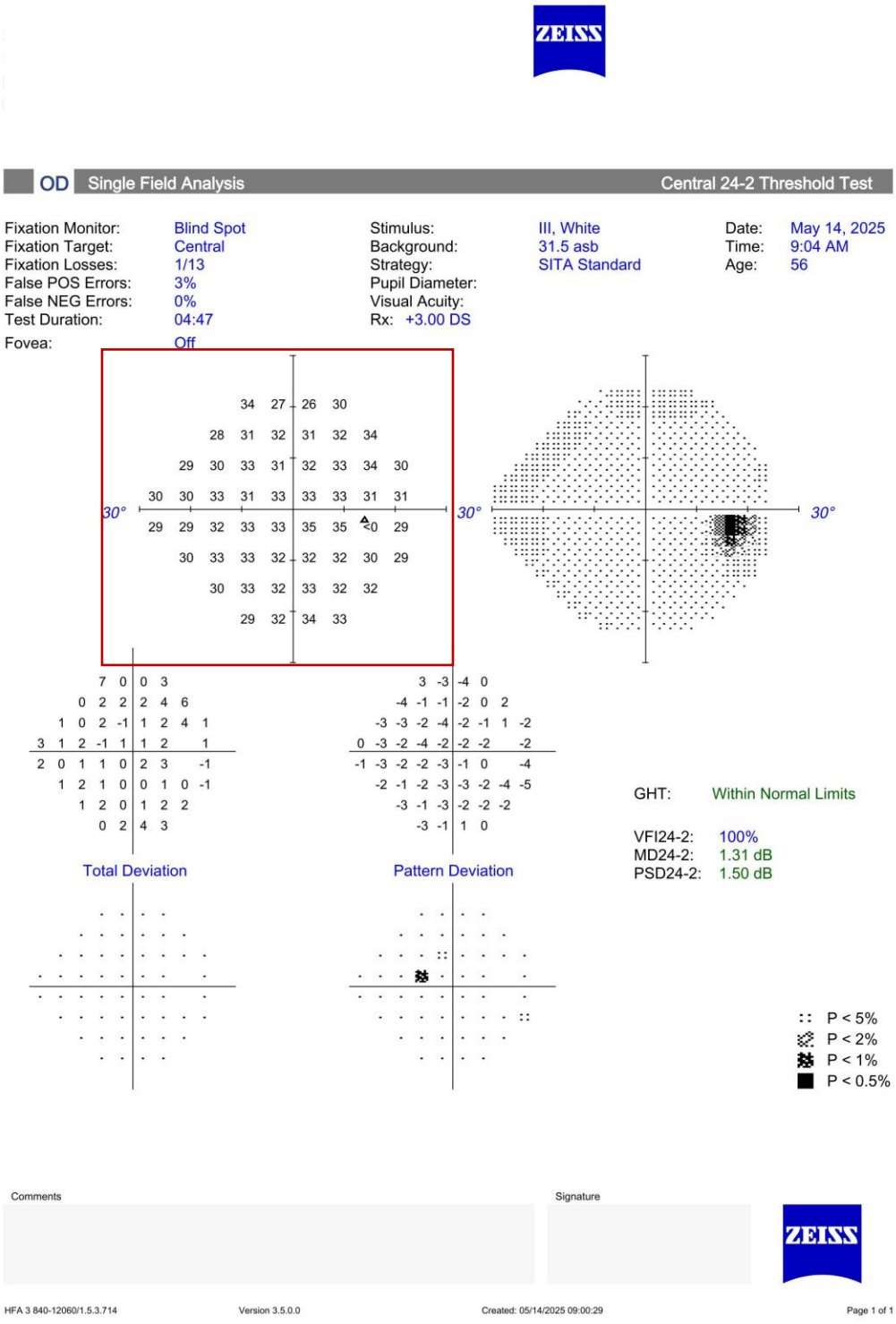
Figure 3.6. Bland–Altman analysis comparing visual field indices per eye measured by the RVF device and the HFA. (A) Mean deviation (MD): The mean bias (solid red line) was −1.12 dB, indicating that the RVF device measured slightly lower MD values than the HFA (n = 404). The 95% limits of agreement (dashed green lines), calculated as the mean bias ±1.96 standard deviations, ranged from −4.76 to 2.52 dB. (B) Pattern standard deviation (PSD): The mean bias was −0.72 dB, indicating that the RVF device measured slightly lower PSD values than the HFA. The 95% limits of agreement ranged from −3.51 to 2.08 dB. Each point represents a paired measurement from an individual participant.

3.5. Pointwise Delta Heatmaps Comparison between OM and HFA

Evaluation of pointwise sensitivity differences (Δ dB = OM − HFA) demonstrated that agreement between the OM device and HFA varies with both the location and severity of visual field loss. In a normal participant, the sensitivity delta map showed overall strong agreement between devices, with predominantly neutral values indicating minimal threshold deviation (Figure 3.7A−C). Scattered peripheral points demonstrated minor, non-systematic positive deviations, consistent with small physiological variability in otherwise normal visual fields.

In a patient with early-stage glaucoma with a nasal step defect, the sensitivity delta map revealed a localized cluster of negative deviations corresponding to the superior-nasal field (Figure 3.8A−C). This indicates lower sensitivity estimates in the OM system relative to HFA within the region of early structural loss, with clearer divergence confined to the affected area. In a case of advanced glaucoma with a dense arcuate defect extending toward central vision, the sensitivity delta map demonstrated marked spatial heterogeneity (Figure 3.9A−C). The core of the defect showed close inter-device agreement consistent with absolute field loss, while transitional regions at the margins of the scotoma exhibited clusters of negative deviations, indicating greater underestimation of sensitivity by the OM system relative to HFA in areas of partial loss. Overall, OM demonstrated good agreement with HFA in both normal retinal regions and zones of absolute field loss. However, in areas of partial or transitional field defect, OM showed greater deviation, with a tendency toward lower relative sensitivity estimates compared with the standard HFA.

A


OD Single Field Analysis
Central 24-2 Threshold Test
Fixation Monitor: Blind Spot
Fixation Target: Central
Fixation Losses: 1/13
False POS Errors: 3%
False NEG Errors: 0%
Test Duration: 04:47
Fovea: Off
Stimulus: III, White
Background: 31.5 asb
Strategy: SITA Standard
Pupil Diameter:
Visual Acuity:
Rx: +3.00 DS
Date: May 14, 2025
Time: 9:04 AM
Age: 56
30°
30°
30°
Total Deviation
Pattern Deviation
GHT: Within Normal Limits
VFI24-2: 100%
MD24-2: 1.31 dB
PSD24-2: 1.50 dB
P < 5%
P < 2%
P < 1%
P < 0.5%
Comments
Signature
© 2019 Carl Zeiss Meditec, Inc. All rights reserved.
HFA 3 840-12060/1.5.3.714
Version 3.5.0.0
Created: 05/14/2025 09:00:29
Page 1 of 1

B

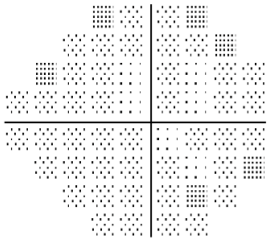
order of magnitude

OD 24-2-Full

Fixation Monitor :	Gaze Monitor and Blindspot Monitor	Stimulus :	III, white	Date :	May 14, 2025
		Background :	1.3 cd/m^2	Time :	09:36 a.m.
Fixation Target :	Central	Strategy :	Custom-made Staircase	Age :	57 years
Fixation Losses :	22.22% (9)	Pupil Diameter :	3		
False Positive :	57.14% (14)	Visual Acuity :	0		
False Negative :	55.56% (9)	Rx :	0		
Test Duration :	0 h : 07 m: 43 s				
Fovea Threshold :	40 dB				

			27	30	29	26			
		30	30	29	30	30	27		
	28	31	30	33	31	33	31	30	
31	30	32	30	33	32	33	32	31	
29	29	29	30	32	34	30	31	30	
	29	30	31	31	32	33	31	28	
		29	30	32	31	28	30		
			30	31	30	31			

Numerical Sensitivity Plot



Grayscale Map

			-3	0	-2	-4			
		0	-1	-4	-1	-1	-3		
	-4	0	-2	1	-3	-1	-1	0	
0	0	-1	-3	-1	-1	-1		0	
-2	-2	-3	-3	-2	0	-4		-2	
	-3	-3	-1	-3	-1	-1	-1	-2	
		-4	-2	1	0	-4	-1		
			-1	0	0	0			

Total Deviation

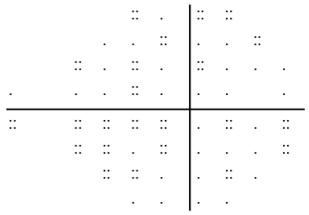
			-3	0	-2	-4			
		0	-1	-4	-1	-1	-3		
	-4	0	-2	1	-3	-1	-1	0	
0	0	-1	-3	-1	-1	-1		0	
-2	-2	-3	-3	-2	0	-4	1	-2	
	-3	-3	-1	-3	-1	-1	-1	-2	
		-4	-2	1	0	-4	-1		
			-1	0	0	0			

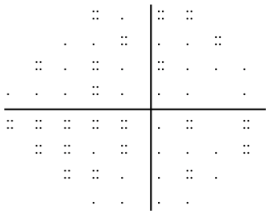
Pattern Deviation

GHT: Within normal limits

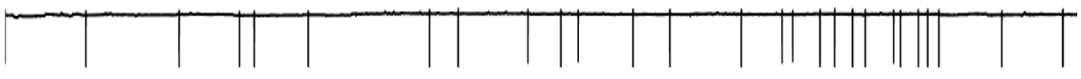
VFI 24-2: 100 %
MD 24-2: -1.50
PSD 24-2: 1.3

Test Reliability





:: P< 5%,
P < 2%,
P < 1% ,
P < 0.5%



C

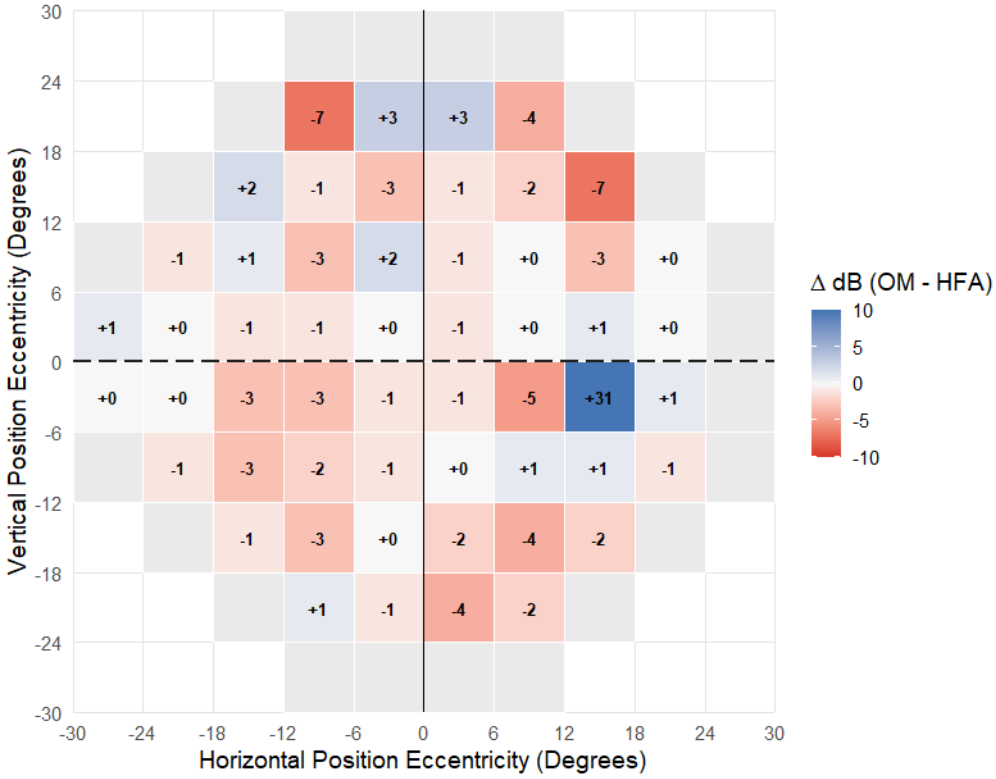
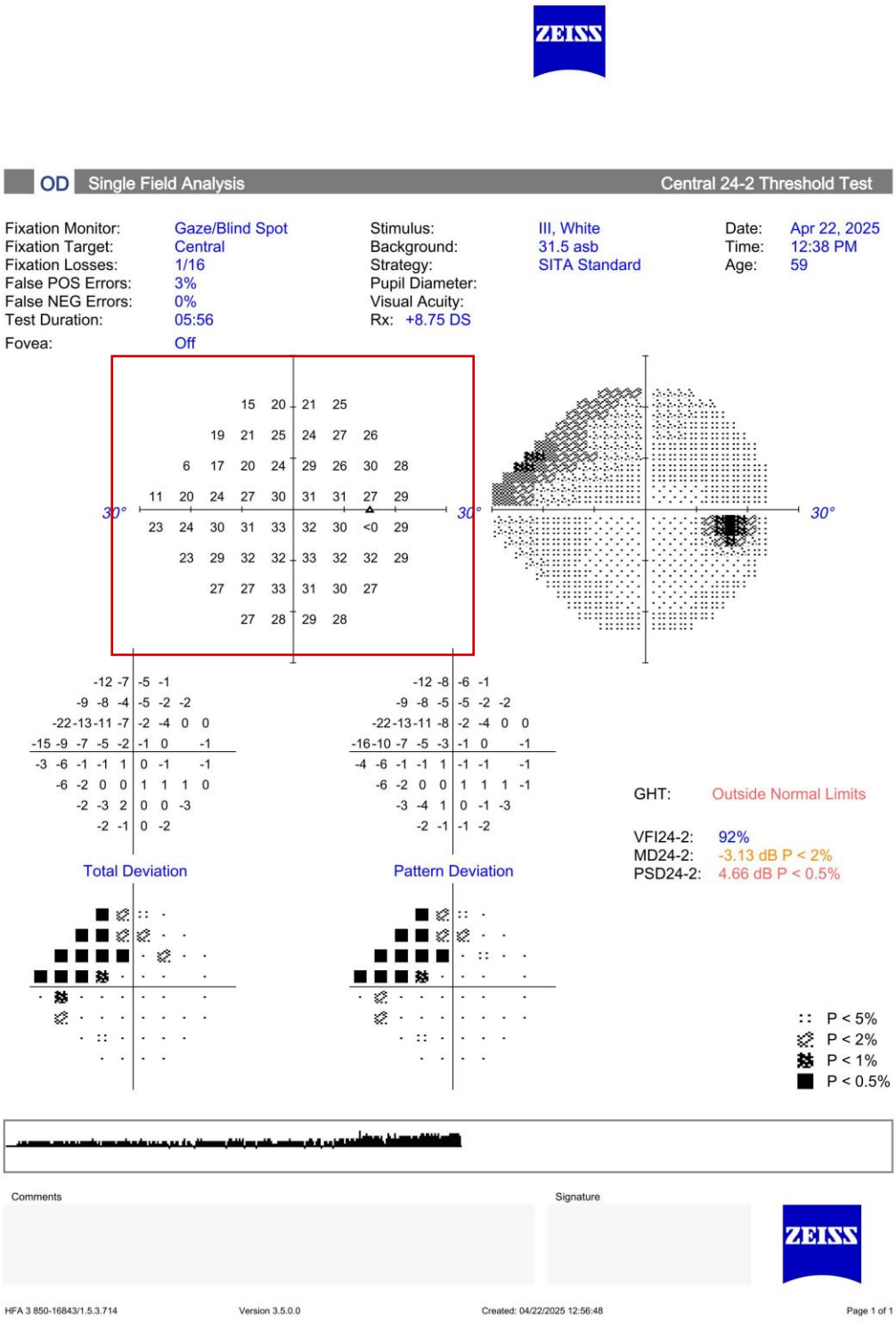


Figure 3.7. Sensitivity mapping comparison for a normal participant between OM and HFA. Perimetric outputs are displayed for the standard-of-care HFA and the OM platform. (A) Standard HFA diagnostic printout and (B) corresponding OM VR diagnostic printout displaying sensitivity threshold maps (red). (C) Spatial delta heatmaps establishing the localized pointwise point-by-point sensitivity differences (Δ dB = OM − HFA) between the two testing modalities across the tested visual field points. The diverging color scale represents the magnitude and directional variance of the threshold sensitivity deviations between the two devices. Blue signifies positive values indicating threshold overestimation by the VR platform relative to the HFA, and red signifies negative values indicating threshold underestimation.

A


OD Single Field Analysis
Central 24-2 Threshold Test
Fixation Monitor: Gaze/Blind Spot
Fixation Target: Central
Fixation Losses: 1/16
False POS Errors: 3%
False NEG Errors: 0%
Test Duration: 05:56
Fovea: Off
Stimulus: III, White
Background: 31.5 asb
Strategy: SITA Standard
Pupil Diameter:
Visual Acuity:
Rx: +8.75 DS
Date: Apr 22, 2025
Time: 12:38 PM
Age: 59
30°
30°
30°
Total Deviation
Pattern Deviation
GHT: Outside Normal Limits
VFI24-2: 92%
MD24-2: -3.13 dB P < 2%
PSD24-2: 4.66 dB P < 0.5%
P < 5%
P < 2%
P < 1%
P < 0.5%
Comments
Signature
© 2019 Carl Zeiss Meditec, Inc. All rights reserved.
HFA 3 850-16843/1.5.3.714
Version 3.5.0.0
Created: 04/22/2025 12:56:48
Page 1 of 1

B

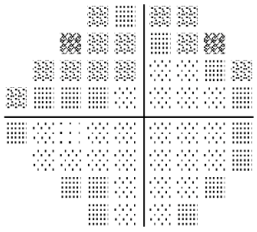
order of magnitude

OD 24-2-Full

Fixation Monitor :	Gaze Monitor and Blindspot Monitor	Stimulus :	III, white	Date :	Apr 22, 2025
		Background :	1.3 cd/m^2	Time :	12:14 p.m.
Fixation Target :	Central	Strategy :	Custom-made Staircase	Age :	60 years
Fixation Losses :	100.00% (9)	Pupil Diameter :	3		
False Positive :	30.00% (10)	Visual Acuity :	0		
False Negative :	88.89% (9)	Rx :	0		
Test Duration :	0 h : 09 m: 51 s				
Fovea Threshold :	28 dB				

			21	25	22	24			
		18	24	23	26	23	20		
	22	21	24	23	29	29	27	24	
22	25	28	27	32	32	31	30	27	
28	29	34	29	32	31	30	30	28	
	30	30	31	32	30	29	30	28	
		28	28	30	30	29	28		
			27	29	32	28			

Numerical Sensitivity Plot



Grayscale Map

			-9	-5	-9	-6			
		-12	-7	-10	-5	-8	-10		
	-10	-10	-8	-9	-5	-5	-5	-6	
-9	-5	-5	-6	-2	-1	-3		-4	
-3	-2	2	-4	-2	-3	-4		-4	
	-2	-3	-1	-2	-3	-5	-2	-2	
		-5	-4	-1	-1	-3	-3		
			-4	-2	2	-3			

Total Deviation

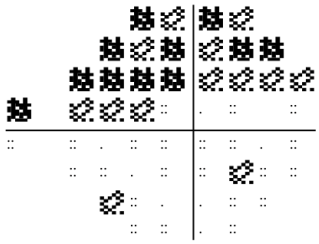
			-8	-4	-8	-5			
		-11	-6	-9	-4	-7	-9		
	-9	-9	-7	-8	-4	-4	-4	-5	
-8	-4	-4	-5	-1	0	-2		-3	
-2	-1	3	-3	-1	-2	-3	1	-3	
	-1	-2	0	-1	-2	-4	-1	-1	
		-4	-3	0	0	-2	-2		
			-3	-1	3	-2			

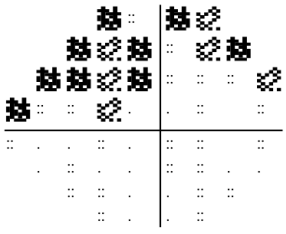
Pattern Deviation

GHT: Outside normal limits

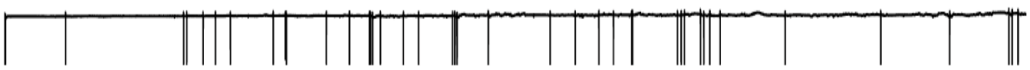
VFI 24-2: 92 %
MD 24-2: -4.48
PSD 24-2: 2.88

Test Reliability





:: P< 5%,
P < 2%,
P < 1% ,
P < 0.5%



C

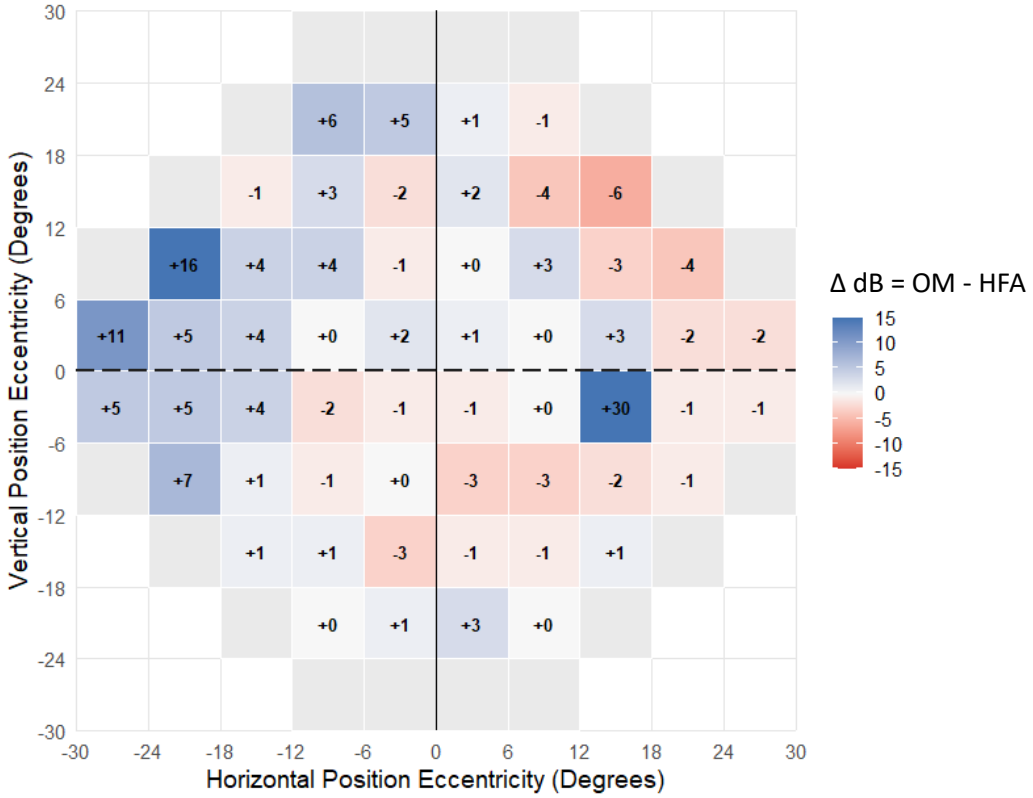
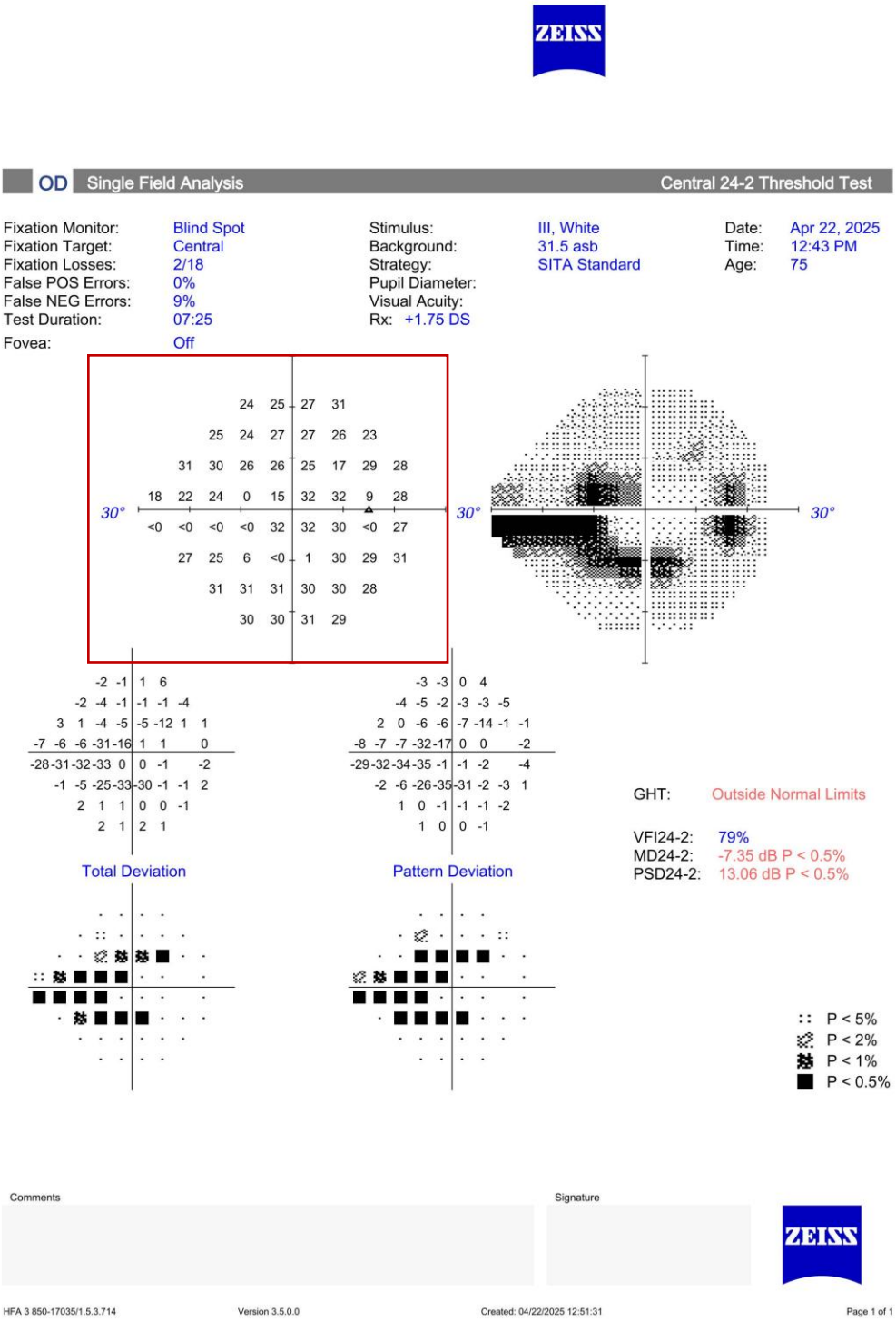


Figure 3.8. Sensitivity mapping comparison for a patient with a nasal step from early-stage glaucoma between OM and HFA. Perimetric outputs are displayed for the standard-of-care HFA and the OM platform. (A) Standard HFA diagnostic printout and (B) corresponding OM VR diagnostic printout displaying sensitivity threshold maps (red). (C) Spatial delta heatmaps establishing the localized pointwise point-by-point sensitivity differences (Δ dB = OM − HFA) between the two testing modalities across the tested visual field points. The diverging color scale represents the magnitude and directional variance of the threshold sensitivity deviations between the two devices. Blue signifies positive values indicating threshold overestimation by the VR platform relative to the HFA, and red signifies negative values indicating threshold underestimation.

A

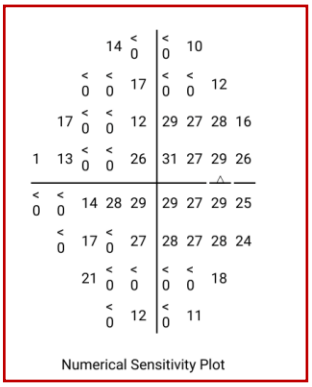


B

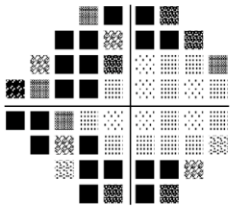
order of magnitude

OD 24-2-Full

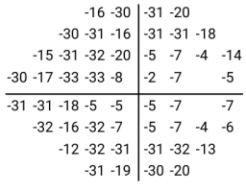
Fixation Monitor :	Gaze Monitor and Blindspot Monitor	Stimulus :	III, white	Date :	Apr 22, 2025
Fixation Target :	Central	Background :	1.3 cd/m^2	Time :	01:25 p.m.
Fixation Losses :	88.89% (9)	Strategy :	Custom-made Staircase	Age :	76 years
False Positive :	68.42% (19)	Pupil Diameter :	3		
False Negative :	100.00% (9)	Visual Acuity :	0		
Test Duration :	0 h : 17 m: 36 s	Rx :	0		
Fovea Threshold :	17 dB				



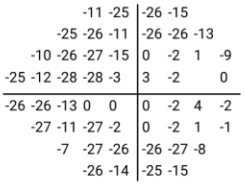
Numerical Sensitivity Plot



Grayscale Map



Total Deviation

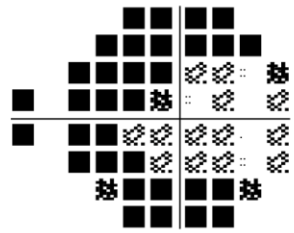


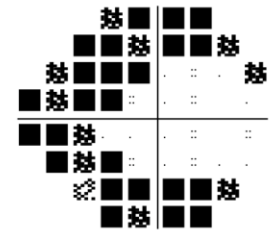
Pattern Deviation

GHT: Outside normal limits

VFI 24-2: 49 %
MD 24-2: -18.37
PSD 24-2: 11.3

Test Reliability





:: P< 5%,
P < 2%,
P < 1% ,
P < 0.5%

C

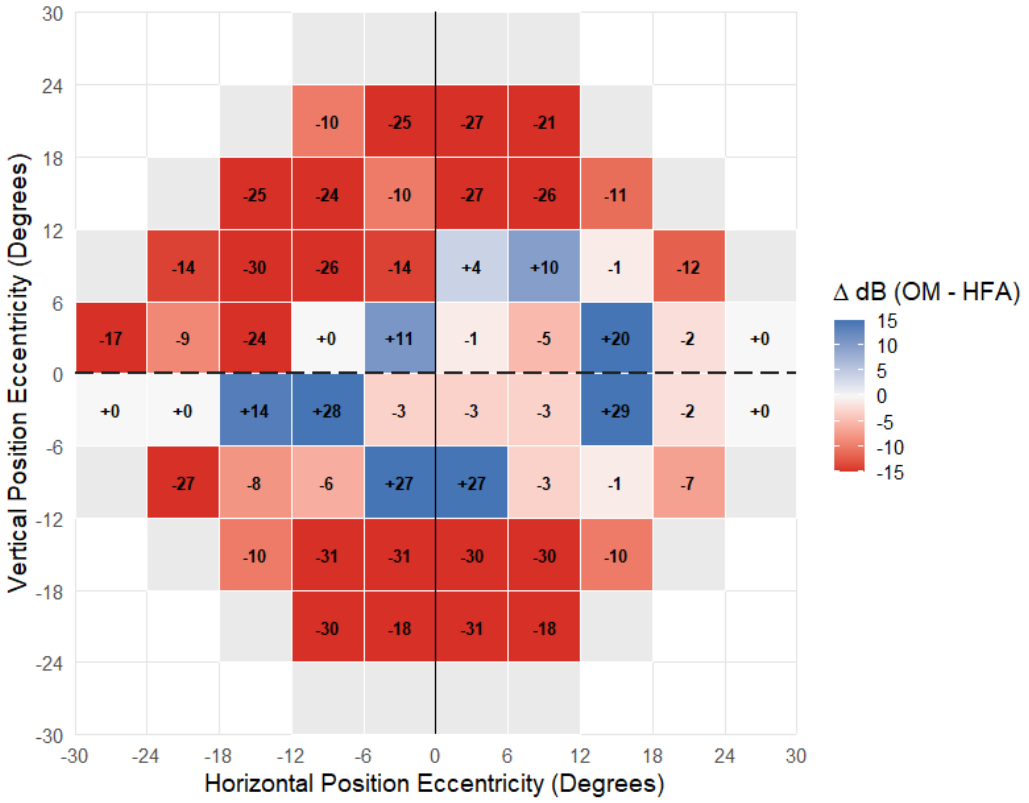


Figure 3.9. Sensitivity mapping comparison for a patient with arcuate defects from advanced-stage glaucoma between OM and HFA. Perimetric outputs are displayed for the standard-of-care HFA and the OM platform. (A) Standard HFA diagnostic printout and (B) corresponding OM VR diagnostic printout displaying sensitivity threshold maps (red). (C) Spatial delta heatmaps establishing the localized pointwise point-by-point sensitivity differences (Δ dB = OM − HFA) between the two testing modalities across the tested visual field points. The diverging color scale represents the magnitude and directional variance of the threshold sensitivity deviations between the two devices. Blue signifies positive values indicating threshold overestimation by the VR platform relative to the HFA, and red signifies negative values indicating threshold underestimation.

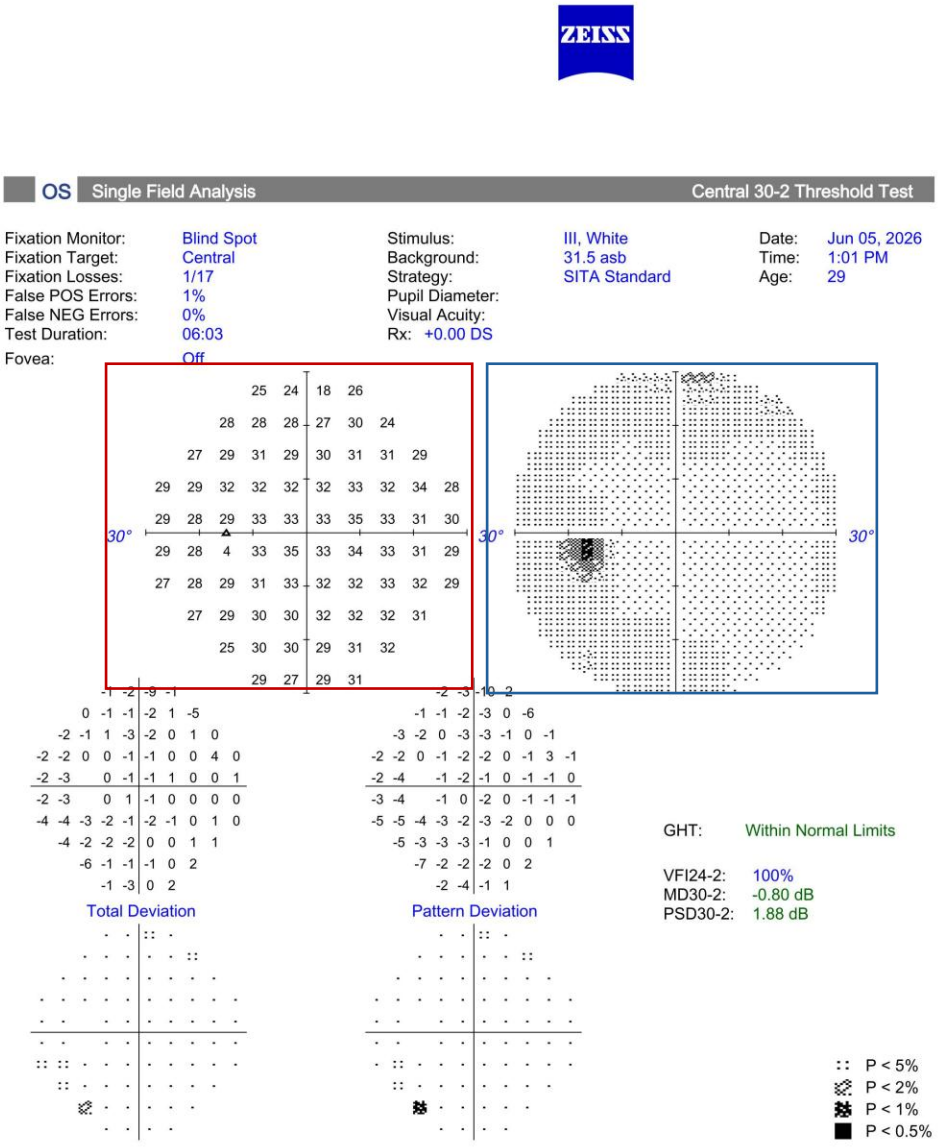
3.6. Pointwise Delta Heatmaps and Spatial Profiling Comparison for Representative Cases between RVF and HFA

Normal Case

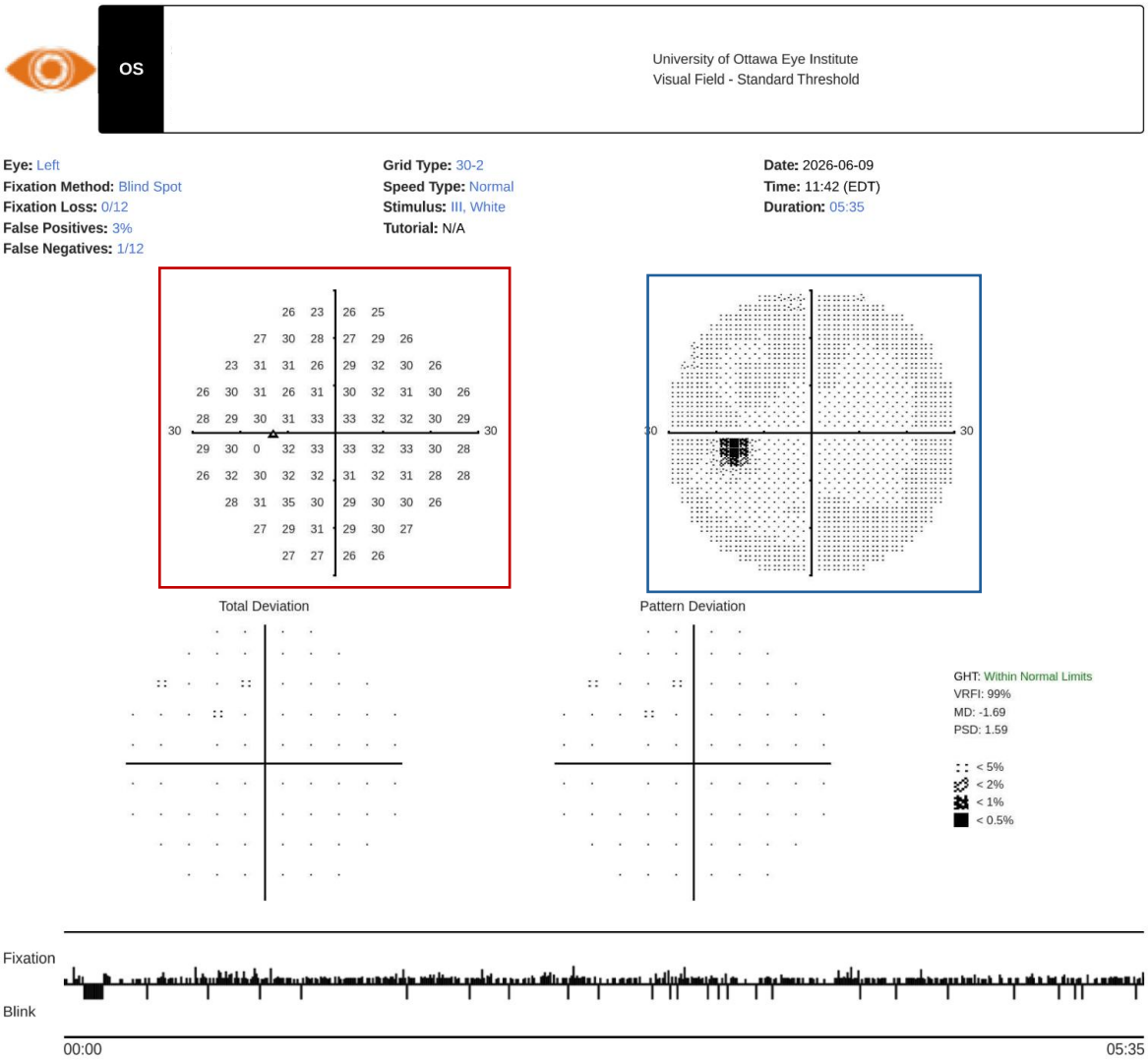
In the normal participant, the delta map demonstrated values tightly clustered around 0 dB across the visual field, with most locations ranging between approximately −1 dB and +2 dB (Figure 3.10A−C) A small number of peripheral outliers were observed, reaching up to +8 dB at the superior edge. Overall, this indicates close pointwise agreement between the RVF and HFA platforms in a normal visual field, with minimal systematic deviation.

An evaluation of threshold sensitivity profiles as a function of visual field eccentricity (mapped using 2° spatial bins) was also performed to assess how visual field defect depth, defined as 0% representing normal physiological sensitivity and 100% representing an absolute scotoma (0 dB), influences spatial concordance between the VR perimetry platforms and the standard HFA. Spatial profiling of a normal participant case demonstrated close agreement between RVF ad HFA, with defect depths remaining below 5% throughout the surrounding visual field. Along the horizontal profile, the machines localized the blind spot between approximately −20° and −10° eccentricity, with peak defect depths exceeding 95%. The RVF profile showed a slightly broader transition at the blind spot boundaries compared with HFA (Figure. 3.10D). The vertical profile demonstrated similar agreement, with both machines capturing the sharp reduction in sensitivity corresponding to the blind spot (Figure. 3.10E). The RVF profile also showed slightly broader transitions at the defect margins compared with HFA. Overall, the RVF machine demonstrated strong agreement to the HFA in the localization and depth of the physiological blind spot, with only minor differences at the defect boundaries.

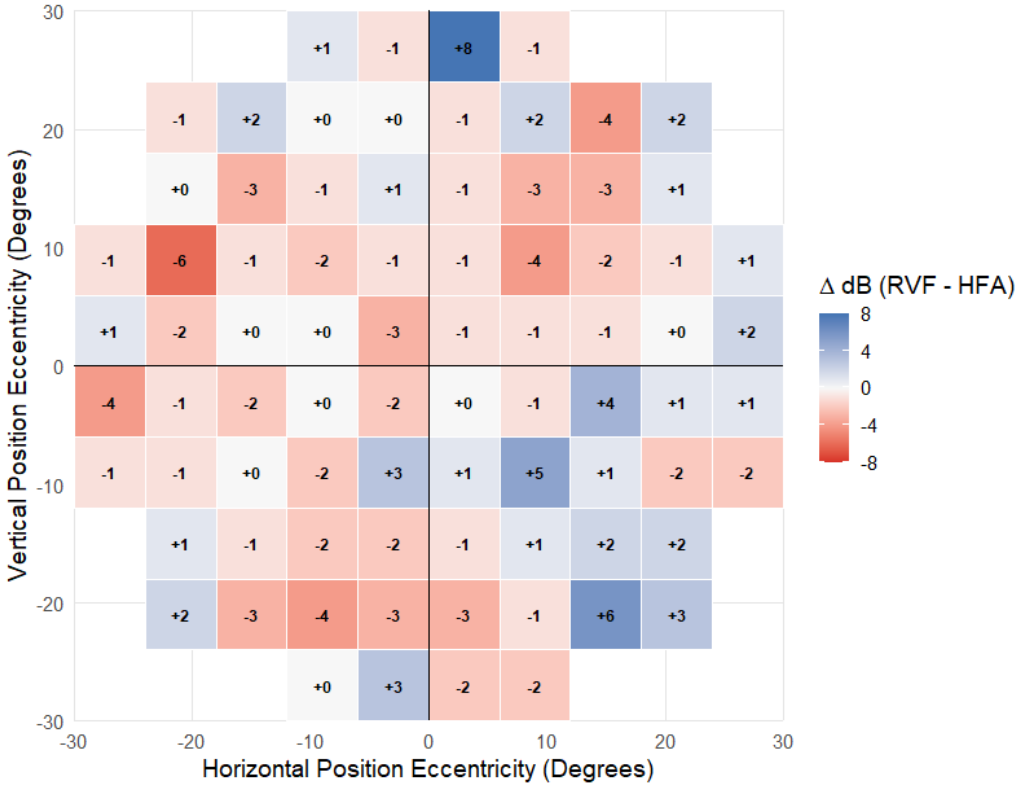
A



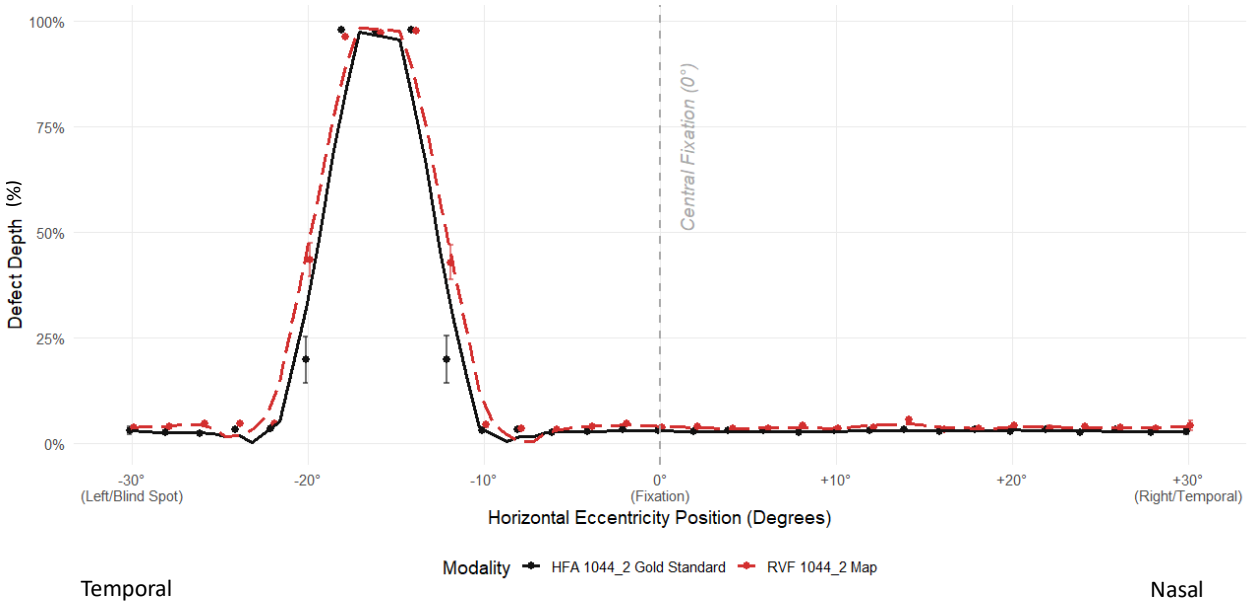
B


OS
University of Ottawa Eye Institute
Visual Field - Standard Threshold
Eye: Left
Fixation Method: Blind Spot
Fixation Loss: 0/12
False Positives: 3%
False Negatives: 1/12
Grid Type: 30-2
Speed Type: Normal
Stimulus: III, White
Tutorial: N/A
Date: 2026-06-09
Time: 11:42 (EDT)
Duration: 05:35
Total Deviation
Pattern Deviation
GHT: Within Normal Limits
VRFI: 99%
MD: -1.69
PSD: 1.59
< 5%
< 2%
< 1%
< 0.5%
Fixation
Blink
00:00
05:35

C



D



Temporal Nasal

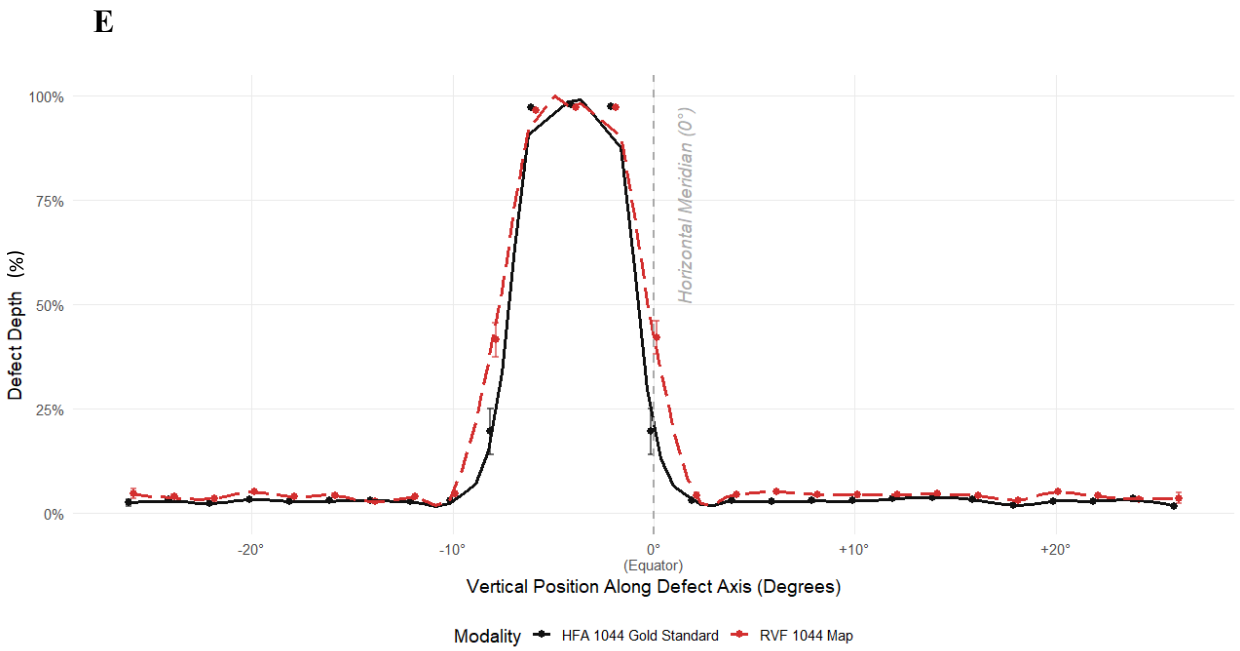
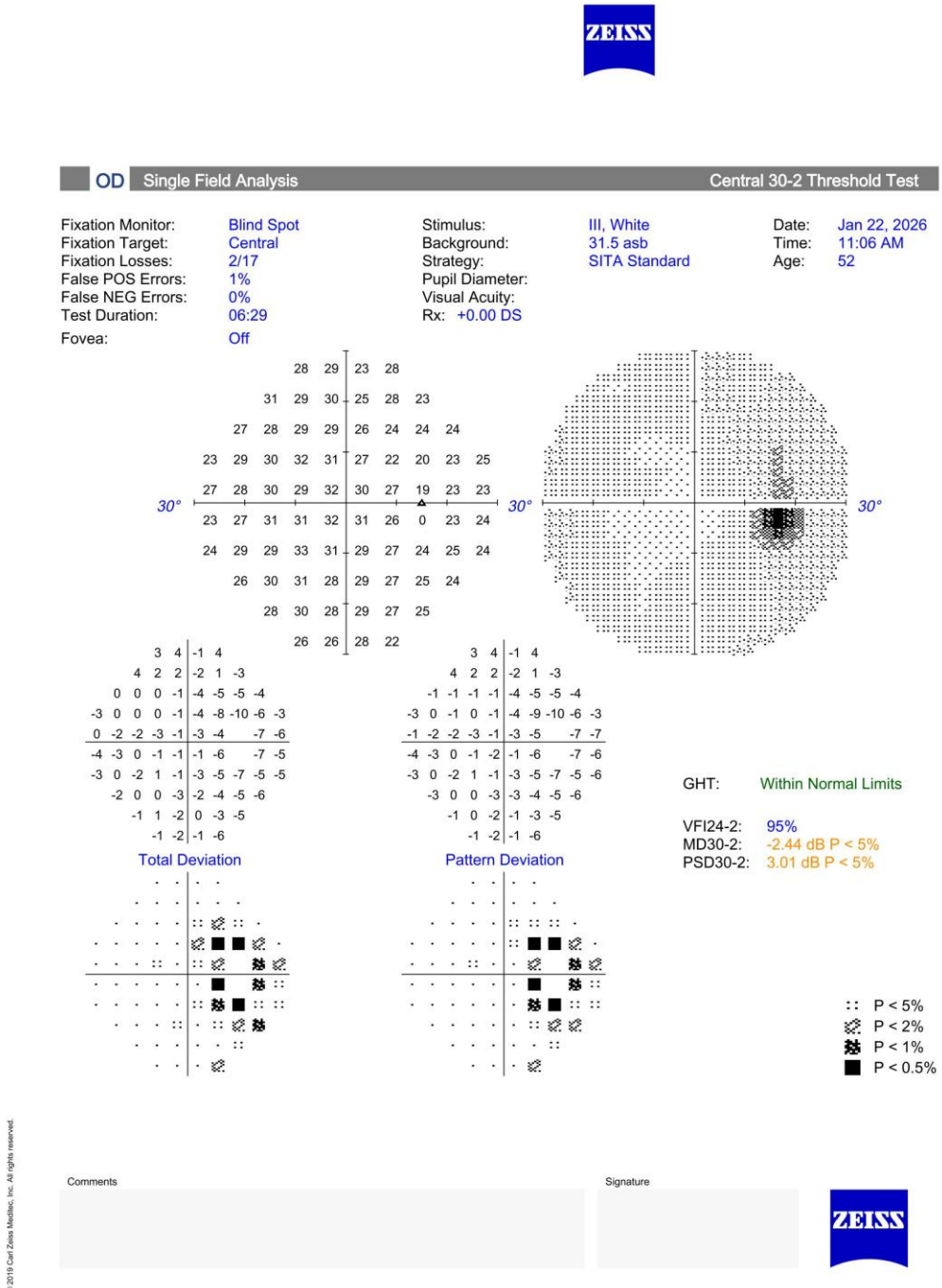


Figure 3.10. Sensitivity mapping comparison for a normal participant of the left eye between RVF and HFA. Perimetric outputs are displayed for the standard-of-care HFA and the RVF platform. (A) Standard HFA diagnostic printout and (B) corresponding RVF diagnostic printout displaying sensitivity (red) and greyscale maps (blue). (C) Spatial delta heatmaps establishing the localized pointwise point-by-point sensitivity differences (Δ dB = RVF − HFA) between the two testing modalities across the tested visual field points. The diverging color scale represents the magnitude and directional variance of the threshold sensitivity deviations between the two devices. Blue signifies positive values indicating threshold overestimation by the VR platform relative to the HFA, and red signifies negative values indicating threshold underestimation. (D) The spatial mapping of visual field defect depth and eccentricity. Defect depth is plotted as a function of horizontal eccentricity position (x = −18° to −14°) from temporal (left) to nasal (right). (E) Corresponding vertical slice analysis showing the distribution of perimetric sensitivity values (dB) as a function of visual field eccentricity (y = −4° to −1°). Black points represent HFA measurements and red points represent RVF measurements. Trendlines illustrate the decline in sensitivity from central fixation toward peripheral eccentricities using 2° spatial bins.

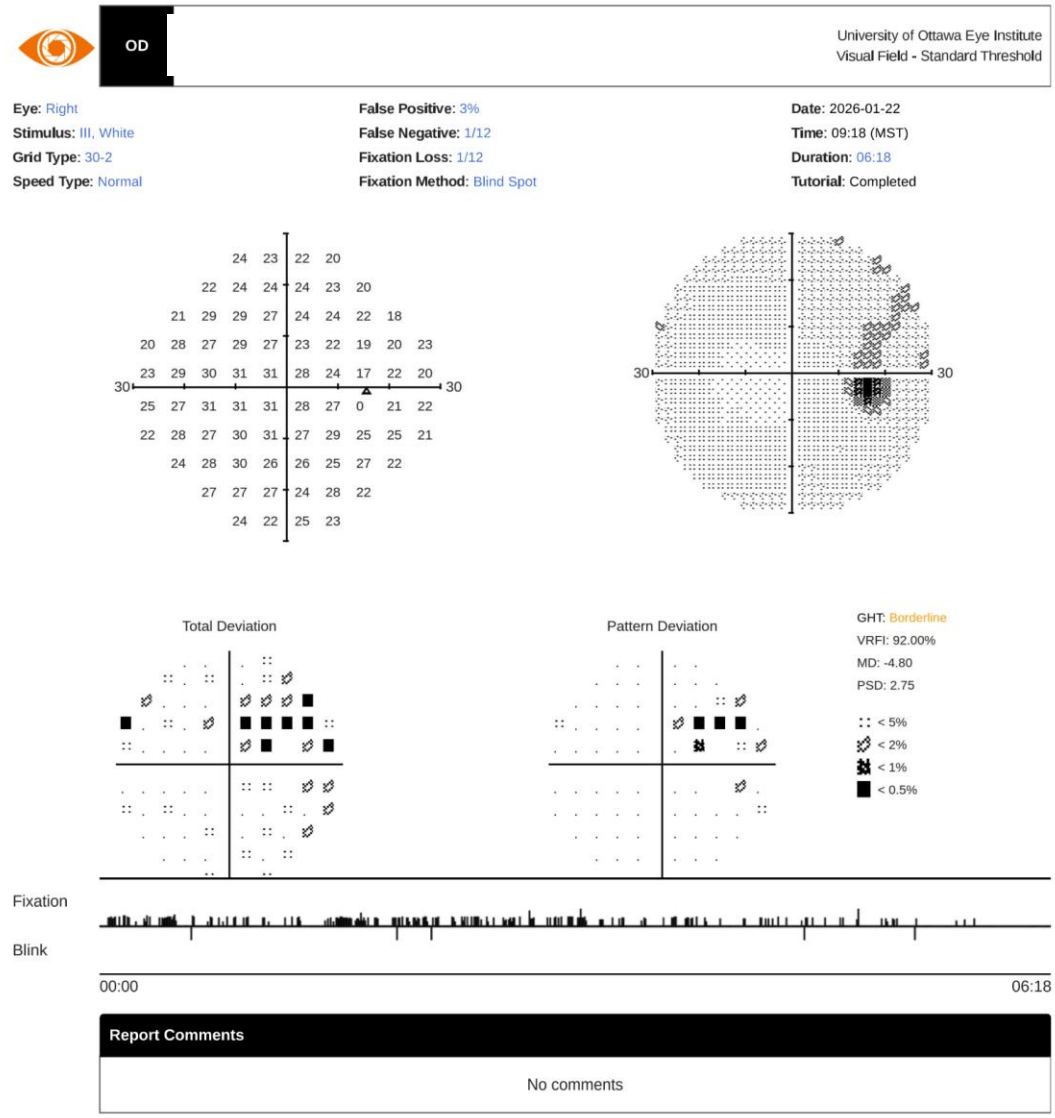
Blind Spot Enlargement Case

In the case of an enlarged blind spot, the delta deviation map showed a mild global negative shift, with most values ranging between approximately −2 dB and −4 dB (Figure 3.11A−C). More pronounced localized deviations were observed at the periphery of the affected region, reaching approximately −9 dB. This pattern reflects generally consistent spatial agreement between platforms, with increased deviation primarily localized to regions corresponding to the enlarged physiological blind spot. Cross-sectional profiling of an enlarged blind spot case showed close agreement between platforms, with both localizing the enlarged blind spot between approximately +10° and +22° eccentricity in the horizontal profile (Figure. 3.11D). However, the RVF profile exhibited broader defect boundaries, with an earlier increase in defect depth beginning at approximately +7° and a more gradual return to baseline beyond +22° compared with the HFA. The vertical profile showed close agreement within the dense core of the enlarged blind spot (−10° to 0°) (Figure. 3.11E). Differences were observed in the superior visual field (0° to +18°), where the HFA demonstrated a partial recovery in defect depth before returning to baseline at approximately +15°, whereas the RVF profile maintained a broader region of elevated defect depth before returning to baseline beyond +18°. The RVF platform displayed localization of the enlarged blind spot, with RVF displaying broader defect margins relative to the HFA.

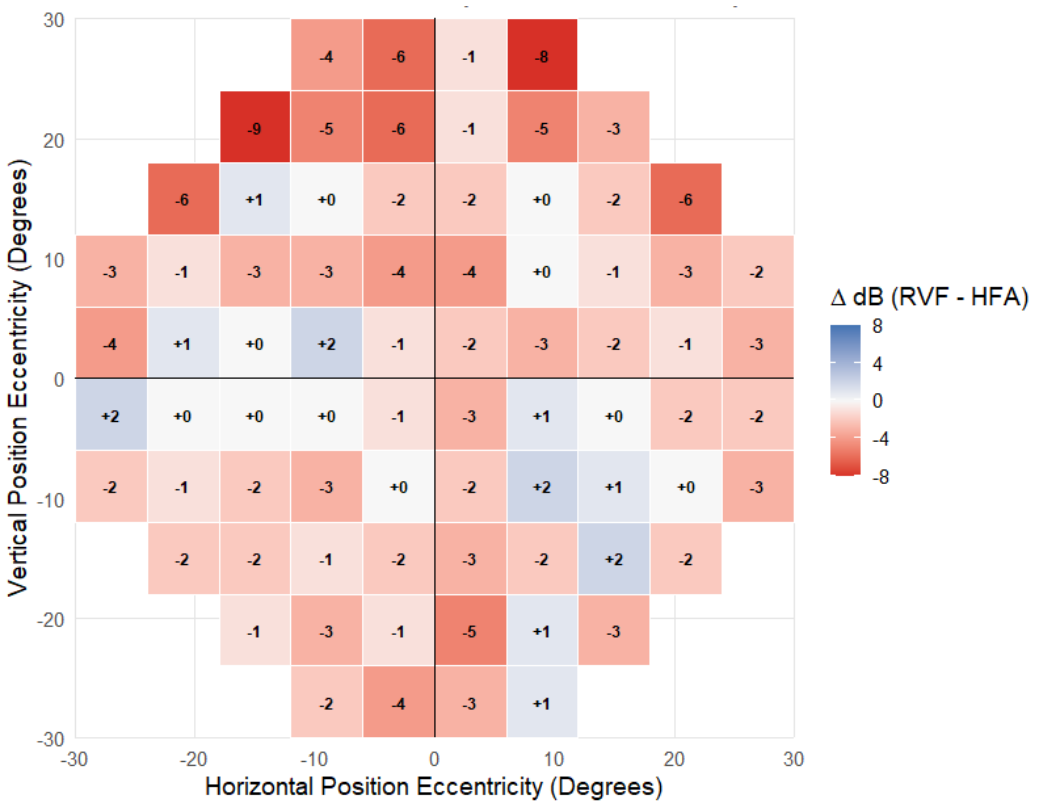
A



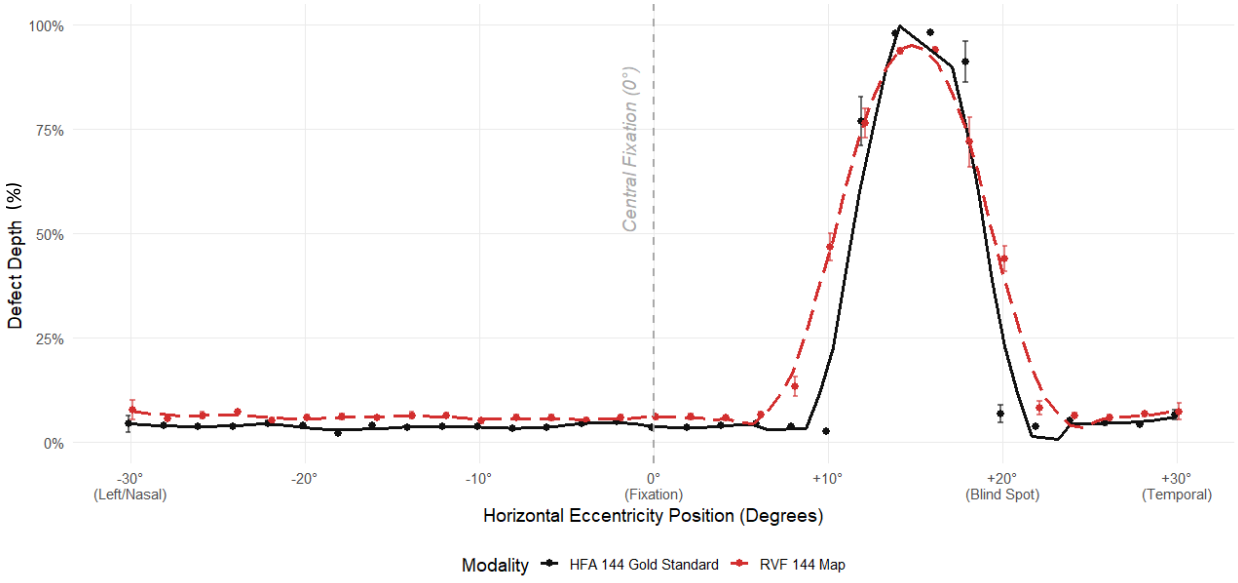
B


OD
University of Ottawa Eye Institute
Visual Field - Standard Threshold
Eye: Right
Stimulus: III, White
Grid Type: 30-2
Speed Type: Normal
False Positive: 3%
False Negative: 1/12
Fixation Loss: 1/12
Fixation Method: Blind Spot
Date: 2026-01-22
Time: 09:18 (MST)
Duration: 06:18
Tutorial: Completed
24 23 22 20
22 24 24 24 23 20
21 29 29 27 24 24 22 18
20 28 27 29 27 23 22 19 20 23
23 29 30 31 31 28 24 17 22 20
30
30
25 27 31 31 31 28 27 0 21 22
22 28 27 30 31 27 29 25 25 21
24 28 30 26 26 25 27 22
27 27 27 24 28 22
24 22 25 23
30
30
Total Deviation
Pattern Deviation
GHT: Borderline
VRFI: 92.00%
MD: -4.80
PSD: 2.75
< 5%
< 2%
< 1%
< 0.5%
Fixation
Blink
00:00
06:18
Report Comments
No comments

C


Vertical Position Eccentricity (Degrees)
Horizontal Position Eccentricity (Degrees)
Δ dB (RVF - HFA)
8
4
0
-4
-8

D


Defect Depth (%)
100%
75%
50%
25%
0%
Central Fixation (0°)
-30°
(Left/Nasal)
-20°
-10°
0°
(Fixation)
+10°
+20°
(Blind Spot)
+30°
(Temporal)
Horizontal Eccentricity Position (Degrees)
Modality
HFA 144 Gold Standard
RVF 144 Map

E

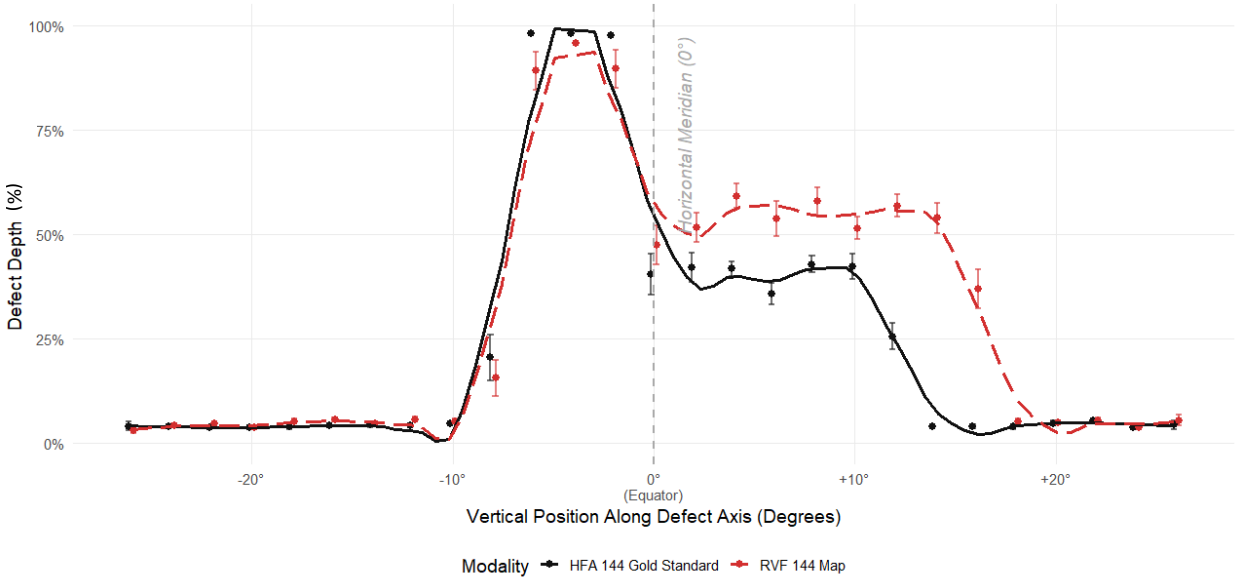
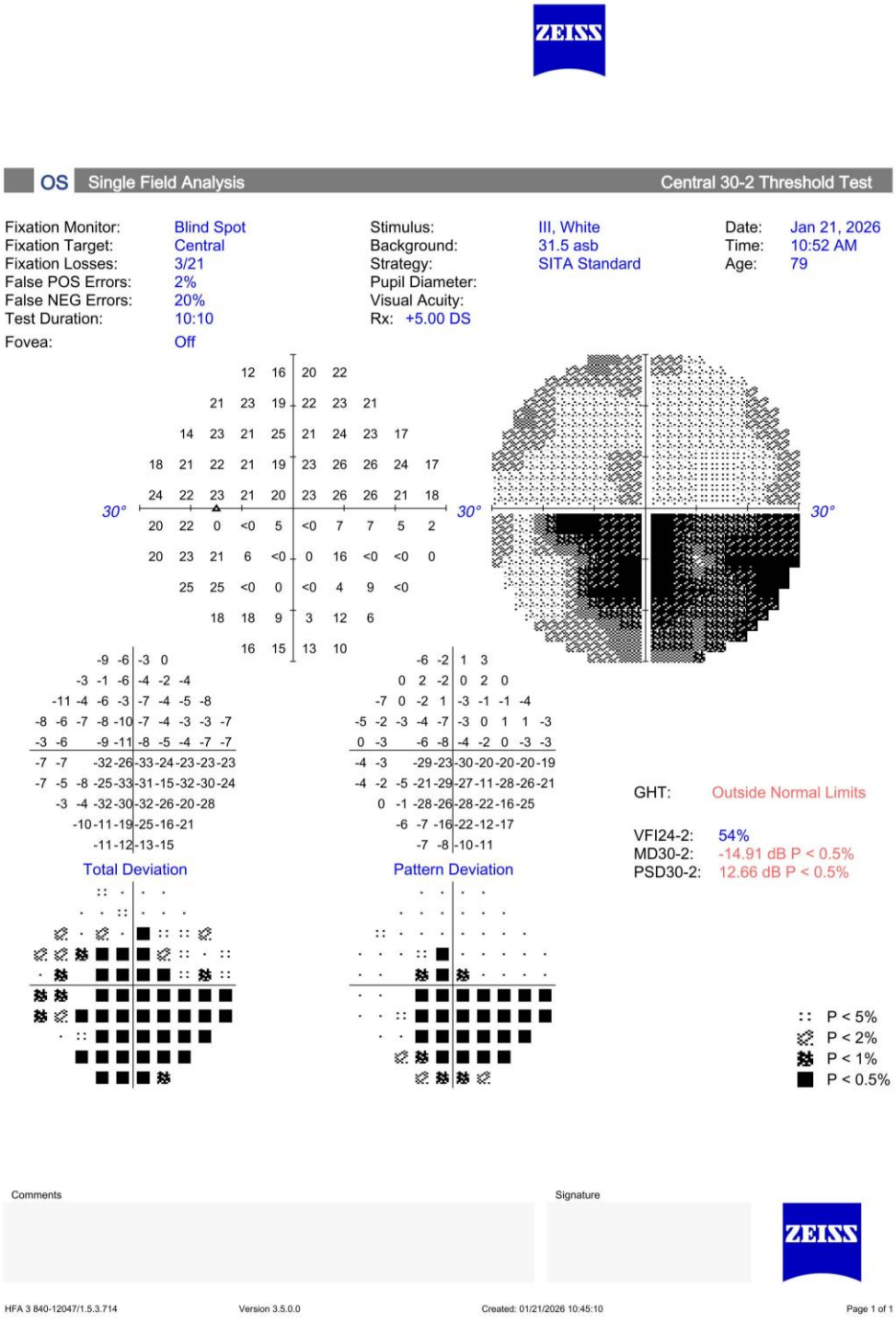


Figure 3.11. Sensitivity mapping comparison for a case of blind spot enlargement in the right eye between RVF and HFA. Perimetric printouts are displayed for the standard-of-care HFA and the RVF platform. (A) Standard HFA diagnostic printout and (B) corresponding RVF diagnostic printout displaying sensitivity (red) and greyscale maps (blue). (C) Spatial delta heatmaps establishing the localized pointwise point-by-point sensitivity differences (Δ dB = RVF − HFA) between the two testing modalities across the tested visual field points. The diverging color scale represents the magnitude and directional variance of the threshold sensitivity deviations between the two devices. Blue signifies positive values indicating threshold overestimation by the VR platform relative to the HFA, and red signifies negative values indicating threshold underestimation. (D) The spatial mapping of visual field defect depth and eccentricity. Defect depth is plotted as a function of horizontal eccentricity position ($y = -4°$ to $-8°$). (E) Corresponding vertical slice analysis showing the distribution of perimetric sensitivity values as a function of visual field eccentricity (x = +14° to +18°). Black points represent HFA measurements and red points represent RVF measurements. Trendlines illustrate the decline in sensitivity from central fixation toward peripheral eccentricities using 2° spatial bins.

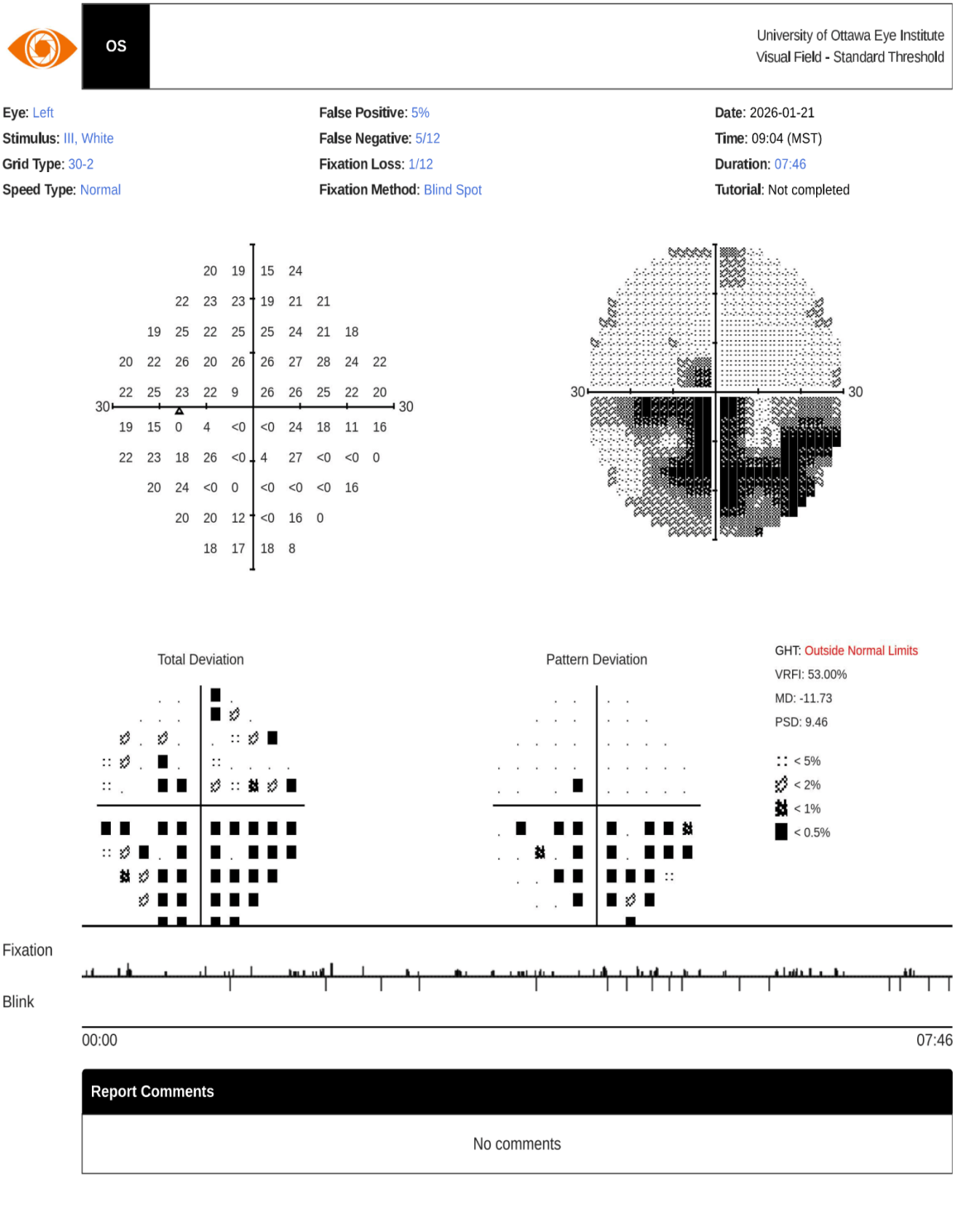
Advanced Glaucoma Case (Inferior Altitudinal Defect)

The patient case with an inferior arcuate defect caused by advanced stage glaucoma demonstrated marked spatial heterogeneity within the inferior-nasal quadrant (Figure 3.12A−C). Pronounced positive deviations were observed in clustered regions (approximately +11 dB to +17 dB), interspersed with negative deviations reaching approximately −11 dB. This indicates reduced pointwise concordance within areas of arcuate field loss, with increased variability along regions of steep spatial transition. Comparing the spatial defect profile of an advanced glaucoma case, the horizontal profile showed close agreement across the temporal visual field, with the RVF and the HFA machines demonstrating a sharp increase in defect depth at approximately −15° (Figure 3.12D). Differences were observed in the nasal visual field (+7.5° to +15.5°), where the RVF profile showed lower defect depths than the HFA. Beyond +17.5°, both machines again demonstrated similar defect depths (>90%). The vertical profile showed strong agreement in the superior visual field and a sharp transition to dense inferior field loss beginning at the horizontal meridian (Figure. 3.12E). The RVF was able to capture this transition; however, it demonstrated a more gradual decline and lower defect depths in the inferior peripheral field.

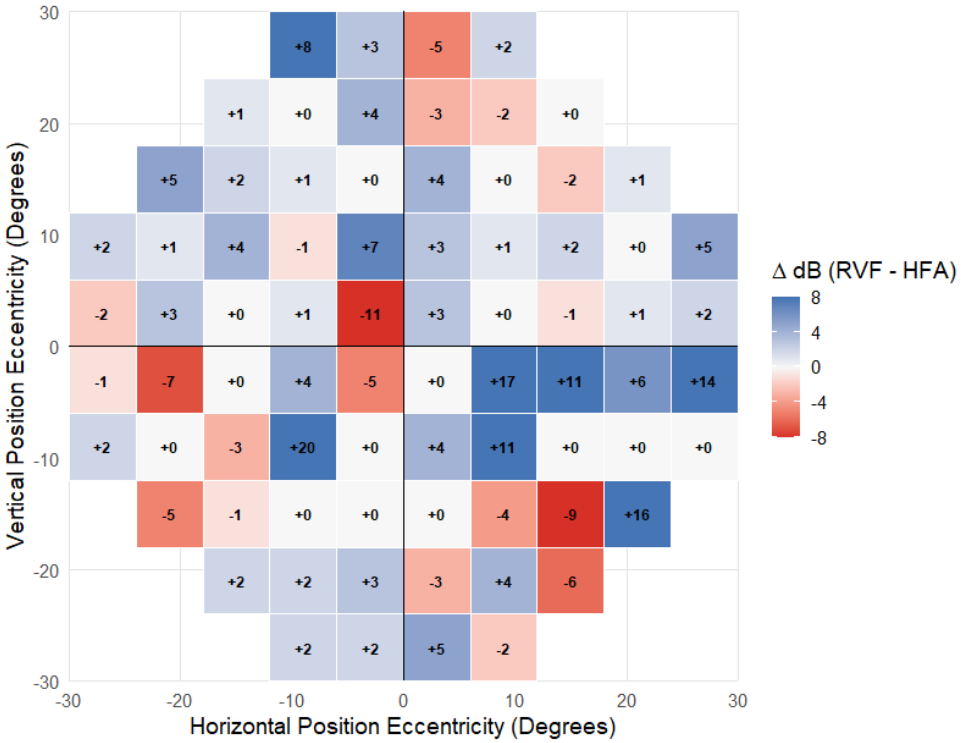
A


OS Single Field Analysis
Central 30-2 Threshold Test
Fixation Monitor: Blind Spot
Fixation Target: Central
Fixation Losses: 3/21
False POS Errors: 2%
False NEG Errors: 20%
Test Duration: 10:10
Fovea: Off
Stimulus: III, White
Background: 31.5 asb
Strategy: SITA Standard
Pupil Diameter:
Visual Acuity:
Rx: +5.00 DS
Date: Jan 21, 2026
Time: 10:52 AM
Age: 79
30°
30°
30°
Total Deviation
Pattern Deviation
GHT: Outside Normal Limits
VFI24-2: 54%
MD30-2: -14.91 dB P < 0.5%
PSD30-2: 12.66 dB P < 0.5%
P < 5%
P < 2%
P < 1%
P < 0.5%
Comments
Signature
© 2019 Carl Zeiss Meditec, Inc. All rights reserved.
HFA 3 840-12047/1.5.3.714
Version 3.5.0.0
Created: 01/21/2026 10:45:10
Page 1 of 1

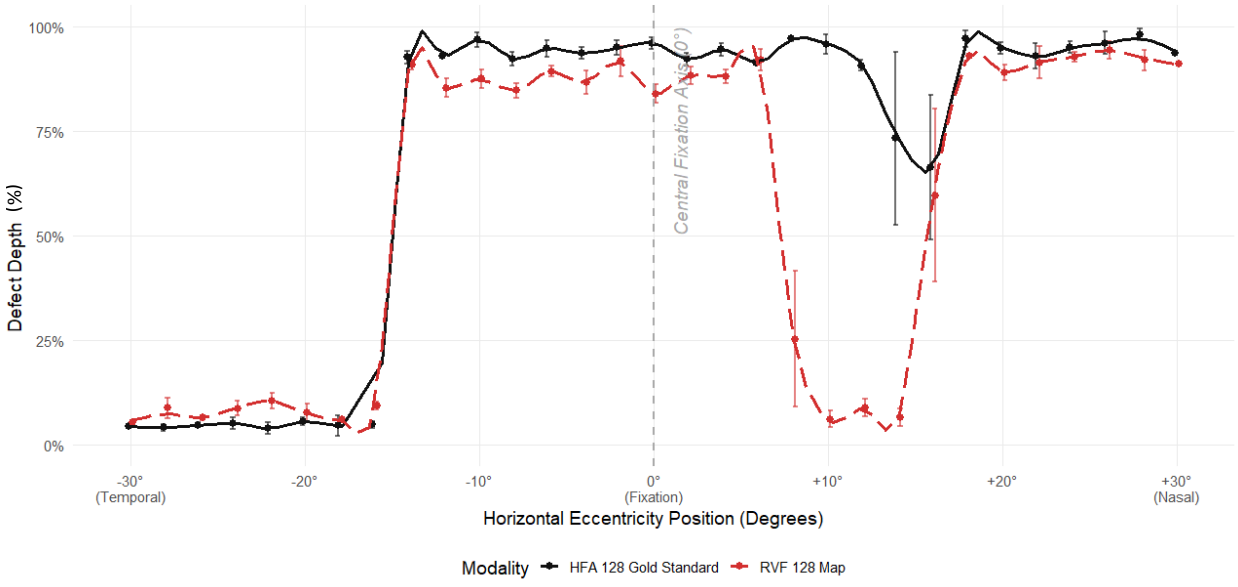
B


OS
University of Ottawa Eye Institute
Visual Field - Standard Threshold
Eye: Left
Stimulus: III, White
Grid Type: 30-2
Speed Type: Normal
False Positive: 5%
False Negative: 5/12
Fixation Loss: 1/12
Fixation Method: Blind Spot
Date: 2026-01-21
Time: 09:04 (MST)
Duration: 07:46
Tutorial: Not completed
20 19 15 24
22 23 23 19 21 21
19 25 22 25 25 24 21 18
20 22 26 20 26 26 27 28 24 22
22 25 23 22 9 26 26 25 22 20
30 30
19 15 0 4 <0 <0 24 18 11 16
22 23 18 26 <0 4 27 <0 <0 0
20 24 <0 0 <0 <0 <0 16
20 20 12 <0 16 0
18 17 18 8
30 30
Total Deviation
Pattern Deviation
GHT: Outside Normal Limits
VRFI: 53.00%
MD: -11.73
PSD: 9.46
:: < 5%
< 2%
< 1%
< 0.5%
Fixation
Blink
00:00
07:46
Report Comments
No comments

C


Vertical Position Eccentricity (Degrees)
Horizontal Position Eccentricity (Degrees)
Δ dB (RVF - HFA)
8
4
0
-4
-8

D


Defect Depth (%)
100%
75%
50%
25%
0%
Central Fixation Axis (0°)
-30°
(Temporal)
-20°
-10°
0°
(Fixation)
+10°
+20°
+30°
(Nasal)
Horizontal Eccentricity Position (Degrees)
Modality
HFA 128 Gold Standard
RVF 128 Map

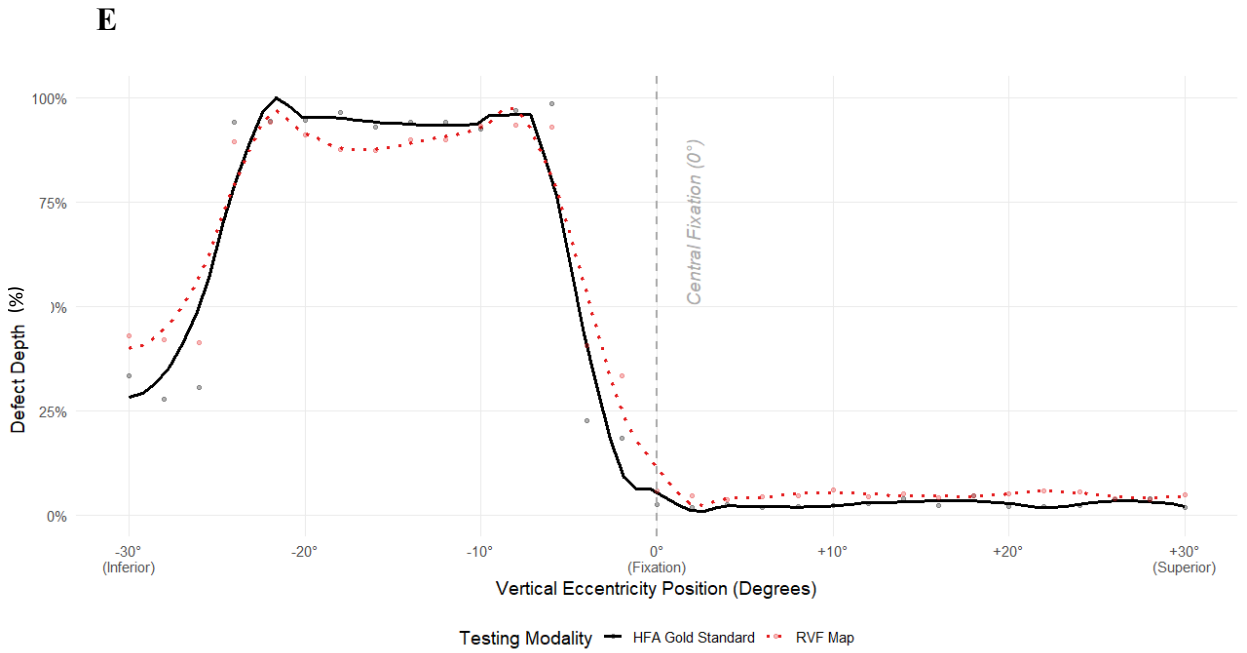


Figure 3.12. Sensitivity mapping comparison for a patient with inferior altitudinal defect secondary to advanced-stage glaucoma between RVF and HFA. Perimetric outputs are displayed for the standard-of-care HFA and the RVF platform. (A) Standard HFA diagnostic printout and (B) corresponding RVF diagnostic printout displaying sensitivity (red) and greyscale maps (blue). (C) Spatial delta heatmaps establishing the localized pointwise point-by-point sensitivity differences (Δ dB = RVF − HFA) between the two testing modalities across the tested visual field points. The diverging color scale represents the magnitude and directional variance of the threshold sensitivity deviations between the two devices. Blue signifies positive values indicating threshold overestimation by the VR platform relative to the HFA, and red signifies negative values indicating threshold underestimation. (D) The spatial mapping of visual field defect depth and eccentricity. Defect severity is plotted as a function of horizontal eccentricity position (y = −8° to −12°). (E) Corresponding vertical slice analysis showing the distribution of perimetric sensitivity values as a function of visual field eccentricity (x = −2° to +2°). Black points represent HFA measurements and red points represent RVF measurements. Trendlines illustrate the decline in sensitivity from central fixation toward peripheral eccentricities.

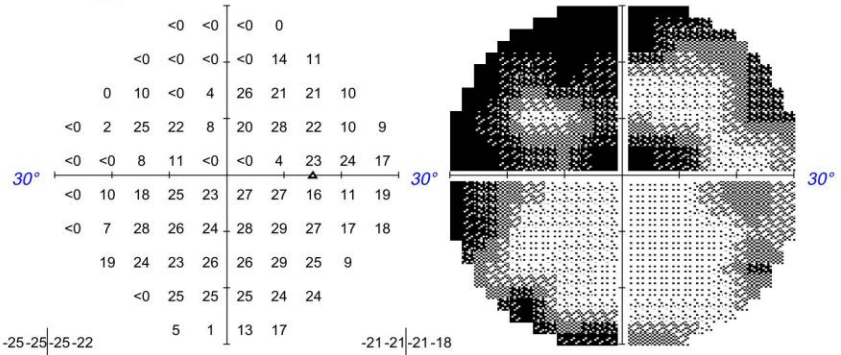
NAION Case (Superior and Central Defects)

In the case with superior and central defects secondary to NAION, the deviation map showed large positive deviations across the superior hemifield and central region, reaching approximately +20 dB to +27 dB (Figure 3.13A−C). These findings were accompanied by localized negative deviations in adjacent paracentral regions (approximately −20 dB to −22 dB). Overall, these results indicate reduced spatial agreement between RVF and HFA in regions of complex central and superior field loss, with marked variability at defect boundaries. Spatial analysis of a case of NAION with superior and central field defect patterns revealed substantial differences in the horizontal profile (y = 0°) between the RVF and the HFA (Figure. 3.13D). Across the nasal visual field (−30° to −10°), the HFA showed a consistently greater defect depth than the RVF. Within the central field (−5° to +5°), the RVF demonstrated a marked increase in defect depth compared with HFA, followed by a rapid reduction at approximately +10° eccentricity. The vertical profile (x = 0°) showed close agreement within the inferior visual field and at the superior altitudinal transition (Figure. 3.13E). However, throughout the superior visual field, the HFA demonstrated persistently greater defect depths, whereas the RVF showed lower defect depths with a gradual reduction toward the superior periphery. This resulted in reduced superimposition of regression trendlines, suggesting decreased concordance in regions characterized by complex or steeply varying relative sensitivity loss. In this case, the RVF demonstrated reduced agreement compared to the HFA, with the largest differences observed in the central, nasal, and superior regions of the visual field.

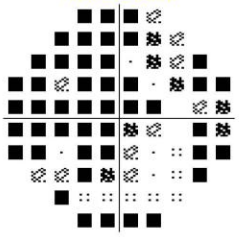
A

OD Single Field Analysis — Central 30-2 Threshold Test

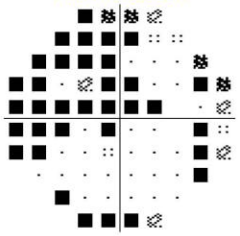
Fixation Monitor:	Blind Spot	Stimulus:	III, White	Date:	Jan 27, 2026
Fixation Target:	Central	Background:	31.5 asb	Time:	10:38 AM
Fixation Losses:	5/22 XX	Strategy:	SITA Standard	Age:	70
False POS Errors:	10%	Pupil Diameter:			
False NEG Errors:	20%	Visual Acuity:			
Test Duration:	13:28	Rx:	+5.00 DS		
Fovea:	Off				



Total Deviation



Pattern Deviation



GHT: Outside Normal Limits

VFI24-2: 60%
MD30-2: -13.18 dB P < 0.5%
PSD30-2: 11.70 dB P < 0.5%

*** Low Test Reliability ***

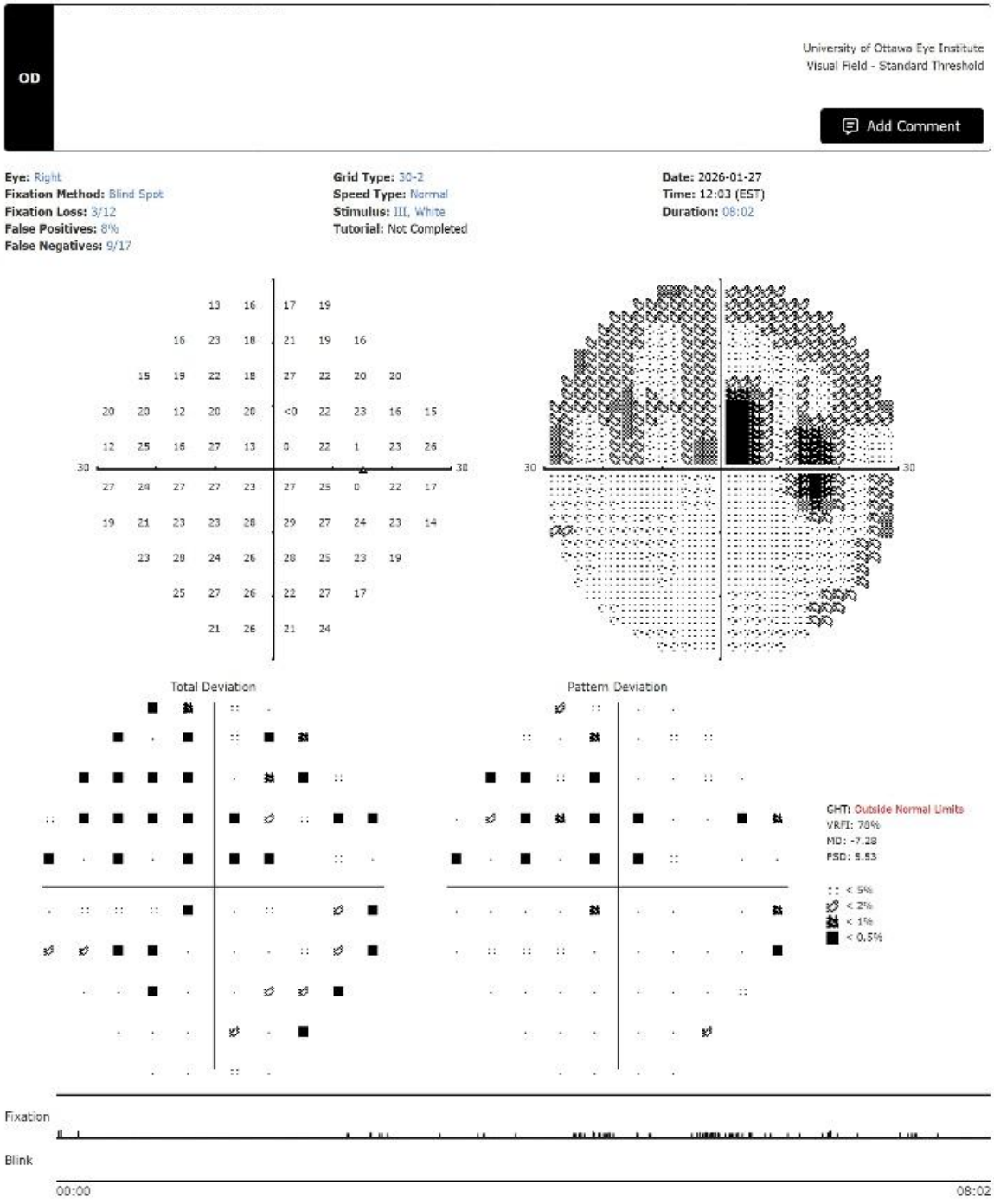
:: P < 5%
P < 2%
P < 1%
P < 0.5%

Comments

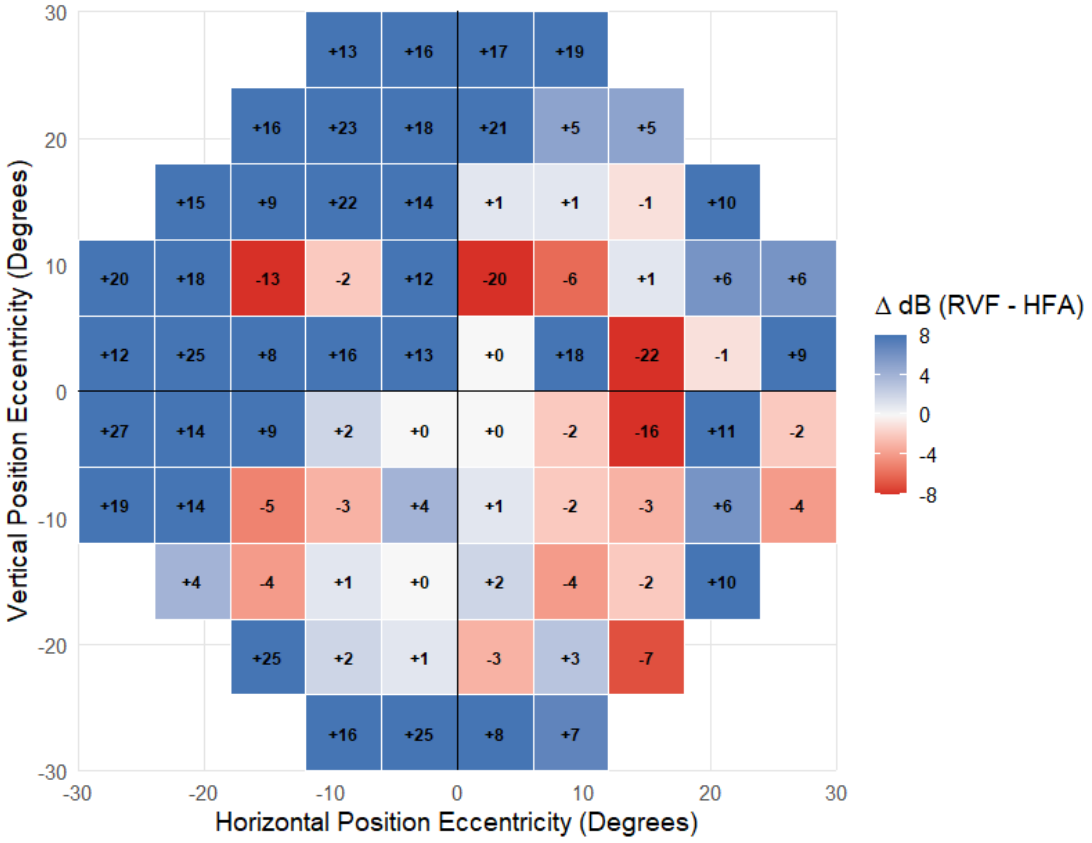
Signature

HFA 3 840-12047/1.5.3.714 Version 3.5.0.0 Created: 01/27/2026 10:48:36 Page 1 of 1

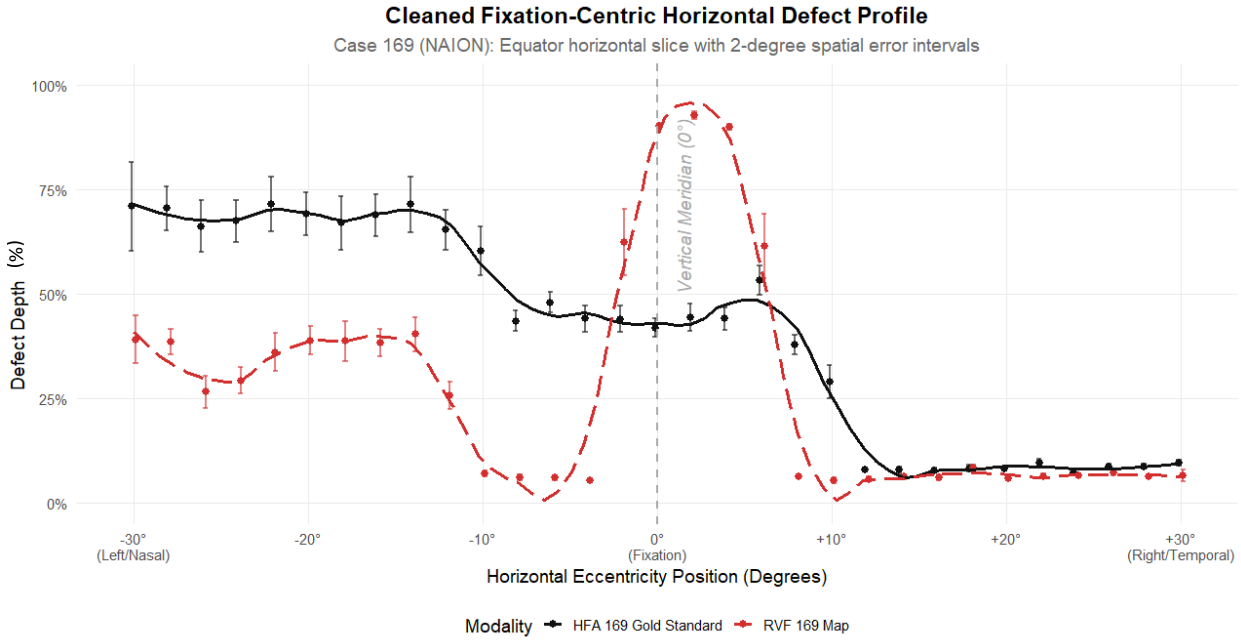
B


OD
University of Ottawa Eye Institute
Visual Field - Standard Threshold
Add Comment
Eye: Right
Fixation Method: Blind Spot
Fixation Loss: 3/12
False Positives: 8%
False Negatives: 9/17
Grid Type: 30-2
Speed Type: Normal
Stimulus: III, White
Tutorial: Not Completed
Date: 2026-01-27
Time: 12:03 (EST)
Duration: 08:02
Total Deviation
Pattern Deviation
GHT: Outside Normal Limits
VRFI: 70%
MD: -7.28
PSD: 5.53
< 5%
< 2%
< 1%
< 0.5%
Fixation
Blink
00:00
08:02

C


Vertical Position Eccentricity (Degrees)
Horizontal Position Eccentricity (Degrees)
Δ dB (RVF - HFA)
8
4
0
-4
-8
30
20
10
0
-10
-20
-30
+13 +16 +17 +19
+16 +23 +18 +21 +5 +5
+15 +9 +22 +14 +1 +1 -1 +10
+20 +18 -13 -2 +12 -20 -6 +1 +6 +6
+12 +25 +8 +16 +13 +0 +18 -22 -1 +9
+27 +14 +9 +2 +0 +0 -2 -16 +11 -2
+19 +14 -5 -3 +4 +1 -2 -3 +6 -4
+4 -4 +1 +0 +2 -4 -2 +10
+25 +2 +1 -3 +3 -7
+16 +25 +8 +7

D


Cleaned Fixation-Centric Horizontal Defect Profile
Case 169 (NAION): Equator horizontal slice with 2-degree spatial error intervals
Defect Depth (%)
100%
75%
50%
25%
0%
Vertical Meridian (0°)
-30°
(Left/Nasal)
-20°
-10°
0°
(Fixation)
+10°
+20°
+30°
(Right/Temporal)
Horizontal Eccentricity Position (Degrees)
Modality
HFA 169 Gold Standard
RVF 169 Map

E

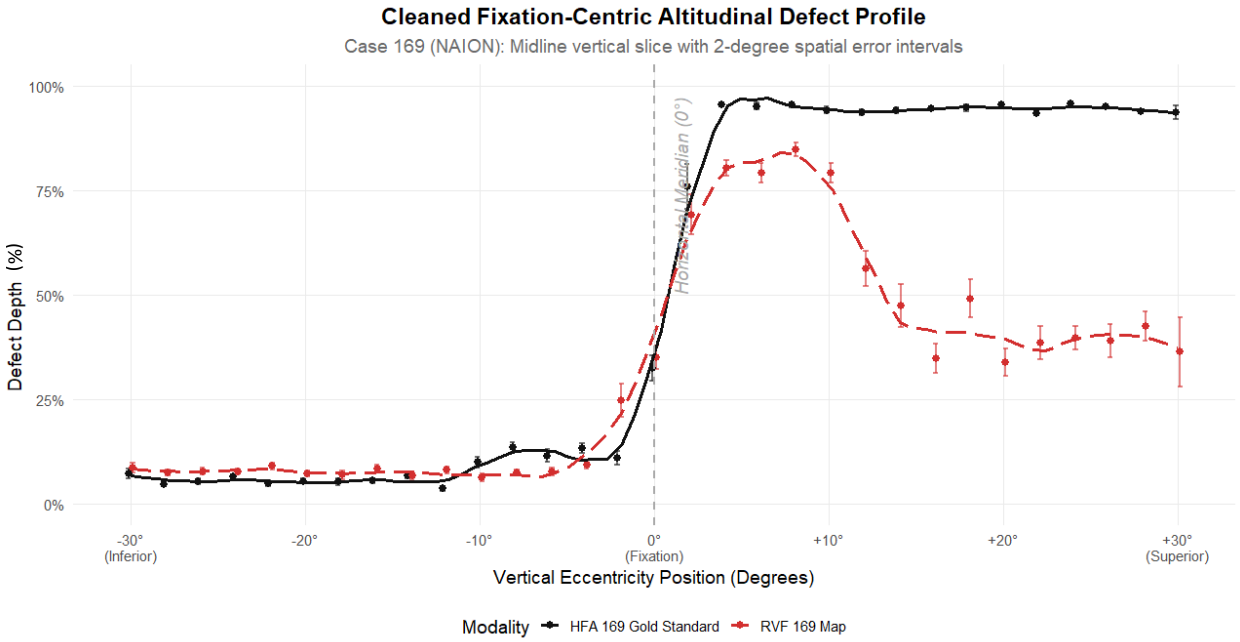


Figure 3.13. Sensitivity mapping comparison for a patient with superior and central defects secondary to NAION between RVF and HFA. Perimetric outputs are displayed for the standard-of-care HFA and the RVF platform. (A) Standard HFA diagnostic printout and (B) corresponding RVF diagnostic printout displaying sensitivity (red) and greyscale maps (blue). (C) Spatial delta heatmaps establishing the localized pointwise point-by-point sensitivity differences (Δ dB = RVF − HFA) between the two testing modalities across the tested visual field points. The diverging color scale represents the magnitude and directional variance of the threshold sensitivity deviations between the two devices. Blue signifies positive values indicating threshold overestimation by the VR platform relative to the HFA, and red signifies negative values indicating threshold underestimation. (D) The spatial mapping of visual field defect depth and eccentricity. Defect severity is plotted as a function of horizontal eccentricity position (y = +6° to +10°). (E) Corresponding vertical slice analysis showing the distribution of sensitivity values as a function of vertical field eccentricity (x = +3° to +7°). Black points represent HFA measurements and red points represent RVF measurements. Trendlines illustrate the decline in sensitivity from central fixation toward peripheral eccentricities using 2° spatial bins.

Advanced Glaucoma Case (Severe Peripheral Constriction)

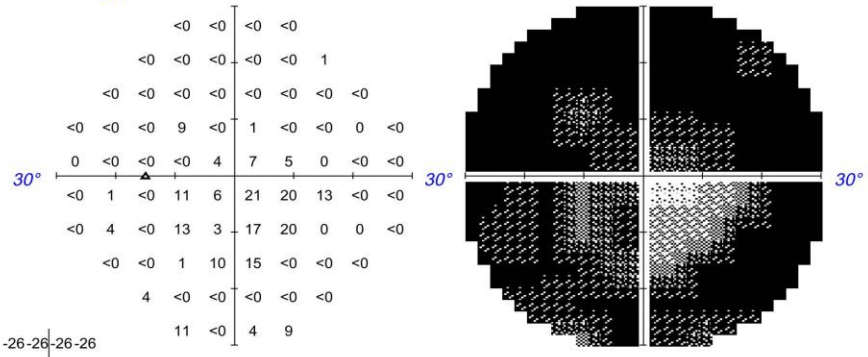
In the case of severe field constriction due to advanced glaucoma, peripheral regions demonstrate near-zero deviation (approximately 0 dB), indicating strong agreement between platforms in areas of absolute field loss (Figure 3.14A−C). In contrast, the central and paracentral regions showed predominantly negative deviations ranging from approximately −5 dB to −13 dB, interspersed with isolated positive outliers reaching up to +23 dB. This suggests high concordance in severely affected peripheral regions, alongside increased pointwise differences within the residual central island of vision. When comparing the spatial representation of a severe glaucomatous field constriction case, the RVF demonstrated close agreement to the HFA in the overall pattern of visual field loss, identifying dense peripheral defects (>90% defect depth) surrounding a residual central island of vision across both the vertical (x = 0°) and horizontal (y = 0°) profiles (Figures. 3.14D, E). Differences were observed within the transition between the central preserved field and the peripheral scotoma. The HFA demonstrated a steeper transition to a minimum defect depth of approximately 15% at fixation, whereas the RVF profile showed a more gradual transition and a higher minimum defect depth of approximately 23%. The RVF showed good agreement compared to HFA in localizing the residual central island and peripheral scotoma, although the RVF demonstrated a broader transition zone and reduced central preservation compared with the HFA.

Overall, the concordance between VR perimetry platforms and HFA was strongly dependent on defect structure and spatial complexity. Higher levels of agreement were observed in normal visual fields and in uniform absolute scotomas, including physiological blind spots and advanced constriction. However, reduced agreement was evident in neuro-ophthalmic defects characterized by steep spatial gradients and heterogeneous relative scotomas, where VR-derived profiles showed attenuated representation of abrupt sensitivity transitions compared with HFA.

A

OS Single Field Analysis — Central 30-2 Threshold Test

Fixation Monitor:	Blind Spot	Stimulus:	III, White	Date:	Jan 21, 2026
Fixation Target:	Central	Background:	31.5 asb	Time:	2:38 PM
Fixation Losses:	1/16	Strategy:	SITA Standard	Age:	55
False POS Errors:	13%	Pupil Diameter:			
False NEG Errors:	0%	Visual Acuity:			
Test Duration:	10:04	Rx:	+1.50 DS +1.50 DC X 38		
Fovea:	Off				



			<0	<0	<0	<0			
		<0	<0	<0	<0	<0	1		
	<0	<0	<0	<0	<0	<0	<0	<0	
<0	<0	<0	9	<0	1	<0	<0	0	<0
0	<0	<0	<0	4	7	5	0	<0	<0
<0	1	<0	11	6	21	20	13	<0	<0
<0	4	<0	13	3	17	20	0	0	<0
	<0	<0	1	10	15	<0	<0	<0	
		4	<0	<0	<0	<0	<0		
			11	<0	4	9			

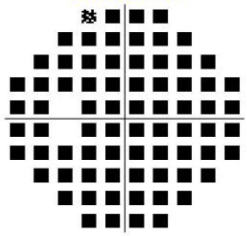
MD Threshold exceeded.
See Total Deviation plot.
Consider 10-2 testing.

GHT: Outside Normal Limits

VFI24-2: 16%
MD30-2: -26.87 dB P < 0.5%
PSD30-2: 7.40 dB P < 0.5%

			-26	-26	-26	-26			
		-28	-29	-29	-29	-29	-26		
	-29	-30	-31	-31	-32	-31	-31	-29	
-30	-31	-32	-21	-33	-31	-33	-32	-29	-28
-29	-32		-34	-28	-25	-27	-31	-31	-29
-31	-29		-21	-26	-12	-13	-18	-32	-29
-31	-26	-33	-18	-29	-15	-12	-31	-29	-28
	-32	-32	-30	-21	-16	-33	-32	-30	
		-25	-32	-32	-32	-31	-30		
			-17	-30	-23	-18			

Total Deviation



Pattern Deviation

MD Threshold exceeded.
See Total Deviation plot.
Consider 10-2 testing.

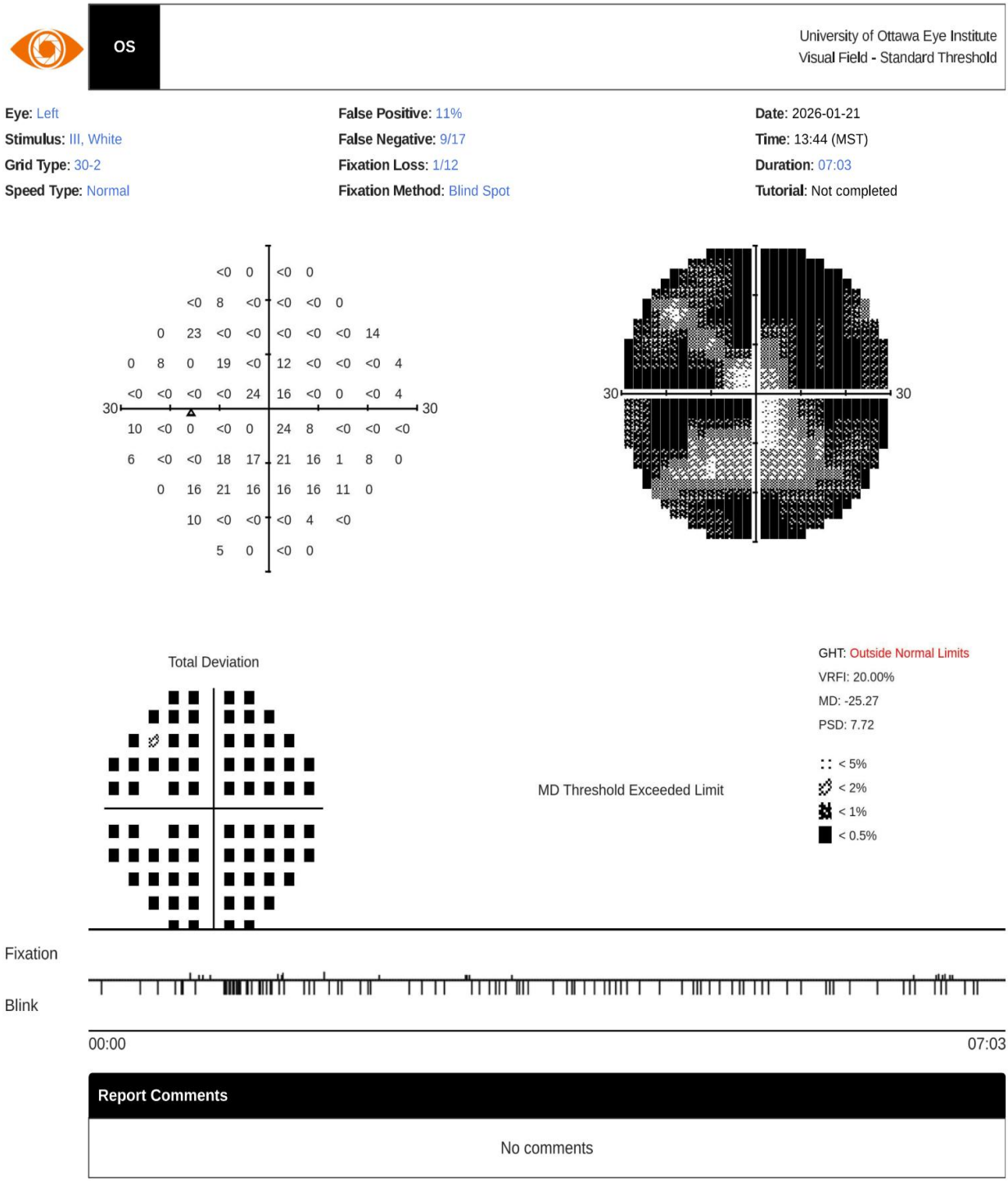
:: P < 5%
P < 2%
P < 1%
P < 0.5%

Comments

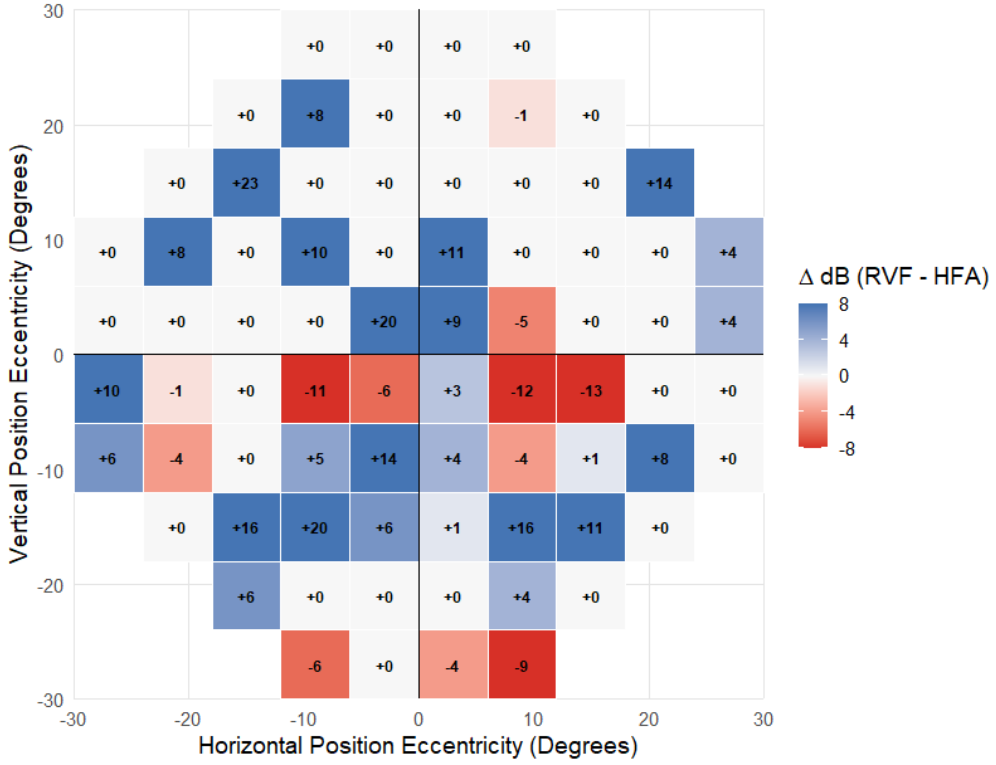
Signature

HFA 3 850-16843/1.6.0.694 Version 3.5.0.0 Created: 01/21/2026 14:51:44 Page 1 of 1

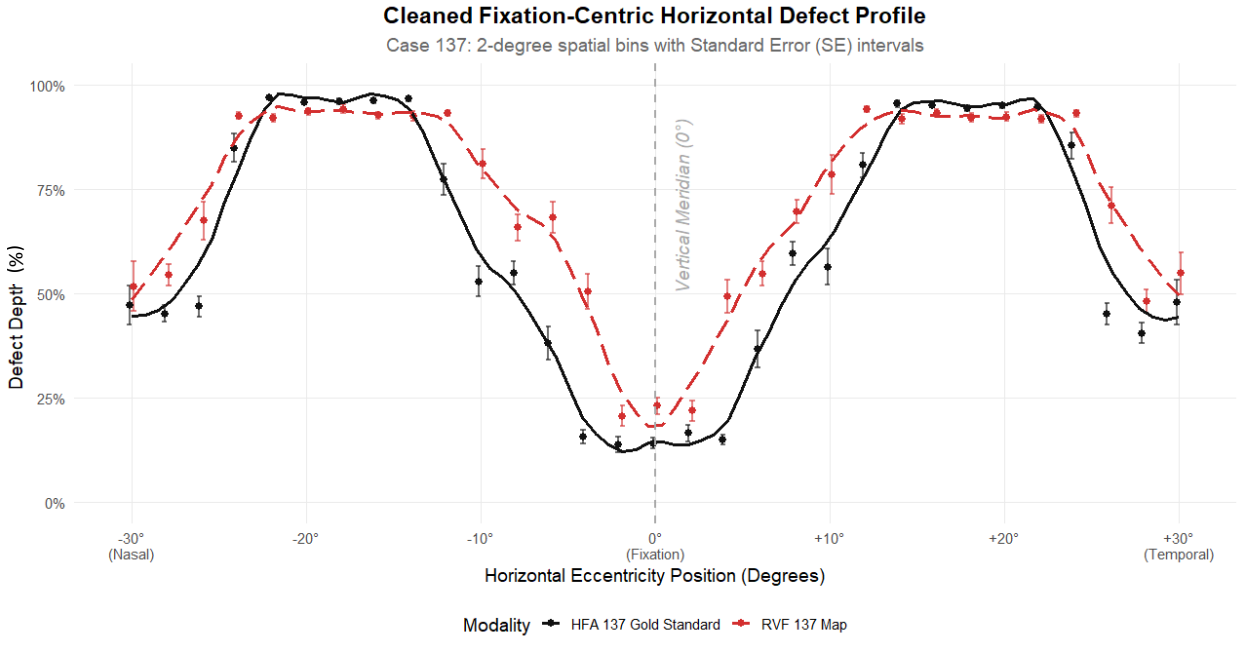
B


OS
University of Ottawa Eye Institute
Visual Field - Standard Threshold
Eye: Left
Stimulus: III, White
Grid Type: 30-2
Speed Type: Normal
False Positive: 11%
False Negative: 9/17
Fixation Loss: 1/12
Fixation Method: Blind Spot
Date: 2026-01-21
Time: 13:44 (MST)
Duration: 07:03
Tutorial: Not completed
30
30
Total Deviation
MD Threshold Exceeded Limit
GHT: Outside Normal Limits
VRFI: 20.00%
MD: -25.27
PSD: 7.72
< 5%
< 2%
< 1%
< 0.5%
Fixation
Blink
00:00
07:03
Report Comments
No comments

C


Vertical Position Eccentricity (Degrees)
Horizontal Position Eccentricity (Degrees)
Δ dB (RVF - HFA)
8
4
0
-4
-8

D


Cleaned Fixation-Centric Horizontal Defect Profile
Case 137: 2-degree spatial bins with Standard Error (SE) intervals
Defect Depth (%)
Vertical Meridian (0°)
Horizontal Eccentricity Position (Degrees)
(Nasal)
(Fixation)
(Temporal)
Modality
HFA 137 Gold Standard
RVF 137 Map

E

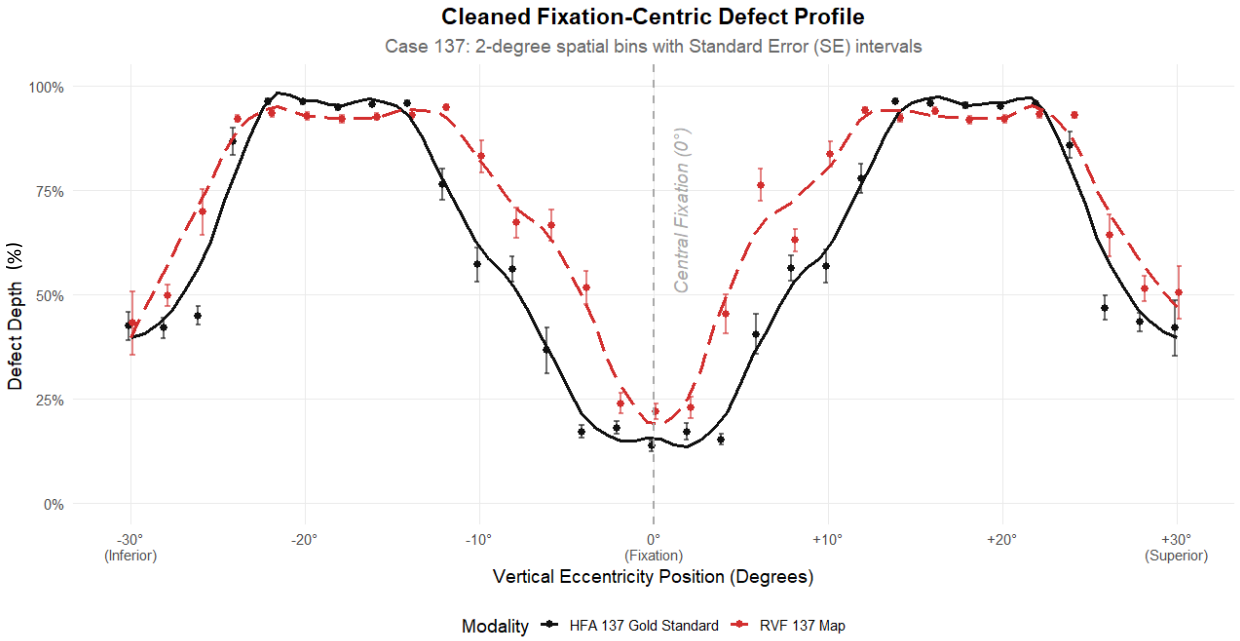


Figure 3.14. Sensitivity mapping comparison for a patient with severe constriction of field secondary to advanced-stage glaucoma between RVF and HFA. Perimetric outputs are displayed for the standard-of-care HFA and the RVF platform. (A) Standard HFA diagnostic printout and (B) corresponding RVF diagnostic printout displaying sensitivity (red) and greyscale maps (blue). (C) Spatial delta heatmaps establishing the localized pointwise point-by-point sensitivity differences (Δ dB = RVF − HFA) between the two testing modalities across the tested visual field points. The diverging color scale represents the magnitude and directional variance of the threshold sensitivity deviations between the two devices. Blue signifies positive values indicating threshold overestimation by the VR platform relative to the HFA, and red signifies negative values indicating threshold underestimation. (D) The spatial mapping of visual field defect depth and eccentricity using 2° spatial bins. Defect severity is plotted as a function of horizontal eccentricity position (y = −7° to −3 °). (E) Corresponding vertical slice analysis showing the distribution of perimetric sensitivity values as a function of visual field eccentricity (y = +3° to +7°). Black points represent HFA measurements and red points represent RVF measurements. Trendlines illustrate the decline in sensitivity from central fixation toward peripheral eccentricities using 2° spatial bins.

3.7. Pupil Data Subset – Baseline Characteristics of the Study Population

A total of 186 participants were prospectively enrolled in the study. Following the exclusion of 45 participants due to unrecoverable data, 141 participants (282 eyes) were included in the final analysis. The entire study cohort comprised 68 females (65.4%) and 36 males (34.6%), with a mean age of 52.3 ± 18.4 years (range, 18–89 years) (Table 3.3). The study cohort included both healthy participants and patients with a broad range of ocular and neuro-ophthalmic conditions affecting the anterior visual pathway, including optic neuropathy, optic neuritis, optic disc edema, meningioma, pituitary adenoma, and other compressive, inflammatory, and ischemic disorders.

Analysis of baseline study population characteristics demonstrated no statistically significant differences between the healthy control group (n = 20) and the patient group (n = 121) with respect to age (44.2 ± 12.3 vs. 50.1 ± 11.2 years; p = 0.056) or sex distribution (70.0% vs. 65.3% female; p = 0.820). Healthy control eyes (n = 40) had a mean best-corrected visual acuity (BCVA) of 0.017 ± 0.04 logMAR, whereas patient eyes (n = 242) had a mean BCVA of 0.20 ± 0.06 logMAR. Inter-eye BCVA did not differ significantly within the healthy control group (p = 0.77); however, a significant inter-eye difference was observed in the patient group (p < 0.001).

Table 3.3. Baseline demographics and clinical characteristics of the study population

Characteristic	Healthy Control Group (n = 20, 40 eyes)	Patient Group (n = 121, 242 eyes)	*p*-value
Age (years; mean ± SD)	44.2 ± 12.3	50.1 ± 11.2	0.056
Sex (female, %)	70	65	0.82[1]
BCVA (logMAR; mean ± SD)	0.017 ± 0.04	0.20 ± 0.06	<0.001***
Right Eye BCVA (logMAR; mean ± SD)	0.015 ± 0.03	0.25 ± 0.07	
Left Eye VA (logMAR; mean ± SD)	0.019 ± 0.03	0.16 ± 0.04	
***p*-value**	0.77	<0.001***	

Independent two-tailed t-test for all p-values; [1]Chi-square test; Statistically significant: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

3.8. Comparison of Baseline Pupillary Dynamics

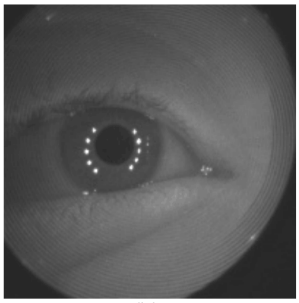
A comparative evaluation of baseline pupillary dynamics was conducted to assess agreement between the Pupil X VR platform and standard manual examination. Sample printouts from normal and affected participants obtained using the Pupil X system are shown in Figures 3.15 and 3.16. The mean testing time for the Pupil X protocol was approximately 30 seconds per participant, compared with approximately 4–5 minutes for the standard manual pupillary examination. Comparative analysis across the full sample (n = 282 eyes) demonstrated a statistically significant difference between measurements obtained by the Pupil X VR platform and manual clinical examination ($p < 0.001$) (Table 3.4). For minimum pupil diameter, the automated platform recorded smaller values than manual assessment (3.06 ± 0.96 mm vs. 3.75 ± 1.13 mm). Similarly, maximum pupil diameter measurements were lower with the automated platform (4.12 ± 1.04 mm vs. 5.06 ± 1.25 mm). Standard deviations were comparable between modalities across both measures.

A.

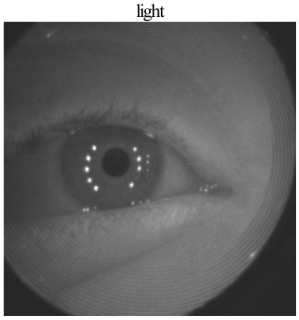
PUPIL

Age	47 year(s)	**Source**	omlabs
Name	Rp0303328	**Protocol**	4, 125 level(s)
Gender	F	**Cycles**	3, 1000/3000 ms
Mrn	RP0330328	**Hospital**	L V Prasad Eye Institute
		Date	30 Mar 2026
		Time	10:23 PM

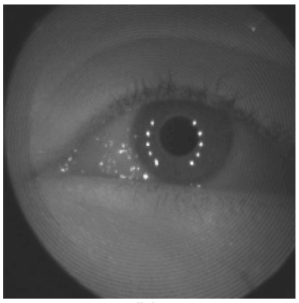
dark



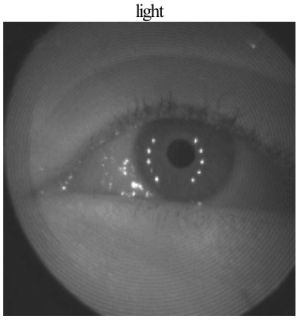
light



dark



light



- **Eye:** Right eye (OD)
- **Automated decision support:**
  1. **Pre-light stimulation -** 4.14 mm
  2. **Post light stimulation -** 2.99 mm
  3. **Direct reflex -** 1.14 mm
  4. **Consensual reflex -** 1.08 mm
  5. **Diagnosis -** Normal
  6. **Rapd grade -** None
  7. **Percentage change -** 27.78%

- **Eye:** Left eye (OS)
- **Automated decision support:**
  1. **Pre-light stimulation -** 4.13 mm
  2. **Post light stimulation -** 2.96 mm
  3. **Direct reflex -** 1.16 mm
  4. **Consensual reflex -** 1.11 mm
  5. **Diagnosis -** Normal
  6. **Rapd grade -** None
  7. **Percentage change -** 28.33%

Right eye  Left eye

B.

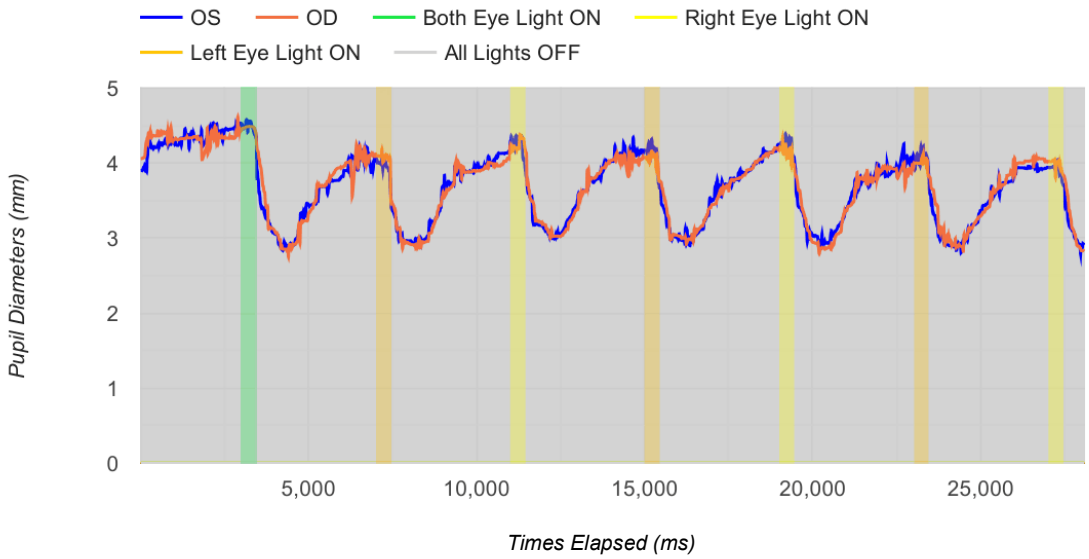


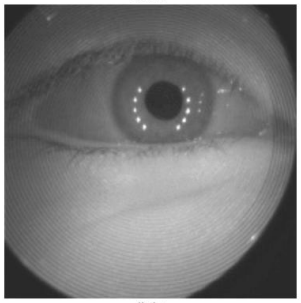
Figure 3.15. Sample printout of Pupil X from a normal participant. (A) Infrared pupillometry tracking and symmetrical waveform analysis in a healthy control participant. Static infrared images acquired during the standardized Pupil X testing protocol, demonstrating automated pupil segmentation under baseline dark adaptation and light stimulation. (B) Synchronized pupillometry waveforms of the right eye (OD, orange) and left eye (OS, blue) recorded over a 25,000 ms trial. Grey shaded bands denote periods of dark intervals. Pupillary constriction and recovery curves remained closely aligned during binocular (green), left-eye (orange), and right-eye (yellow) stimulation intervals. Measured direct constriction amplitudes were 1.14 mm for the right eye and 1.16 mm for the left eye, with corresponding percentage constrictions of 27.78% and 28.33%, respectively. These symmetric pupillary responses were classified by the Pupil X algorithm as no relative afferent pupillary defect (RAPD grade: None).

C.

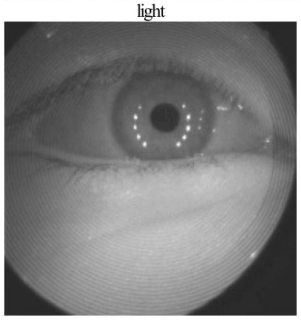
PUPIL

Age	36 year(s)	**Source**	omlabs
Name	RP0323311	**Protocol**	4, 125 level(s)
Gender	F	**Cycles**	3, 1000/3000 ms
Mrn	RP0323311	**Hospital**	L V Prasad Eye Institute
		Date	24 Mar 2026
		Time	1:12 AM

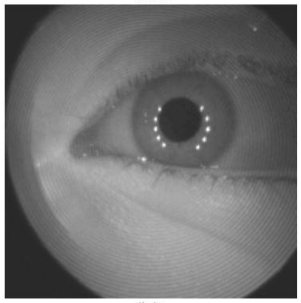
dark



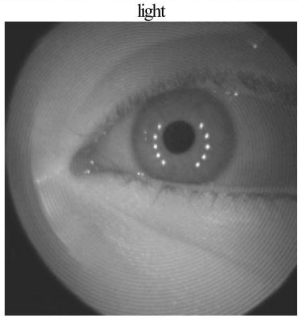
light



dark



light



- **Eye:** Right eye (OD)
- **Automated decision support:**
  1. **Pre-light stimulation -** 4.39 mm
  2. **Post light stimulation -** 3.23 mm
  3. **Direct reflex -** 1.16 mm
  4. **Consensual reflex -** 0.81 mm
  5. **Diagnosis -** Normal
  6. **Rapd grade -** None
  7. **Percentage change -** 26.42%

- **Eye:** Left eye (OS)
- **Automated decision support:**
  1. **Pre-light stimulation -** 3.74 mm
  2. **Post light stimulation -** 3.03 mm
  3. **Direct reflex -** 0.71 mm
  4. **Consensual reflex -** 1.01 mm
  5. **Diagnosis -** RAPD present
  6. **Rapd grade -** 3
  7. **Percentage change -** 18.98%

Right eye  Left eye

D.

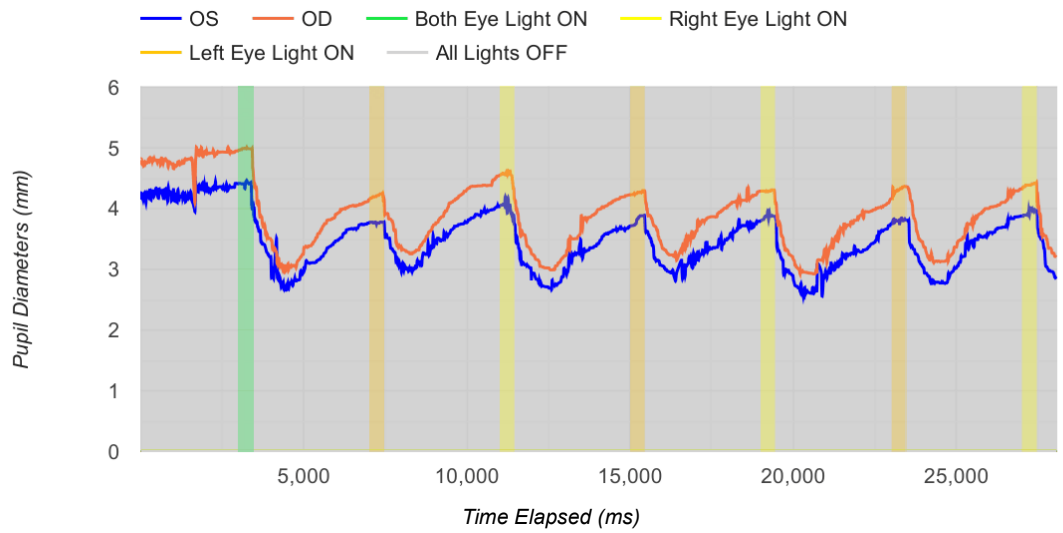


Figure 3.16. Sample printout of Pupil X for a RAPD positive patient. (A) Infrared pupillometry tracking and asymmetrical waveform analysis of a confirmed relative afferent pupillary defect (RAPD). Static infrared images acquired during the Pupil X testing protocol demonstrating real-time pupil segmentation under baseline dark-adaptation and light-stimulation phases. The circular infrared LED array enables continuous tracking of the pupil boundaries throughout the examination. (B) Synchronized pupillometry waveforms of the right eye (OD, orange) and left eye (OS, blue) recorded over a 25 000 ms trial. Grey shaded bands denote periods of dark intervals. During left-eye stimulation intervals, the left eye demonstrated a reduced direct pupillary constriction amplitude and more attenuated wave amplitudes (0.71 mm; 18.98% constriction) compared with the right eye (1.16 mm; 26.42% constriction). These asymmetric pupillary responses were classified by the Pupil X algorithm as a Grade 3 relative afferent pupillary defect.

Table 3.4. Comparative analysis of baseline pupillary dynamics between the Pupil X and manual method (n = 282 eyes).

Metric	Manual Examination	Pupil X VR	*p*-value
Minimum Pupil Diameter (mm; mean ± SD)	3.75 ± 1.13	3.06 ± 0.96	< 0.001***
Maximum Pupil Diameter (mm; mean ± SD)	5.06 ± 1.25	4.12 ± 1.04	< 0.001***

Independent two-tailed t-test for all p-values; Statistically significant: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

3.9. Agreement between Pupil X and Clinical Manual Methods

Inter-method agreement was assessed using Bland–Altman analysis for minimum and maximum pupil diameter measurements. For minimum pupil diameter under light stimulation, the Pupil X platform demonstrated a mean bias of −0.69 mm (Figure 3.17), indicating that the VR system systematically underestimated pupil diameter relative to manual assessment. The 95% limits of agreement ranged from −2.86 mm to 1.48 mm. The mean bias was slightly negative, indicating that the VR system consistently produced lower measurements compared to manual pupillometry. The dispersion of the data points remained relatively constant across the range of measured pupil diameters, suggesting homoscedastic measurement error with no apparent proportional bias.

For maximum pupil diameter under dark condition intervals, the mean bias was −0.94 mm (Figure 3.18), with 95% limits of agreement ranging from −2.92 mm to 1.03 mm. The mean bias was again found to be slightly negative, indicating overall lower pupil diameter measurements by the Pupil X platform relative to the standard manual examination. The dispersion of the data points appears uniform across the range of measured pupil diameters, suggesting that the measurement

error is homoscedastic and independent of the absolute pupil size. Furthermore, a single outlier was identified in both the minimum constriction and maximum dilation Bland–Altman plots, in which the Pupil X VR system recorded substantially larger pupil diameters than the corresponding manual measurements. No other observations demonstrated comparable deviation from the overall pattern of agreement.

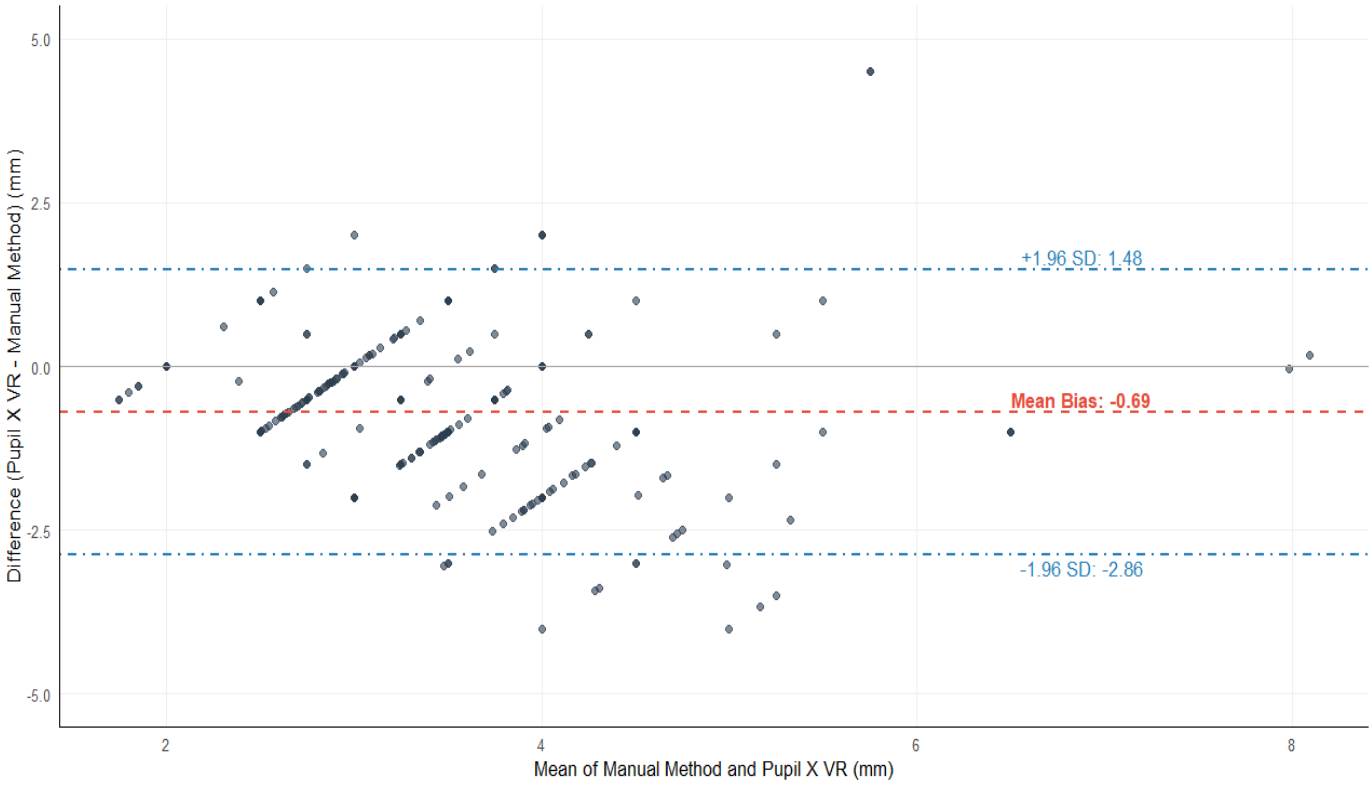


Figure 3.17. Bland-Altman plot comparing minimum constriction diameters between Pupil X platform and the manual method. Individual data points (N = 282 eyes) represent the difference between paired measurements plotted against their mean minimum constriction diameter (mm). The dashed horizontal line indicates the mean bias (−0.69 mm), and the dash-dotted lines represent the 95% limits of agreement (±1.96 SD). Measurements span a minimum constriction diameter range of approximately 1.5 – 8.0 mm across the study cohort.

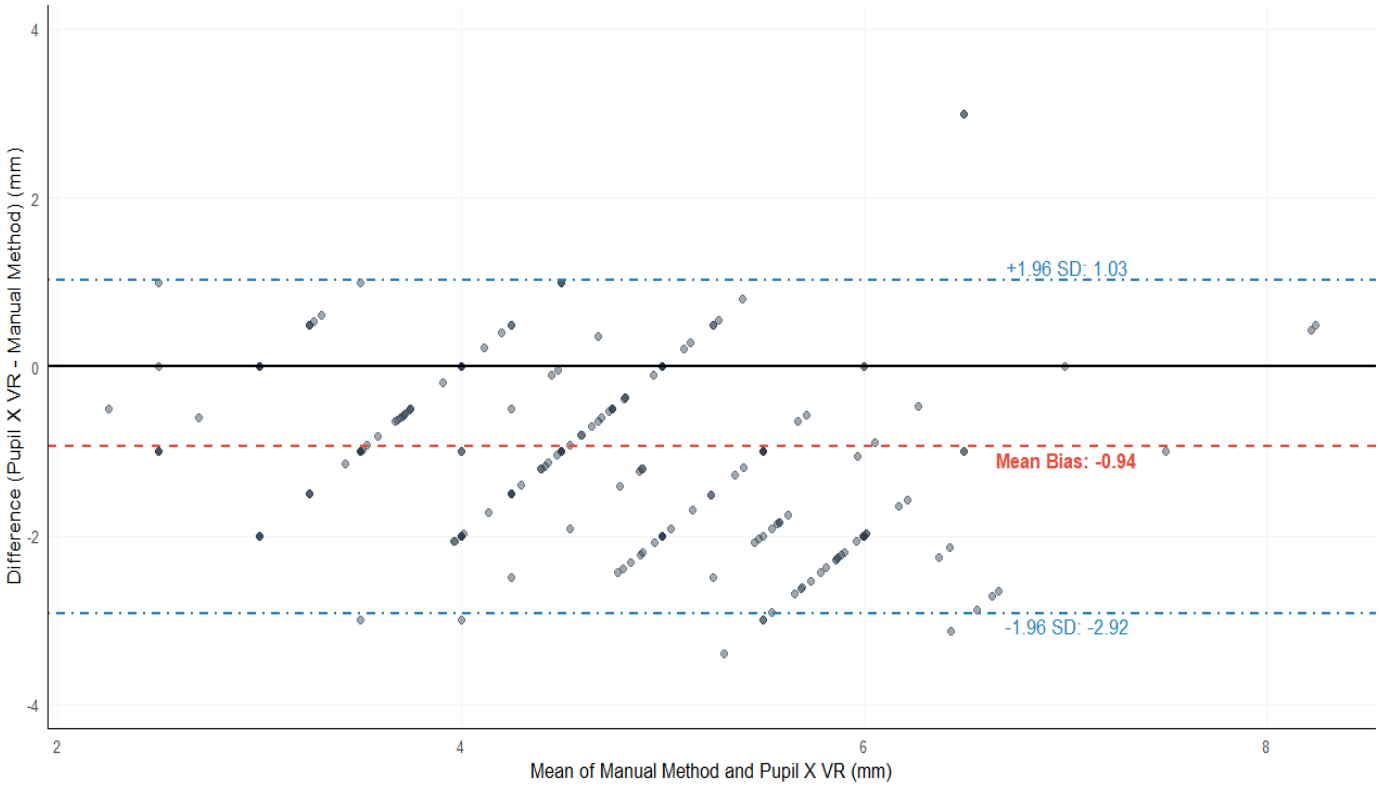


Figure 3.18. Bland-Altman plot comparing maximum pupil diameters measurements obtained with Pupil X platform and the manual method. Individual data points (N = 282 eyes) represent the difference between paired measurements plotted against their mean maximum pupil diameter (mm). The dashed horizontal line indicates the mean bias (−0.94 mm), and the dash-dotted lines represent the 95% limits of agreement (±1.96 SD). Measurements span a maximum pupil diameter range of 2.0–8.5 mm across the study cohort.

3.10. Relative Afferent Pupillary Defect (RAPD) Detection

The Pupil X VR platform identifies and grades RAPDs by continuously comparing inter-eye pupillary responses during alternating monocular light stimulation. It computes an inter-eye asymmetry index from real-time kinetic parameters, including constriction amplitude and percentage change from baseline diameters. Based on previous studies with this device, an inter-eye asymmetry percentage change cut-off value of 1.17 was identified as the optimal threshold for distinguishing RAPD-positive cases from healthy controls (Negi et al., 2023). Values at or below this threshold (≤ 1.17) are classified as normal inter-eye symmetry (i.e., no RAPD present), whereas values exceeding 1.17 are classified as RAPD positive. The magnitude of the inter-eye

asymmetry index is subsequently used to assign categorical severity grades (e.g., Grade 1+ to 4+ RAPD).

In a normal participant case, pupillary responses were closely matched between eyes, with minimal inter-eye variability. For example, constriction amplitudes of 1.14 mm (27.78%) and 1.16 mm (28.33%) were recorded in the right and left eye, respectively (Figure 3.15). This case demonstrated inter-eye differences under the system's RAPD detection threshold, and as such were classified as "*No RAPD*". In a RAPD-positive patient case, a reduction in constriction amplitude and response strength was observed in the affected eye during stimulation cycles (Figure 3.16). For instance, left-eye stimulation produced a reduced response of 0.71 mm (18.98%), compared with 1.16 mm (26.42%) during right-eye stimulation. These inter-eye discrepancies exceed the platform's classification threshold and thus were graded as a "*Grade 3 RAPD*".

To evaluate the diagnostic accuracy of the Pupil X platform against the clinical standard, data from n = 141 patients were compiled into a confusion matrix. As shown in Table 3.5 and 3.6, the Pupil X platform demonstrated a sensitivity of 76.67% (95% CI: 59.07%–88.21%), correctly identifying 23 of 30 clinically confirmed RAPD-positive cases, and a specificity of 63.06% (95% CI: 53.79%–71.46%). The negative predictive value (NPV) was high at 90.91% (95% CI: 82.40%–95.53%), indicating strong performance for ruling out normal inter-eye symmetry. In contrast, the positive predictive value (PPV) was lower at 35.94% (95% CI: 25.29%–48.18%), reflecting 41 false-positive classifications in the non-pathological group.

Table 3.5. Contingency matrix for RAPD detection between VR and manual methods

	Clinical Reference (RAPD present)	Clinical Reference (RAPD Absent)	Total
VR Positive	23 (True Positive)	41 (False Positive)	**64**
VR Negative	7 (False Negative)	(70 True Negative)	**77**
Total	**30**	**111**	**141**

Table 3.6. Diagnostic performance metrics derived from contingency matrix

Metrics	Value (%; 95% CI)
Sensitivity	76.67; (59.07 – 88.21)
Specificity	63.06; (53.79 – 71.46)
Positive predictive value (PPV)	35.94; (25.29 – 48.18)
Negative predictive value (NPV)	90.91; (82.40 – 95.53)

3.11. Analysis of Diagnostic Discordance – False Positive Characterization

Forty-one false-positive cases were identified in which Pupil X classified a RAPD that was not detected on clinical examination. A comprehensive chart review revealed that many of these cases occurred in patients with coexisting ocular conditions, including Adie's tonic pupil, longstanding ptosis, diplopia with strabismus, cataracts and other media opacities, and glaucoma. Several false-positive cases also involved multiple concurrent ocular abnormalities, representing clinically complex presentations. These findings characterize the clinical features of the false-

positive cohort and identify common ocular comorbidities present in cases misclassified by the device (Table. 3.7).

3.11.1 False Negative Characterization

Seven false-negative cases were identified in which Pupil X failed to detect a clinically apparent RAPD. One notable outlier involved a patient with a severe (+4) RAPD and coexisting active optic nerve glioma and high pathological myopia, representing the only severe RAPD within the false-negative cohort. The remaining six false-negative cases occurred exclusively in eyes with clinically mild or borderline RAPDs.

Table 3.7. Clinical breakdown of discordant cases (false positives and false negatives)

Error type	Sample size (*n*)	Diagnosed ocular pathologies	Proposed reason for misclassification
False Positives	41	Idiopathic intracranial hypertension Strabismus/diplopia Cataracts/media opacities Optic Neuropathy Ptosis	**Eye tracking artifacts:** ▪ Ocular misalignment and ptosis can alter the infrared camera's tracking of the pupillary margin, simulating an afferent defect. ▪ Media opacities can induce asymmetric light attenuation.
False Negatives	7	Optic nerve glioma Pathologic myopia Papilledema Meningioma Cataract Dry eye	**Fixation/Luminance floor effects:** ▪ High refractive errors (pathologic myopia) may distort peripheral infrared reflection ▪ Mild cases of RAPDs may fall below the algorithm's initial stimulus-response noise threshold.

The clinical mapping in Table 3.7 showed that longstanding ptosis, strabismus, high myopia, and other ocular abnormalities were frequently represented among misclassified cases. These clinical characteristics were observed across both false-positive and false-negative

classifications and provide a summary of the ocular conditions present within the misclassified cohort.

3.11.2. Diagnostic Agreement for RAPD Detection and Classification

To quantify diagnostic agreement between the Pupil X VR platform and the manual clinical reference standard for binary RAPD classification (present vs. absent), a Cohen's kappa analysis was performed (n = 141) (Table 3.8). The observed agreement was 66 % (n = 93/141), compared with an expected chance agreement of 53%, yielding a Cohen's κ of 0.28 (95% CI: 0.12–0.45; $p < 0.001$), indicating fair agreement. The kappa magnitude was influenced by asymmetric marginal distributions, driven in part by 41 cases classified as RAPD-positive by the automated system but considered normal on clinical examination.

An ordinal agreement analysis was performed using a quadratic weighted kappa (K_w) across the five-tier classification scale (Grades 0–4) (Table 3.8). The analysis yielded a K_w of 0.48 (95% CI: 0.33–0.63), indicating a statistically significant moderate level of agreement ($p < 0.001$). Most discordant classifications represented minor disagreement, with single-grade differences predominantly involving Grade 1 classifications (n = 32/41; 78%). However, disagreement at the extremes of the scale was also observed. This included Grade 4 false negatives (n = 4) and Grade 4 false positives (n = 4), reflecting misclassification at the highest severity level and contributing to reduced overall ordinal agreement.

Table 3.8. Agreement analysis between Pupil X VR and manual RAPD grading

Analysis Type	Observed Rate	Chance-Expected Rate	Kappa Index (95% CI)	Agreement Interpretation
Binary classification (RAPD present vs. absent)	**65.96%** (Observed agreement, P_o)	**52.65%** (Expected agreement, P_e)	**0.281** (0.116–0.446)	Fair
Ordinal severity grading (Grades 0–4)	**4.375** (Observed weighted disagreement, D_o)	**8.435** (Expected weighted disagreement, D_e)	**0.481** (0.334–0.628)	Moderate

CHAPTER 4. DISCUSSION

4.1. Overview of principal findings

The primary objective of this study was to evaluate the clinical performance of integrated VR-based perimetry and pupillometry platforms relative to established clinical reference standards. This study evaluated the clinical performance of VR-based perimetry and pupillometry in a diverse cohort of 413 participants, including healthy individuals and patients with a range of ocular and neuro-ophthalmic conditions. The study population included a broad age range (20–89 years) and comparable demographic characteristics across device cohorts, reducing potential confounding related to age and sex. The inclusion of conditions such as glaucoma, idiopathic intracranial hypertension, optic neuritis, and other ophthalmic disorders enabled evaluation across a wide spectrum of visual field and pupillary abnormalities. Overall, this heterogeneous cohort provided a clinically relevant framework for assessing the agreement, diagnostic performance, and potential clinical applications of VR-based ophthalmic technologies. The findings demonstrated differing levels of performance between the VR platforms, with RVF showing strong agreement and OM demonstrating moderate agreement with conventional HFA, while Pupil X showed potential as an objective screening tool for RAPD assessment.

4.2. Comparison of Test Durations and Paradigms across Perimetric Platforms

The evolution of VR-based perimetry reflects a shift in ophthalmic diagnostics toward portable, patient-centered technologies that can extend visual assessment beyond conventional clinic settings. Unlike the HFA, both OM and RVF platforms eliminate the need for dedicated perimetry rooms, bulky stationary equipment, trial lenses, and manual eye patching, allowing testing to be performed under standard room lighting with minimal setup. These practical advantages may improve accessibility in community clinics, emergency departments, bedside

assessments, and teleophthalmology programs, where conventional automated perimetry is often impractical. Similar advantages have been reported for several commercially available and investigational portable perimeters, which have demonstrated clinically comparable visual field assessments while reducing infrastructure requirements and expanding opportunities for community-based, bedside, and remote testing (McLaughlin et al., 2022; Selvan et al., 2023; Tsapakis et al., 2018). Collectively, these findings suggest that portability and accessibility are emerging as some of the main advantages of VR perimetry without substantially compromising diagnostic performance.

These advantages, however, come with important trade-offs. Both VR systems provided less detailed diagnostic outputs than the HFA. The OM generated lower-resolution visual field maps, while the RVF omitted numerical total and pattern deviation values that are routinely used to identify and monitor localized visual field defects. Consequently, although these devices appear well suited for screening, remote monitoring, and situations where portability is prioritized, they may currently provide less detailed information for longitudinal disease monitoring than conventional automated perimetry. This trade-off between accessibility and analytical detail has also been recognized across other portable perimetric systems, where simplified hardware and software designs improve usability but often provide fewer diagnostic metrics than conventional bowl perimeters (Selvan et al., 2024).

4.2.1. Agreement Analysis with the Standard Humphrey Field Analyzer

The Bland–Altman analyses demonstrated that the RVF platform showed closer agreement with the HFA than the OM platform, with smaller mean biases and narrower limits of agreement for both MD and PSD. These findings indicate that RVF more consistently reproduced HFA-derived estimates of both diffuse and localized visual field loss, whereas the greater variability

observed with OM suggests that its measurements should be interpreted with greater caution, particularly when monitoring disease progression over time.

The higher level of agreement observed with RVF is consistent with previous literature evaluating portable and VR-based perimetry. VR perimeters in the past have demonstrated high score correlation and agreement with conventional automated perimetry, although performance varied according to testing algorithms, hardware design, and patient population (Chen et al., 2022; Heinzman et al., 2022; Phu & Kalloniatis, 2021). Studies evaluating these devices have similarly reported good agreement with HFA-derived global indices, supporting the growing role of VR-based perimetry as an alternative for screening and remote assessment while acknowledging that conventional perimetry remains the benchmark for longitudinal disease monitoring.

The comparatively poorer agreement observed with OM likely reflects differences in the operational aspect of the program rather than limitations inherent to VR technology. The OM employs a full-threshold strategy, whereas the HFA and RVF both use adaptive thresholding algorithms that reduce examination time and minimize fatigue-related variability. Longer testing durations have consistently been associated with decreased reliability due to patient fatigue, fixation losses, and response variability, all of which may contribute to the wider limits of agreement observed in this study (Johnson et al., 1988; Wild et al., 1991). Consequently, the present findings suggest that the performance of VR perimetry depends not only on the hardware platform but also on the underlying thresholding strategy used to estimate visual field sensitivity.

From a clinical perspective, the narrow limits of agreement observed with the RVF suggest that it may be suitable for screening and routine monitoring of visual field loss, particularly in settings where access to the conventional HFA is limited. Nevertheless, the wider limits of agreement observed with the OM indicate that individual measurements should be interpreted

cautiously, especially when monitoring small longitudinal changes in visual field sensitivity. Overall, these findings support the potential of RVF as a clinically useful portable perimeter that more closely reproduces HFA measurements than OM. Nevertheless, despite the encouraging agreement demonstrated by RVF and other contemporary VR perimeters, the HFA remains the reference standard for monitoring subtle longitudinal changes until further evidence establishes the long-term repeatability and interchangeability of these emerging technologies.

4.2.2. Spatial Fidelity and Agreement of Visual Field Defects

Beyond comparisons of global indices, spatial agreement is critical because clinical decision-making often depends on the accurate localization and characterization of visual field defects in addition to changes to visual field indices like MD and PSD. Accurate representation of defect morphology is particularly important for differentiating glaucomatous from neuro-ophthalmic disease, monitoring progression, and correlating functional deficits with structural imaging (Lucy & Wollstein, 2016). The qualitative case comparisons and pointwise ΔdB heat maps demonstrated that both VR platforms generally preserved the overall spatial distribution of visual field defects identified by the HFA, although quantitative agreement varied according to the severity and complexity of visual field loss.

For the OM platform, agreement was greatest in normal visual fields and in regions of absolute visual field loss, where pointwise sensitivity measurements differed little from those obtained with the HFA. Greater variability was consistently observed within transitional regions bordering scotomas and areas of partial sensitivity loss. These regions contain the steepest sensitivity gradients and are inherently more susceptible to threshold variability, making small differences in testing algorithms or patient responses more likely to produce measurable pointwise discrepancies. Similar findings have been reported in studies of conventional automated perimetry,

where variability increases near defect margins and in areas of moderate sensitivity loss, regardless of the testing platform (Gardiner et al., 2014). Consequently, the wider pointwise differences observed with OM likely reflect the inherent difficulty of threshold estimation in transitional zones rather than a failure to detect the underlying defect.

The RVF platform demonstrated similar spatial trends while consistently preserving HFA-defined defect morphology across a broad range of pathological conditions. Physiological blind spots, arcuate defects, glaucomatous constriction, and diffuse neuro-ophthalmic field loss were reproduced with adequate spatial correspondence, suggesting that the platform reliably captures clinically meaningful patterns of visual field loss. Although quantitative differences increased in advanced diseases, particularly along defect borders and within regions of heterogeneous sensitivity, the overall configuration and location of visual field defects remained largely preserved. These findings are consistent with previous evaluations of VR-based perimeters, which have been shown to generally reproduce clinically relevant visual field patterns despite modest pointwise differences compared with conventional automated perimetry (Chia et al., 2022, 2024). Collectively, these studies suggest that preservation of defect topology is as clinically important as relative pointwise agreement when the objective is disease detection or remote assessment.

The reduced quantitative agreement observed in advanced disease likely reflects several factors common to automated perimetry. At very low sensitivity levels, threshold measurements become increasingly variable because of greater response fluctuation, fixation instability, fatigue, and the limited dynamic range of perimetric algorithms (Miranda & Henson, 2008; Montolio et al., 2012b; Stewart & Hunt, 1993). Consequently, broader transition zones observed in the VR-derived profiles are not unexpected and have also been described with conventional perimetry when assessing advanced glaucomatous damage. Importantly, despite this increased variability,

both VR platforms continued to identify the presence, location, and overall extent of clinically significant visual field defects.

These findings suggest that one of the strengths of VR perimetry lies in its potential to preserve spatial patterns of visual field loss rather than replicate every pointwise threshold measured by the HFA. This distinction is particularly relevant for screening, triage, teleophthalmology, and community-based eye care, where identifying the presence and distribution of a defect is often more important than obtaining identical threshold values. However, for longitudinal monitoring of progressive diseases such as glaucoma, where subtle changes at defect margins may influence management decisions, the higher spatial resolution and more comprehensive analytical outputs of conventional automated perimetry remain advantageous. Future improvements in VR perimetry should therefore focus on enhancing spatial resolution, optimizing thresholding algorithms, and improving quantitative accuracy within transitional regions while maintaining the accessibility and portability that distinguish these systems.

4.3. Agreement of Baseline Pupillometry Metrics

Comparative analysis of baseline pupil diameter measurements demonstrated a systematic difference between Pupil X VR and manual assessment, with Pupil X consistently producing smaller pupil diameter measurements. Unlike previous reports demonstrating strong agreement between VR-based pupillometers and manual measurements, our results indicate the presence of a systematic measurement bias (Karaer et al., 2025; Park et al., 2026c). In addition, a notable feature of the Bland–Altman plots were the presence of distinct, parallel diagonal bands. This pattern can arise when a discrete measurement method is compared with a continuous system and reflects differences in measurement resolution rather than true disagreement between techniques (Bland & Altman, 1999). In the present study, manual pupil measurements were

recorded in rounded increments (e.g., to the nearest 0.5 mm or 1.0 mm), whereas Pupil X continuously tracked pupil size with finer resolution. Since the Bland–Altman plot displayed the difference between methods against their mean, observations with the same rounded manual measurement can form diagonal bands when paired with continuously varying Pupil X measurements. Thus, the observed pattern can also reflect differences in measurement granularity. Clinically, the finer resolution of Pupil X may allow detection of subtle changes in pupil size that are not captured by routine manual assessment, although the clinical significance of these small differences remains uncertain.

The observed systematic bias may reflect differences in illumination conditions, image acquisition protocols, and device-specific image processing algorithms used to identify the pupil boundary, all of which have been shown to influence pupillary measurements and agreement between measurement techniques (W. Chen et al., 2026; Steinhauer et al., 2022). Automated infrared and VR-based pupillometers estimate pupil diameter from infrared video images using software-based pupil detection and segmentation algorithms, whereas manual pupillary assessment is influenced by examiner positioning, ambient illumination, ruler alignment, and subjective identification of the pupillary margin (Kelbsch et al., 2019; Smith et al., 2020). Thus, the difference between Pupil X and manual measurements likely reflects a modality-specific measurement bias rather than reduced measurement reliability. Device-specific reference ranges and calibration will be important to consider when interpreting pupil measurements across different assessment methods.

4.3.1 Diagnostic Performance of Pupil X and RAPD Screening and Grading

Pupil X demonstrated moderate sensitivity and high negative predictive value for RAPD detection, supporting its potential role as an objective screening tool. These findings are consistent

with recent VR-based pupillometry studies, which have demonstrated the feasibility of automated RAPD assessment and highlighted the potential of these systems for rapid evaluation in community, emergency, and teleophthalmology settings (Negi et al., 2024; Park et al., 2025). The high negative predictive value observed in this study suggests that Pupil X may be particularly useful for identifying patients unlikely to have clinically significant RAPD and prioritizing those requiring further neuro-ophthalmic evaluation.

The lower specificity and positive predictive value likely reflect challenges associated with threshold-based automated classification. VR pupillometers rely on predefined inter-eye asymmetry thresholds to improve objectivity; however, these thresholds may also detect subtle physiological differences that overlap with borderline clinical responses (Negi et al., 2024). Additionally, the higher quadratic weighted kappa compared with binary kappa found in this study further suggests that discrepancies were primarily related to differences in severity grading rather than complete classification disagreement, with most false-positive results representing mild Grade 1 deviations.

In this cohort, discordant cases were observed among participants with strabismus, ptosis, media opacities, and high myopia. These conditions may reduce the accuracy of VR pupillometry by impairing eye alignment, obscuring visualization of the pupil, or reducing image quality, thereby affecting automated pupil detection and tracking algorithms (Philibert & Milea, 2024). This interpretation is consistent with recommendations emphasizing that image quality and acquisition conditions influence the accuracy and reproducibility of automated pupil measurements (Larson & Singh, 2016). Overall, Pupil X demonstrates potential as an accessible VR-based RAPD screening tool; however, further validation and threshold refinement are required to optimize specificity across diverse clinical populations.

4.4. Study Limitations

Several limitations should be considered when interpreting the findings of this study. Although an overall recruitment target of approximately 500 participants was established a priori, recruitment was not achieved equally across the individual VR platforms. Recruitment for the OM platform was prematurely terminated because of device-specific technical and operational limitations, including device malfunction and compatibility issues. Consequently, OM testing was not tested concurrently with RVF testing, and a direct within-cohort comparison between the two VR platforms was not possible. Therefore, differences observed between OM and RVF cannot be attributed solely to device performance, as differences in participant characteristics, recruitment period, and clinical case mix may also have contributed. Future studies may evaluate both platforms concurrently within the same participant population and use standardized testing protocols to enable a direct head-to-head comparison and more robust assessment of relative device performance.

For visual field testing, agreement analyses were limited to cross-sectional comparisons at a single time point. Consequently, while the study demonstrated agreement with the reference standards, it could not evaluate test–retest repeatability or long-term reproducibility, both of which are essential for monitoring disease progression in clinical practice. Although longitudinal follow-up was initially planned, participant attrition and technical challenges with the investigational devices limited the availability of follow-up data.

For pupillometry, the heterogeneous study cohort included participants with a wide range of structural, optical, and neuro-ophthalmic conditions, including strabismus, diplopia, ptosis, cataracts, other media opacities, high myopia, papilledema, optic neuropathies, and intracranial lesions. These conditions can influence pupillary measurements and may have affected the

diagnostic performance of Pupil X, limiting the ability to determine its accuracy within specific disease populations (Larson & Singh, 2016; Philibert & Milea, 2024). Another notable limitation of the device was the proportion of participants with incomplete datasets. Of the 186 participants enrolled for pupillometry, 45 (24%) had incomplete data. Importantly, incomplete datasets were not attributable to participant inability or unwillingness to complete the assessment. Deidentified data was collected on-site and subsequently stored on an internationally hosted cloud-based platform; however, a subset of recorded data was subsequently lost and could not be recovered. As a result, some participants who had completed the assessment did not have complete datasets available for analysis. Although this data loss does not necessarily reflect the technical performance of the VR pupillometry system itself, it represents an important operational limitation that may affect the overall feasibility of its implementation in clinical practice. This finding highlights the importance of robust data-management, storage, backup, and retrieval procedures when integrating cloud-based diagnostic technologies into clinical workflows. Accordingly, the relatively high proportion of incomplete datasets should be considered when interpreting the feasibility of the VR-based pupillary assessment. Future implementations should incorporate redundant data-storage mechanisms, routine verification of successful data synchronization, and established data-recovery procedures to minimize the risk of data loss and ensure the reliability and continuity of clinical data.

4.5. Future Directions

Future research should focus on multicenter studies with larger, more balanced cohorts to improve the generalizability of these findings and enable disease-specific analyses. Stratifying participants by conditions such as glaucoma, optic neuropathies, and other neuro-ophthalmic disorders would provide a better understanding of device performance across different clinical

populations and help establish condition-specific diagnostic thresholds. Such validation has been identified as an important step toward the clinical adoption of VR-based ophthalmic technologies (Ahuja et al., 2025; Selvan et al., 2024).

Longitudinal studies are also needed to establish the test–retest repeatability and long-term reproducibility of VR-based perimetry and pupillometry. Repeated assessments would help distinguish true disease progression from normal measurement variability and determine whether these technologies are suitable for monitoring chronic ophthalmic and neuro-ophthalmic conditions. The portability and automated nature of VR-based ophthalmic devices also support their potential use in teleophthalmology, community screening, and remote patient monitoring, particularly in underserved regions with limited access to specialist eye care. Future implementation studies should therefore evaluate their clinical workflow, patient feedback, and cost-effectiveness to determine their feasibility in routine practice.

Finally, continued improvements in hardware and software may further enhance diagnostic performance. Advances in eye tracking, infrared imaging, and image-processing algorithms may improve measurement accuracy, thereby enhancing the reliability of VR-based visual function testing.

4.6. Conclusion

In conclusion, the RVF platform demonstrated strong agreement with the HFA for visual field assessment, supporting its potential as an alternative method for visual field testing. On the other hand, the OM platform and Pupil X demonstrated moderate agreement with conventional clinical assessments, indicating that further refinement and validation are warranted before routine clinical implementation. These findings support the potential of VR-based perimetry as an accessible approach for detecting clinically relevant visual field defects. Although VR devices are

not intended to replace conventional automated perimetry, their cost-effectiveness, portability, accessibility, and reduced infrastructure requirements provide important advantages for use in remote, clinical, and bedside settings. By eliminating the need for dedicated testing rooms and specialized equipment, VR-based platforms may improve access to visual field assessment and serve as effective initial screening tools to identify patients who require further comprehensive ophthalmic evaluation. Future studies evaluating longitudinal performance, clinical workflow integration, and cost-effectiveness will be important to establish their role in routine ophthalmic care.

REFERENCES

Ahuja, A. S., Paredes, A. A., Eisel, M. L. S., Ahuja, S. A., Wagner, I. V., Vasu, P., Dorairaj, S., Miller, D., & Abubaker, Y. (2025). The Utility of Virtual Reality in Ophthalmology: A Review. *Clinical Ophthalmology (Auckland, N.Z.)*, *19*, 1683–1692. https://doi.org/10.2147/OPTH.S517974

All-in-one standalone VR Headset | Pico Neo 2 Eye - Tobii. (n.d.). Retrieved July 1, 2026, from https://www.tobii.com/products/integration/xr-headsets/device-integrations/pico-neo-2-eye

Bell, R. A., Waggoner, P. M., Boyd, W. M., Akers, R. E., & Yee, C. E. (1993). Clinical Grading of Relative Afferent Pupillary Defects. *Archives of Ophthalmology*, *111*(7), 938–942. https://doi.org/10.1001/archopht.1993.01090070056019

Biousse, V., & Newman, N. J. . (2016). *Neuro-ophthalmology illustrated*.

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. Statistical methods in medical research, 8(2), 135–160. https://doi.org/10.1177/096228029900800204

Bowling, B. (2024). *Kanski's Clinical Ophthalmology, Eighth Edition (2024)*. https://doi.org/10.1016/B978-0-7020-5572-0.00025-8

Bryan, B. T., Pomeranz, H. D., & Smith, K. H. (2014). Complete binasal hemianopia. *Proceedings (Baylor University. Medical Center)*, *27*(4), 356. https://doi.org/10.1080/08998280.2014.11929158

Buratto, L., Brint, S. F., & Sacchi, L. (1990). The Pupils. *Cataract Surgery: Introduction and Preparation*, 39–42. https://doi.org/10.1201/9781003522928-7

Chen, W., Jiang, X., Guo, X., Lin, J., Dou, N., & Tang, M. (2026). Comparison of automated ultrasound pupillometry assessment vs. infrared pupillary in critically ill patients. *Frontiers in Medicine*, *12*, 1722645. https://doi.org/10.3389/FMED.2025.1722645

Chen, Y. T., Yeh, P. H., Cheng, Y. C., Su, W. W., Hwang, Y. S., Chen, H. S. L., Lee, Y. S., & Shen, S. C. (2022). Application and Validation of LUXIE: A Newly Developed Virtual Reality Perimetry Software. *Journal of Personalized Medicine*, *12*(10), 1560. https://doi.org/10.3390/JPM12101560

Chia, Z. K., Kong, A. W., Turner, M. L., Backus, B. T., Deiner, M., & Ou, Y. (2022). Comparison of a virtual reality-based visual field test to conventional perimetry and OCT. *Investigative Ophthalmology & Visual Science*, *63*(7), 3103–3103.

Chia, Z. K., Kong, A. W., Turner, M. L., Saifee, M., Damato, B. E., Backus, B. T., Blaha, J. J., Schuman, J. S., Deiner, M. S., & Ou, Y. (2024). Assessment of Remote Training, At-Home Testing, and Test-Retest Variability of a Novel Test for Clustered Virtual Reality Perimetry. *Ophthalmology Glaucoma*, *7*(2), 139–147. https://doi.org/10.1016/j.ogla.2023.08.006

Choplin, N. T. ., & Edwards, R. P. . (1998). *Visual fields*. 255.

Dalley, A. F. ., Agur, A. M. R. ., & Moore, K. L. . (2023). *Moore's Clinically Oriented Anatomy, 9e*. Lippincott Williams & Wilkins, a Wolters Kluwer business.

Farrell, K., & MacDougall, D. (2023). An Overview of Clinical Applications of Virtual and Augmented Reality. *An Overview of Clinical Applications of Virtual and Augmented Reality: CADTH Horizon Scan*. https://www.ncbi.nlm.nih.gov/books/NBK596298/

Gardiner, S. K., Swanson, W. H., Goren, D., Mansberger, S. L., & Demirel, S. (2014). Assessment of the reliability of standard automated perimetry in regions of glaucomatous damage. *Ophthalmology*, *121*(7), 1359. https://doi.org/10.1016/J.OPHTHA.2014.01.020

Global leader in eye tracking for over 20 years - Tobii. (n.d.). Retrieved July 1, 2026, from https://www.tobii.com/?creative=805108963399&keyword=tobii&matchtype=p&network=g&device=c&utm_source=google&utm_medium=cpc{ifdisplay:display}{ifvideo:video}&utm_campaign=&utm_term=tobii&utm_content=g&gad_source=1&gad_campaignid=19079851648&gbraid=0AAAADcpHK8BJkbZc-AAkS_-DEZLS6Xvx&gclid=CjwKCAjwmJjSBhB-EiwAkZgxiwQevjsZOVBwKplrNK6D5lPTW09OS3GaoBswvNaPeeges8bmwVGpHBoCR3MQAvD_BwE

Gupta, M., Ireland, A. C., Omole, A. E., & Bordoni, B. (2026). Neuroanatomy, Visual Pathway. *StatPearls*. https://www.ncbi.nlm.nih.gov/books/NBK553189/

Hamann, S., Obaid, H. G., & Celiz, P. L. (2015). Binasal hemianopia due to bilateral internal carotid artery atherosclerosis. *Acta Ophthalmologica*, *93*(5), 486–487. https://doi.org/10.1111/AOS.12565

Heinzman, Z., Alawa, K., Marín-Franch, I., Turpin, A., & Wall, M. (2022). Validation of visual field results of a new open-source virtual reality headset. *Investigative Ophthalmology & Visual Science*, *63*(7), 1259-A0399-1259 – A0399.

Hutchins, J., Aune, L., Bevan, C., Anderson, J. T., & Rebarchik, E. (2024). *Introduction to Neuroscience*. Weber State University. https://uen.pressbooks.pub/introneuro/

Johnson, C. A., Adams, C. W., & Lewis, R. A. (1988). Fatigue effects in automated perimetry. *Applied Optics*, *27*(6), 1030. https://doi.org/10.1364/AO.27.001030

Kaeser, P. F., & Kawasaki, A. (2010). Disorders of Pupillary Structure and Function. *Neurologic Clinics*, *28*(3), 657–677. https://doi.org/10.1016/j.ncl.2010.03.007

Kandel, E. R. ., Koester, John., Mack, Sarah., & Siegelbaum, Steven. (2021). *Principles of neural science*. 1646.

Karaer, I., Yoon, H.-J., Ma, R., Savant, R., Rodwell, V., Shenoy, R., Tu, Z., Arshad, Q., Mukaetova-Ladinska, E. B., & Thomas, M. G. (2025). *Evaluation of a Virtual Reality-Based Eye Tracker for Neuro-Ophthalmic Assessment: A Feasibility, Reliability and Reproducibility Study*. https://doi.org/10.21203/RS.3.RS-6966745/V1

Katz, J., & Sommer, A. (1986). Asymmetry and variation in the normal hill of vision. *Archives of Ophthalmology (Chicago, Ill. : 1960)*, *104*(1), 65–68. https://doi.org/10.1001/ARCHOPHT.1986.01050130075023

Kelbsch, C., Strasser, T., Chen, Y., Feigl, B., Gamlin, P. D., Kardon, R., Peters, T., Roecklein, K. A., Steinhauer, S. R., Szabadi, E., Zele, A. J., Wilhelm, H., & Wilhelm, B. J. (2019). Standards in Pupillography. *Frontiers in Neurology*, *10*. https://doi.org/10.3389/FNEUR.2019.00129

Kumar, A. (2018). *How to examine Confrontational Visual Fields*. https://buos.co.uk/wp-content/uploads/2020/05/BUOS-HowToExamineConfrontVisualFields.pdf

Lewis, R. J., Chu, B. X., Winters, T. M., & Spicer, C. M. (2025). Visual Field Assessment and Disability Evaluation. *Visual Field Assessment and Disability Evaluation*, 1. https://doi.org/10.17226/29124

Liu, Volpe, and Galetta's Neuro-Ophthalmology - 3rd Edition | Elsevier Shop. (n.d.). Retrieved June 27, 2026, from https://shop.elsevier.com/books/liu-volpe-and-galetta-s-neuro-ophthalmology/liu/978-0-323-34044-1

Lu, S. J., Girgis, S., Shah, P., & Lee, G. A. (2024). Patient Experience and Barriers to the Visual Field Test for Glaucoma. *Journal of Glaucoma*, *33*(11), 835–840. https://doi.org/10.1097/IJG.0000000000002477

Lucy, K. A., & Wollstein, G. (2016). Structural and Functional Evaluations for the Early Detection of Glaucoma. *Expert Review of Ophthalmology*, *11*(5), 367. https://doi.org/10.1080/17469899.2016.1229599

McLaughlin, D., Munshi, H., Savatovsky, E., Vanner, E., Chang, T. C., & Grajewski, A. L. (2022). Visual Field Testing in a Telehealth Setting: Remote Perimetry Using a Head-Mounted

Device in Normal Eyes. *Investigative Ophthalmology & Visual Science*, *63*(7), 1265-A0405-1265 – A0405.

Mijares, G., Gameiro, G., Matosas, M., Arrieta, E. A., Kashem, R., Shousha, M. A., Parrish, R., & McSoley, J. (2024). Assessment of the Reliability of Pupillometer Measurements Obtained from a Virtual Reality Headset: A Control Subject Analysis. *Investigative Ophthalmology & Visual Science*, *65*(7), 5491–5491.

Miller, N. R. ., Subramanian, P. S. ., Patel, V. R. ., Walsh, F. B. ., & Hoyt, W. Fletcher. (2021). *Walsh and Hoyt's Clinical Neuro-Ophthalmology: The Essentials, 4e*. Lippincott Williams & Wilkins, a Wolters Kluwer business.

Miranda, M. A., & Henson, D. B. (2008). Perimetric sensitivity and response variability in glaucoma with single-stimulus automated perimetry and multiple-stimulus perimetry with verbal feedback. *Acta Ophthalmologica*, *86*(2), 202–206. https://doi.org/10.1111/J.1600-0420.2007.01033.X

Montelongo, M., Gonzalez, A., Morgenstern, F., Donahue, S. P., & Groth, S. L. (2021). A Virtual Reality-Based Automated Perimeter, Device, and Pilot Study. *Translational Vision Science & Technology*, *10*(3). https://doi.org/10.1167/TVST.10.3.20

Montolio, F. G. J., Wesselink, C., Gordijn, M., & Jansonius, N. M. (2012a). Factors that influence standard automated perimetry test results in glaucoma: test reliability, technician experience,

time of day, and season. *Investigative Ophthalmology & Visual Science*, *53*(11), 7010–7017. https://doi.org/10.1167/IOVS.12-10268

Montolio, F. G. J., Wesselink, C., Gordijn, M., & Jansonius, N. M. (2012b). Factors That Influence Standard Automated Perimetry Test Results in Glaucoma: Test Reliability, Technician Experience, Time of Day, and Season. *Investigative Ophthalmology & Visual Science*, *53*(11), 7010–7017. https://doi.org/10.1167/IOVS.12-10268

Negi, R., Kalivemula, M., Bisht, K., Bhate, M., Sachdeva, V., & Bharadwaj, S. R. (2024a). Diagnostic accuracy of a modularized, virtual-reality-based automated pupillometer for detection of relative afferent pupillary defect in unilateral optic neuropathies. *Frontiers in Ophthalmology*, *4*. https://doi.org/10.3389/FOPHT.2024.1396511

Negi, R., Kalivemula, M., Bisht, K., Bhate, M., Sachdeva, V., & Bharadwaj, S. R. (2024b). Diagnostic accuracy of a modularized, virtual-reality-based automated pupillometer for detection of relative afferent pupillary defect in unilateral optic neuropathies. *Frontiers in Ophthalmology*, *4*, 1396511. https://doi.org/10.3389/FOPHT.2024.1396511/TEXT

Negi, R., Raviselvan, M., Yarravarapu, D., Chillakala, K., Durai, C. V. R., Baskar, J., Jain, A., Bisht, K., Bhate, M., & Bharadwaj, S. R. (2023). Effect of Light Intensity on the Relative Afferent Pupillary Defect in Unilateral Neuro-ophthalmic Pathology. *Optometry and Vision Science*, *100*(9), 614–624. https://doi.org/10.1097/OPX.0000000000002061

Neuro Ophthalmology | IECRC Singapore. (n.d.). Retrieved June 26, 2026, from https://eyecataractretina.com/eye-conditions/neuro-ophthalmology

Park, C. H., Kim, D., Park, S. W., Park, G., & Heo, H. (2026a). Development and pilot evaluation of a binocular virtual reality Headset-Based pupillometer for quantitative assessment of pupillary light reflexes. *Scientific Reports 2026 16:1*, *16*(1), 1352-. https://doi.org/10.1038/s41598-025-29953-9

Park, C. H., Kim, D., Park, S. W., Park, G., & Heo, H. (2026b). Development and pilot evaluation of a binocular virtual reality Headset-Based pupillometer for quantitative assessment of pupillary light reflexes. *Scientific Reports 2026 16:1*, *16*(1), 1352-. https://doi.org/10.1038/s41598-025-29953-9

Park, C. H., Kim, D., Park, S. W., Park, G., & Heo, H. (2026c). Development and pilot evaluation of a binocular virtual reality Headset-Based pupillometer for quantitative assessment of pupillary light reflexes. *Scientific Reports 2026 16:1*, *16*(1), 1352-. https://doi.org/10.1038/s41598-025-29953-9

(PDF) Aspects of a perimetric learning index Aristeidis Chandrinos Doctor of Philosophy. (n.d.). Retrieved June 26, 2026, from https://www.researchgate.net/publication/338597167_Aspects_of_a_perimetric_learning_index_Aristeidis_Chandrinos_Doctor_of_Philosophy

Philibert, M., & Milea, D. (2024). Basics, benefits, and pitfalls of pupillometers assessing visual function. *Eye (London, England)*, *38*(12), 2415–2421. https://doi.org/10.1038/S41433-024-03151-9

Phu, J., & Kalloniatis, M. (2021). Static automated perimetry using a new head-mounted virtual reality platform, Virtual Field, compared with the Humphrey Field Analyzer in glaucoma and optic nerve disease. *Investigative Ophthalmology & Visual Science*, *62*(8), 3364–3364.

Purves, D., Augustine, G. J., Fitzpatrick, D., Katz, L. C., LaMantia, A.-S., McNamara, J. O., & Williams, S. M. (2001). Neuroscience. *Sinauer Associates, Inc.*, 1–2. https://www.ncbi.nlm.nih.gov/books/NBK10799/

Razeghinejad, R., Gonzalez-Garcia, A., Myers, J. S., & Jay Katz, L. (2020). *Preliminary Report on a Novel Virtual Reality Perimeter Compared With Standard Automated Perimetry*. https://doi.org/10.1097/IJG.0000000000001670

Roy, A. K., Shastry, R., & Rao, A. (1990). Visual Fields. *Ophthalmic Diagnostics: Technology, Techniques, and Clinical Applications*, 243–254. https://doi.org/10.1007/978-981-97-0138-4_21

Ruddy, J., Asuncion, R. M. D., & Cardenas, A. C. (2024). Hemianopsia. *StatPearls*. https://www.ncbi.nlm.nih.gov/books/NBK562262/

Ruia, S., & Tripathy, K. (2025). Humphrey Visual Field. *StatPearls*. https://www.ncbi.nlm.nih.gov/books/NBK585112/

Sander, L., Oommen, G., Brophy, C., Bohus-Roper, S., Chiaro, G., Bremner, F., & Iodice, V. (2025). Validation of Monocular Pupillometry in Healthy Controls and Patients With Autonomic Dysfunction: Pupillary Biomarkers for Autonomic Failure. *European Journal of Neurology*, *32*(8), e70320. https://doi.org/10.1111/ENE.70320

Selvan, K., Mina, M., Abdelmeguid, H., Gulsha, M., Vincent, A., & Sarhan, A. (2023). Virtual reality headsets for perimetry testing: a systematic review. *Eye*, *38*(6), 1041. https://doi.org/10.1038/S41433-023-02843-Y

Selvan, K., Mina, M., Abdelmeguid, H., Gulsha, M., Vincent, A., & Sarhan, A. (2024). Virtual reality headsets for perimetry testing: a systematic review. *Eye (London, England)*, *38*(6), 1041–1064. https://doi.org/10.1038/S41433-023-02843-Y

Smith, J., Flower, O., Tracey, A., & Johnson, P. (2020). A comparison of manual pupil examination versus an automated pupillometer in a specialised neurosciences intensive care unit. *Australian Critical Care : Official Journal of the Confederation of Australian Critical Care Nurses*, *33*(2), 162–166. https://doi.org/10.1016/J.AUCC.2019.04.005

Splittgerber, Ryan., & Snell, R. S. . (2019). *Snell's Clinical Neuroanatomy, 8e*. Lippincott Williams & Wilkins, a Wolters Kluwer business.

Steinhauer, S. R., Bradley, M. M., Siegle, G. J., Roecklein, K. A., & Dix, A. (2022). Publication guidelines and recommendations for pupillary measurement in psychophysiological studies. *Psychophysiology*, *59*(4). https://doi.org/10.1111/PSYP.14035

Stewart, W. C., & Hunt, H. H. (1993). Threshold variation in automated perimetry. *Survey of Ophthalmology*, *37*(5), 353–361. https://doi.org/10.1016/0039-6257(93)90065-F

Tan, J. C. K., Yohannan, J., Ramulu, P. Y., Kalloniatis, M., Crabb, D. P., Crowston, J., & Phu, J. (2025). Visual field testing in glaucoma using the Swedish Interactive Thresholding Algorithm (SITA). *Survey of Ophthalmology*, *70*(1), 141–152. https://doi.org/10.1016/j.survophthal.2024.09.005

Tan, K.-L., Fhun, L.-C., Yaakub, M., Chong, M.-F., & Liza-Sharmini, A. T. (2017). Anxiety and Visual Field Assessment Reliability in Glaucoma Patients. *Asian Journal of Medicine and Health*, *7*(3), 1–8. https://doi.org/10.9734/AJMAH/2017/36396

Tharp, M., Cantor, L., Jacobs, B., Haider, K., & Lind, J. (2023). Assessing Heru, Inc. Virtual Reality Technology for Application in Visual Field Diagnostic Testing. *Investigative Ophthalmology & Visual Science*, *64*(8), 5523–5523.

Tsapakis, S., Papaconstantinou, D., Diagourtas, A., Kandarakis, S., Droutsas, K., Andreanos, K., & Brouzas, D. (2018). Home-based visual field test for glaucoma screening comparison with Humphrey perimeter. *Clinical Ophthalmology (Auckland, N.Z.)*, *12*, 2597. https://doi.org/10.2147/OPTH.S187832

Wall, M. (2021). Perimetry and visual field defects. *Handbook of Clinical Neurology*, *178*, 51–77. https://doi.org/10.1016/B978-0-12-821377-3.00003-9

Wild, J. M., Searle, A. E. T., Dengler-Harles, M., & O'Neill, E. C. (1991). Long-term follow-up of baseline learning and fatigue effects in the automated perimetry of glaucoma and ocular hypertensive patients. *Acta Ophthalmologica*, *69*(2), 210–216. https://doi.org/10.1111/J.1755-3768.1991.TB02713.X

Yotharak, P. (2012). *Correlation between clinical grading and quantification by neutral density filter of relative afferent pupillary defect (RAPD)*. Yotharak, P., & Aui-Aree, N. (2012). Correlation between clinical grading and quantification by neutral density filter of relative afferent pupillary defect (RAPD). Journal of the Medical Association of Thailand = Chotmaihet thangphaet, 95 Suppl 4, S92–S95.