



OPEN ACCESS

EDITED BY

Huaizhou Wang,
Capital Medical University, China

REVIEWED BY

Mohd Izzuddin Hairol,
Universiti Kebangsaan Malaysia, Malaysia
Ward Nieboer,
University Medical Center Groningen,
Netherlands

*CORRESPONDENCE

Alexander Sverstad
✉ alsver@ous-hf.no

RECEIVED 28 May 2025

ACCEPTED 23 September 2025

PUBLISHED 16 October 2025

CITATION

Sverstad A, Helland-Hansen BA,
Kristianslund O, Kolko M, Larsen SE and
Petrovski G (2025) Eye-tracking biomarkers
for glaucoma based on saccadic reaction
time: a controlled clinical study.
Front. Ophthalmol. 5:1636911.
doi: 10.3389/fopht.2025.1636911

COPYRIGHT

© 2025 Sverstad, Helland-Hansen,
Kristianslund, Kolko, Larsen and Petrovski. This
is an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Eye-tracking biomarkers for glaucoma based on saccadic reaction time: a controlled clinical study

Alexander Sverstad ^{1,2*}, Bjørn André Helland-Hansen ¹,
Olav Kristianslund ³, Miriam Kolko ^{4,5},
Stig Einride Larsen ⁶ and Goran Petrovski ¹

¹Centre for Eye Research and Innovative Diagnostics, Oslo University Hospital (OUS), Ullevål, Oslo, Norway, ²Department of Ophthalmology, Vestfold Hospital Trust, Tønsberg, Norway, ³Department of Ophthalmology, Oslo University Hospital and Institute of Clinical Medicine, University of Oslo, Oslo University Hospital (OUS), Ullevål, Oslo, Norway, ⁴Department of Drug Design and Pharmacology, University of Copenhagen, Copenhagen, Denmark, ⁵Department of Ophthalmology, Rigshospitalet, Glostrup, Denmark, ⁶Meddoc Research, Skjetten, Norway

Purpose: Evaluate the validity and reliability of saccadic reaction time (SRT)-based variables obtained using the novel eye-tracking device Bulbicam (BCAM) in differentiating early-to-moderate glaucoma (GLA) from healthy controls (HCs) and to identify potential biomarkers for GLA.

Methods: A controlled clinical study was conducted, involving 18 GLA-patients, and 18 age-matched HCs. Participants underwent BCAM's visual field (VF) test, which measures SRT at 58 symmetrically arranged locations with 6° spacing. Variables were analysed for group differences, within- and between-patient repeatability, and stability. To evaluate their potential as biomarkers, VF locations were aggregated into clusters, quadrants, hemifields, and whole VF analyses.

Results: Significant SRT differences ($p \leq 0.05$) were observed between GLA and HC in 44 of 58 locations in the worst eye and 42 of 58 in the best eye. Eight out of ten clusters met the criteria for BCAM biomarkers having significant group differences, sufficient within- and between-patient repeatability, and adequate stability. All quadrants demonstrated excellent stability and repeatability thereby qualifying as biomarkers. Hemifield SRTs were reliable, however, the absolute difference between hemifields showed poor within-participant repeatability. The mean and standard deviation of SRT for the whole VF were identified as significant biomarkers with excellent stability.

Conclusions: The majority of SRT variables are capable of differentiate glaucomatous eyes from HC while maintaining sufficient reliability and stability for clinical application. 19 of 22 BCAM VF test variables were found to be potential GLA-biomarkers.

Clinical Trial Registration: <https://clinicaltrials.gov/>, identifier NCT05449041.

KEYWORDS

glaucoma, visual field test, eye movement perimetry, saccadic reaction time, reliability, biomarkers, stability index, agreement index

Introduction

Glaucoma (GLA) is a debilitating disease characterised by the loss of retinal ganglion cells (RGCs) and their axons, which may ultimately lead to blindness (1–3). The asymptomatic nature and time-consuming assessment of this disease pose significant challenges for timely diagnosis and management. Despite advancements in technology such as optical coherence tomography (OCT) and standard automated perimetry (SAP), uncertainty in diagnosing and monitoring disease progression persists (4). SAP, the most widely used functional test (5), fails to detect damage until roughly 30% of retinal ganglion cells are lost (6–8), and its variability further complicates clinical decision-making (9–11). These limitations highlight the need for innovative diagnostic strategies capable of detecting GLA earlier and with greater reliability.

Eye-tracking technology has been widely used to study ocular motor control and to investigate the impact of neurological and ophthalmological diseases on eye movement behaviour (12, 13). In GLA, the influence on saccadic reaction time (SRT) has been well documented. Kanjee et al. were the first to explore this in 2012 (14), demonstrating prolonged saccadic latencies across all disease stages, and numerous subsequent studies have confirmed these findings (15–20), including analyses in specific GLA subgroups (21). Compared to SAP, SRT-based perimetry has showed comparable clinical applicability (17). A key distinction between SAP and SRT-based perimetry lies in the mode of patient interaction. SAP relies on subjective button-press responses and sustained, stable fixation, which are vulnerable to inattention, fatigue and response bias, contributing to variability. In contrast, SRT-based perimetry exploits reflexive eye movements that are rapid, and closely tied to visual processing, thereby reducing cognitive load and improving engagement. While SRT is known to be significantly affected by factors such as age and stimulus characteristics, it appears to show small variation with respect to ethnicity, sex, or the presence of cataract (22–24). Importantly, some studies have highlighted the potential of SRT-based perimetry as a promising method for early GLA detection, with evidence showing its ability to detect decreased VF responsiveness in regions that appear normal on SAP (25–27).

Despite this encouraging evidence, SRT-based eye movement perimetry (EMP) has yet to be adopted in routine clinical practice. Barriers include the lack of standardisation, and formal validation across platforms, as well as practical constraints of some existing systems, which can involve more extensive setup procedures and longer testing times. Furthermore, most published studies have prioritized demonstrating differences between GLA and HCs, but relatively little attention has been paid to measurement reliability and stability. To date, only one study, conducted by Pel et al. in 2013, has specifically addressed the validity and repeatability of SRT-based perimetry, limited to a healthy population, demonstrating low variability in SRT across three repeated measurements (28).

Over time, advancements in eye-tracking technology have led to a range of solutions, from desktop display systems to compact, portable devices resembling virtual reality (VR) goggles. One such

innovation is a device called Bulbicam (BCAM), developed by Bulbitech (Trondheim, Norway). Designed as a point-of-care tool, BCAM integrates a high-precision eye-tracking system with dual displays in a VR-goggle-like format, facilitating in-depth assessment of visual and neurological function through a diverse set of tests. BCAM features a user-friendly interface, allowing for rapid test administration, intuitive result interpretation, and flexible adaptation to various clinical and research settings. Of particular relevance for GLA assessment is BCAM's visual field (VF) test, a perimetric test based on SRTs measured across a symmetric grid of 58 locations with 6° spacing. While perimetry exists in other VR-goggle-like platforms, they are typically based on light sensitivity where the responses are given by eye movements or button presses. BCAM is among the first VR-goggle-like platform to incorporate VF responsiveness in the form of SRT-based perimetry.

As with any new diagnostic tool, the value of SRT-based perimetry depends not only on its ability to distinguish patients from HCs but also on the extent to which its measurements are affected by error. Establishing the validity, reliability, and stability of test variables is essential before they can be considered for clinical adoption (29). Reliability reflects the consistency of measurements and requires repeatability within- and between subjects. In practice, within-subject repeatability is often assessed using Bland-Altman plots, whereas between-subject repeatability is commonly expressed using the intraclass correlation coefficient (ICC). Stability describes the consistency of measurements over multiple time points and can uncover training effects or temporal variability that may influence longitudinal monitoring. Together, these properties determine whether a variable has the quality required to function as a clinically meaningful biomarker or not.

While the literature has consistently demonstrated that SRTs are prolonged in GLA, to our knowledge, no study to date has specifically focused on the reliability and stability of these metrics in a GLA population, nor formally assessed their potential as biomarkers. This gap may, in part, reflect challenges in standardising the criteria required to qualify biomarkers (30). Within our framework, a variable qualifies as a biomarker only if it demonstrates sufficient validity and reliability for a clearly specified diagnostic purpose within a defined test procedure.

The aim of this study was to assess BCAM's VF test in terms of validity, reliability, and to explore the potential of 22 predefined variables as biomarkers for GLA.

Materials and methods

Materials

The study sample comprised 18 patients diagnosed with early-to-moderate open-angle GLA, and 18 age-matched healthy controls (HCs) of both genders, at least 18 years of age, and without any other eye disease or other known serious systemic disease. Three screened candidates were excluded prior to enrolment due to age-related macular degeneration, Parkinson's disease, and epiretinal fibrosis and are not counted among the 18 included patients. Out of

these, six eyes were excluded. One due to previous retinal vein occlusion, and the rest due to no detectable GLA changes or advanced GLA. GLA severity was defined by mean deviation (MD) from SAP. Early GLA was defined as MD ≤ 6 dB, and moderate as MD > 6 dB and ≤ 12 dB (5). Excluded participants included those with best corrected visual acuity (BCVA) worse than 1.0 logMAR in either eye, inability to perform eye movements, abnormal visible part of the eye, pupils not able to respond normally to dilation or contraction (e.g., due to damaged nerves or mechanical damage of the pupil). IOP was not used as an inclusion criterion, as participants were already diagnosed and under routine follow-up. Participants were consecutively recruited from the Department of Ophthalmology, Oslo University Hospital Ullevål and Vestfold Hospital Trust. For each participating patient, an age-matched HC without neurological or ophthalmological disease was included. To maintain statistical power, and allow for independent analysis, each participant's eyes was categorised as 'best' or 'worst' based on MD from SAP, with BCVA used as a secondary criterion if MD was equal in both eyes. In the six cases where only one eye was included, the same eye was used in both best and worst eye analyses. This was also applied to the HCs group to better balance the analyses. In total, 18 eyes were classified as mild (out of which 3 were preperimetric) and 12 as moderate GLA. In the best eye analysis, 10 were mild and 8 moderate, while in the worst eye analysis, 9 were mild and 9 moderate. Demographic and clinical characteristics of the study participants can be found in [Table 1](#).

Ethics

All participants gave written informed consent. The study was approved by the institution data protection officer of Oslo University Hospital and Vestfold Hospital Trust. The study was considered by the Regional Ethics committee to be outside their mandate. The study adhered to the tenets of the Declaration of Helsinki and is registered with [ClinicalTrials.gov](#) (NCT05449041).

Methods

This study was designed as an open, non-randomized, controlled clinical study.

Equipment

Data collection was performed using the BCAM device, which employs video-oculography technology, utilising both dark pupil/bright pupil tracking and corneal reflex techniques at a frequency of 400 frames per second (fps) to capture precise gaze direction data. The device features two liquid crystal displays and an infrared eye-tracking camera, enabling presentation of stimuli to one or both eyes and tracking accordingly, based on the test chosen.

Clinical procedure

Participants were seated comfortably and fitted with a mask designed to maintain the optimal distance between the eyes and the displays, while also blocking external light contamination. To further minimise light contamination, BCAM examinations were performed in a dimly lit room. Background noise was kept to a minimum. The mask was magnetically secured to the BCAM, which was positioned on a desk stand and adjusted to ensure participants could comfortably maintain a stable head position. Each participant's interpupillary distance (IPD) and refraction for distance was entered into the BulbiHub software, which automatically calculated the appropriate refraction for the BCAM glasses.

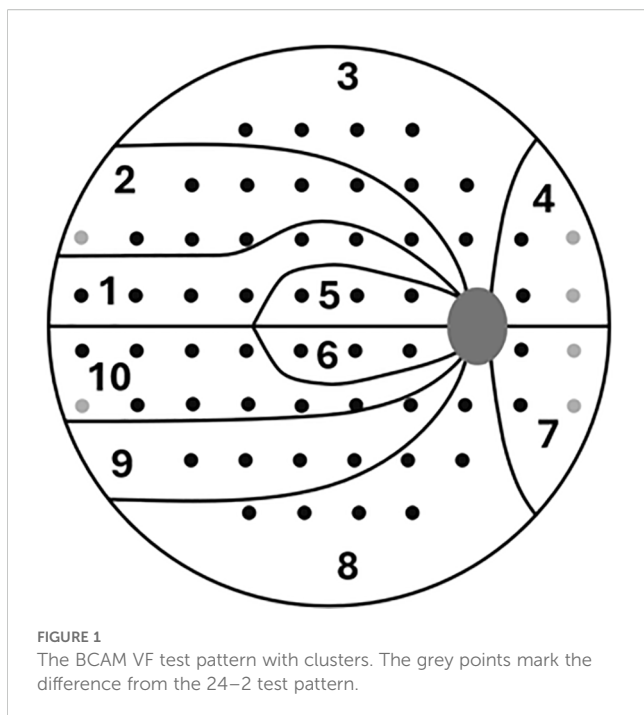
Participants completed a white-on-white, eye movement-based perimetry test, designated as the "VF test". This test employs a grid pattern similar to the commonly used 24-2 layout with 6° spacing. In contrast to the 24-2 layout, it includes two additional peripheral nasal locations and four additional peripheral temporal locations, making a total of 60 test locations. For the analysis, locations at 15° temporal, $\pm 3^\circ$ vertically were excluded as these locations correspond to the blind spot in most individuals (31) ([Figure 1](#)).

The stimulus was presented using the overlap paradigm, with a size of 0.43° (equivalent to Goldman size III) and flicker of 10 Hz, against a background luminance of 10 cd/m². For each test location, the stimulus brightness logarithmically increased from 10cd/m² to a peak of 262 cd/m² over 3 seconds. A green circle, with a size of 0.66°, served as a fixation target; participants were instructed to maintain their gaze on this target until detecting the flickering white stimulus. Upon noticing the stimulus, they were instructed to immediately shift their gaze and fixate on it.

The SRT was measured as the time between the first frame displaying the stimulus, to the first frame in which the participants' gaze deviated outside the fixation target. For an SRT to be accepted,

TABLE 1 Demographic and clinical characteristics of study participants.

Factor/ variables		Glaucoma patients (GLA)	Heathy controls (HC)
Demographic factors	Sex (F/M)	10/8	13/5
	Age (years)	71.7 (58.9 - 84.5)	71.5 (58.9 - 83.7)
	Disease duration (years)	7.0 (0.5 - 20.9)	–
Clinical characteristics	MD	5.7 (-1.3 - 10.6)	2.0 (-0.7 - 5.6)
	IOP (mmHg)	13.9 (9.5 - 22.0)	13.6 (7.0 - 22.0)
	BCVA (logMAR)	0.0 (-0.2 - 0.2)	0.0 (-0.4 - 0.2)



the participant's fixation had to remain within the green circle for a minimum of 120 frames and subsequently on the stimulus for 25 frames. If this criterion was not met, another trial was initiated.

BCAM assessments were completed over 1–3 days depending on participant convenience. Each BCAM VF test took approximately 2 minutes per eye, and participants were encouraged to take breaks as needed. GLA patients completed a total of six test repetitions, while HC completed two. Calibration was done automatically by the device's software. All recordings were done by the same operator (AS). Participants received standardised instructions in Norwegian. To minimize the learning effect, every participant completed a minimum of two practice tests before formal data collection.

The variables in this study were SRTs measured at 58 predefined VF locations with the Bulbicam test (Figure 1). From these, we derived aggregated measures to better evaluate regional susceptibility to glaucomatous damage, enhance the clinical interpretability of spatial patterns in functional loss, and assess how reliability changes with different levels of spatial aggregation.

- Cluster 1 to 10 (Figure 1): Groups of anatomically and functionally related VF points, arranged similar to that used in the EyeSuite software (Haag-Streit Inc., Köniz, Switzerland) (32). These clusters approximate the distribution of RNFL bundles, an established approach in glaucoma evaluation.

- VF quadrants; superonasal (SN), superotemporal (ST), inferonasal (IN), and inferotemporal (IT): Included as an exploratory segmentation to examine whether broader VF divisions reveal differences between GLA and HCs.

- VF halves; superior, inferior, temporal and nasal: The superior-inferior split reflect the typical asymmetry of glaucomatous damage across the horizontal meridian, whereas the nasal-temporal split was included exploratively to assess whether additional asymmetries could be captured.

- Absolute difference of opposing hemifields (superior-inferior, temporal-nasal), included to highlight intra-eye asymmetry.
- Mean and SD of all the VF points was calculated to highlight the overall loss of responsiveness and its variability within the field.

Statistical analysis

All statistical analysis was performed separately for the best and worst eyes.

The power analysis:

The primary purpose of this study was to validate the Bulbicam VF test and identify biomarkers for use in GLA patients. In such studies, it is crucial to minimise false positive biomarker identification while avoiding the oversight of important biomarkers. Thus, the clinically relevant difference (CRD) between patients and HCs was set to 2 standard deviations (SD). With a significance level of 5% ($\alpha=0.05$), and a power of 90% ($\beta=0.90$), a sample size of 12 patients and 12 HCs was required. Validity verification also includes documenting reliability and stability. For this purpose, a slightly larger sample size was considered appropriate. If the CRD was set to 1.5 SD with a corresponding significance level and power, the required number of patients and HCs increased to 16 in each.

Validation: The assumed continuously distributed variables were expressed as mean values with 95% confidence interval (CI). As an index of dispersion, SD or standard error (SE) were provided. All tests were performed two-tailed with a significance level of 5%. Analysis of Variance (ANOVA) and Receiver Operating Characteristic (ROC) analysis were used for group comparisons.

Repeatability: Let SD_w and SD_b denote the SD within and between participants, respectively, and $M1$ and $M2$ represent measurement 1 and 2. The Agreement Index (AI) derived from the Bland-Altman model, was used as a measure of repeatability within participants, defined as $AI = 1 - \frac{2SD_w}{Mean\ of\ M1\ and\ M2}$ (33). Intraclass Correlation Coefficient version 3.1 (ICC) was used as measure of repeatability between participants. ICC values were calculated with the 2-way mixed effects absolute-agreement model

(34), where $ICC = \frac{O_b'^2}{O_b'^2 + O_w'^2}$. This value represents the proportion

of total variance in SRT measurements that is attributable to true differences between participants rather than measurement inconsistency within participants.

Reliability: A variable is considered reliable when both AI and ICC are ≥ 0.50 , in conjunction with sufficient stability.

Stability: Stability was quantified as the Stability Index (SI), defined as $SI = 1 - SD_w/SD_b$, where SD_w and SD_b represent the SD within and between patients, respectively (35). The stability of a variable is considered acceptable when $SI \geq 0.14$. Further details regarding the statistical approach can be found in the paper by Dalbro et al., 2025 (35).

Biomarker: A clinically useful biomarker must be valid and reliable. Validity is shown by group discrimination (ANOVA $p < 0.05$ and/or ROC AUC with 95% CI lower bound > 0.50). Reliability is defined by ICC (between participant repeatability),

AI (within participant repeatability), and SI (temporal stability over several measurements). A variable is a population-level biomarker if validity criteria are met and $ICC \geq 0.50$ and $SI \geq 0.14$. An individual-level biomarker if validity criteria are met and $AI \geq 0.50$ and $SI \geq 0.14$. Variables meeting all four criteria (validity, AI, ICC and SI) qualify at both levels. All thresholds were pre-specified.

Statistical analyses were performed using SAS software version 9.4 (SAS Institute Inc., Cary, NC, USA).

Results

Visual field point analysis

In the worst eye, 44 out of 58 VF locations showed a significant difference compared to HC. No significant differences ($p > 0.05$) were

detected at 14 locations (Figure 2A). In the best eye, significant differences were observed in 42 locations, and 16 showed no significant difference (Figure 2B).

Cluster analysis

SRT was significantly greater in GLA patients compared to HC in the ANOVA analysis, with significant differentiation confirmed by ROC analysis in 9 of the 10 VF clusters in the worst eye (Table 2, Figures 3A, B). For the best eye, significant differences between patients and HC were detected in 8 of the 10 clusters. ROC-analysis confirmed significant differentiation when the lower limit of the 95% confidence interval (CI) for AUC exceeded 0.50. No significant difference between patients and HC was detected in cluster 6 for either eye, or in cluster 5 for the best eye.

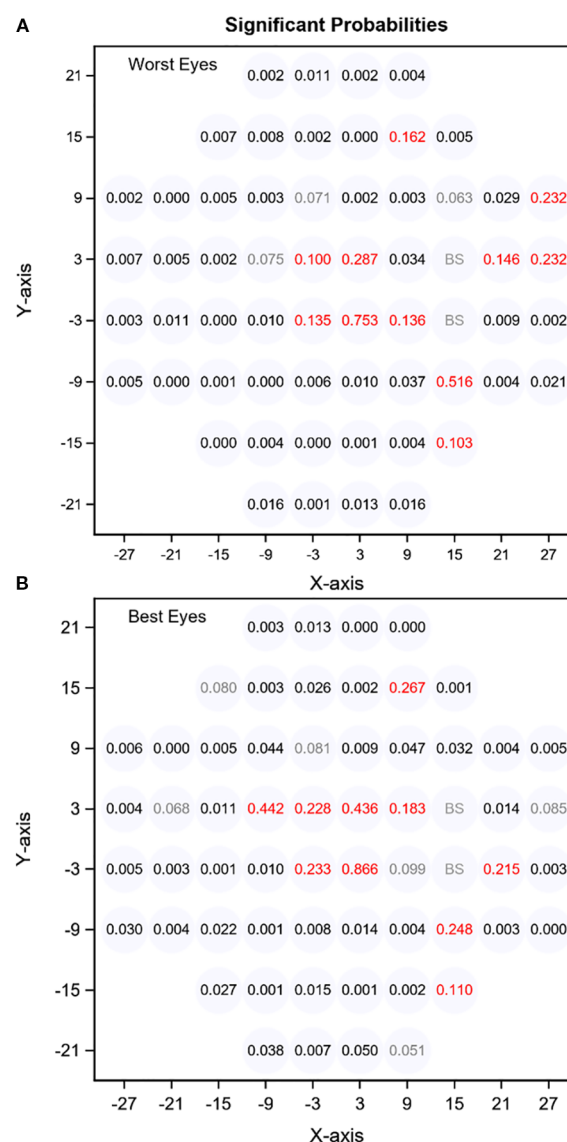


FIGURE 2
ANOVA probability plot of all VF points in the worst (A) and best (B) eyes.

TABLE 2 Validation of the BCAM VF test. Comparison of patients with glaucoma and age-matched HCs including ROC analysis.

Eye	Saccadic reaction time (SRT)		GLA patients	Health controls	GLA-HC mean (95% CI)	ROC analysis	
						AUC	95% CI
Worst Eye	Cluster	1	772.6 (51.7)	498.6 (51.7)	274.0 (128.4-419.7)	0.78	0.67 - 0.89
		2	879.5 (46.6)	526.7 (46.6)	352.7 (221.3 - 484.1)	0.87	0.80 - 0.95
		3	799.9 (36.4)	553.2 (36.4)	246.7 (144.0 - 349.4)	0.81	0.71 - 0.91
		4	617.1 (41.2)	494.9 (41.2)	122.3 (6.0 - 238.5)	0.66	0.53 - 0.79
		5	749.4 (71.6)	531.1 (71.6)	218.3 (16.5 - 420.1)	0.64	0.51 - 0.77
		6	606.2 (47.3)	522.6 (47.3)	83.6 (-49.9 - 217.0)	0.58	0.45 - 0.72
		7	784.8 (47.2)	537.6 (47.2)	247.2 (114.0 - 380.5)	0.72	0.60 - 0.84
		8	828.3 (39.8)	588.9 (39.8)	239.4 (127.3 - 351.5)	0.76	0.65 - 0.87
		9	911.7 (49.6)	553.6 (49.6)	358.1 (218.2 - 498.0)	0.81	0.71 - 0.91
		10	869.6 (57.3)	502.8 (57.3)	366.8 (205.3 - 528.3)	0.84	0.75 - 0.93
	Quadrant	Superotemporal	760.1 (33.4)	521.9 (33.4)	238.3 (144.1 - 332.4)	0.82	0.73 - 0.91
		Inferotemporal	736.6 (29.8)	544.4 (29.8)	192.2 (108.0 - 276.4)	0.76	0.65 - 0.87
		Inferonasal	907.0 (52.9)	538.5 (52.9)	368.5 (219.3 - 517.6)	0.84	0.75 - 0.93
		Superonasal	834.9 (46.2)	528.3 (46.2)	306.6 (176.5 - 436.8)	0.84	0.75 - 0.93
	Half	Inferior	828.5 (37.2)	547.3 (37.2)	281.2 (176.2 - 386.2)	0.84	0.75 - 0.93
		Superior	794.9 (34.8)	529.9 (34.8)	265.0 (166.8 - 363.3)	0.86	0.78 - 0.94
		Abs diff/In-Su/	144.7 (16.4)	40.3 (16.4)	104.3 (58.0 - 150.6)	0.75	0.63 - 0.87
		Nasal	894.1 (48.2)	547.5 (48.2)	346.6 (210.7 - 482.5)	0.87	0.79 - 0.95
		Temporal	732.7 (26.4)	528.8 (26.4)	203.9 (129.5 - 278.4)	0.83	0.73 - 0.92
		Abs diff/Na-Te/	212.4 (27.7)	63.5 (27.7)	148.9 (70.8 - 227.0)	0.80	0.71 - 0.90
	Whole	Mean	811.5 (34.1)	538.5 (34.1)	273.0 (176.8 - 369.2)	0.87	0.79 - 0.95
		SD	491.8 (26.2)	243.6 (26.2)	248.2 (174.3 - 322.1)	0.87	0.79 - 0.95
Best Eye	Cluster	1	670.5 (46.1)	471.1 (46.1)	199.4 (69.4 - 329.4)	0.75	0.63 - 0.86
		2	741.3 (34.7)	499.9 (34.7)	241.4 (143.5 - 339.4)	0.78	0.67 - 0.89
		3	803.8 (41.0)	509.3 (41.0)	294.5 (178.9 - 410.1)	0.82	0.72 - 0.92
		4	624.3 (39.9)	434.7 (39.9)	189.6 (77.2 - 302.0)	0.67	0.54 - 0.80
		5	650.9 (48.4)	548.4 (48.4)	102.5 (-34.1 - 239.1)	0.64	0.51 - 0.76
		6	520.2 (25.1)	475.1 (25.1)	45.1 (-25.8 - 116.0)	0.53	0.39 - 0.66
		7	669.3 (43.3)	508.0 (43.3)	161.3 (39.3 - 283.3)	0.65	0.52 - 0.78
		8	783.4 (46.8)	587.1 (46.8)	196.4 (64.5 - 328.2)	0.64	0.51 - 0.77
		9	807.4 (48.4)	522.8 (48.4)	284.5 (148.0 - 421.0)	0.70	0.58 - 0.83
		10	713.3 (38.0)	468.7 (38.0)	244.6 (137.4 - 351.8)	0.77	0.66 - 0.88
	Quadrant	Superotemporal	746.4 (30.8)	510.5 (30.8)	235.9 (149.0 - 322.9)	0.84	0.76 - 0.93
		Inferotemporal	698.9 (34.2)	536.2 (34.2)	162.6 (66.1 - 259.2)	0.69	0.57 - 0.82
		Inferonasal	743.4 (37.2)	496.4 (37.2)	247.0 (142.2 - 351.8)	0.73	0.61 - 0.85
		Superonasal	703.5 (34.9)	485.4 (34.9)	218.0 (119.5 - 316.6)	0.79	0.68 - 0.89
	Half	Inferior	730.3 (33.7)	517.6 (33.7)	212.7 (117.6 - 307.8)	0.73	0.61 - 0.85

(Continued)

TABLE 2 Continued

Eye	Saccadic reaction time (SRT)		GLA patients	Health controls	GLA-HC mean (95% CI)	ROC analysis	
						AUC	95% CI
		Superior	724.9 (29.2)	499.2 (29.2)	225.7 (143.3 - 308.1)	0.83	0.74 - 0.92
		Abs diff/In-Su/	130.9 (14.5)	38.3 (14.5)	92.5 (51.6 - 133.4)	0.77	0.66 - 0.88
		Nasal	744.0 (35.4)	504.2 (35.4)	239.8 (139.8 - 339.7)	0.76	0.64 - 0.87
		Temporal	710.8 (27.3)	513.4 (27.3)	197.4 (120.5 - 274.3)	0.80	0.70 - 0.90
		Abs diff/Na-Te/	136.3 (15.4)	66.6 (15.4)	69.7 (26.4 - 113.1)	0.67	0.55 - 0.80
	Whole	Mean	727.8 (29.6)	508.5 (29.6)	219.2 (135.9 - 302.6)	0.79	0.69 - 0.90
		SD	430.5 (26.5)	215.9 (26.5)	214.6 (139.7 - 289.4)	0.80	0.70 - 0.91

The results are expressed by mean values, SE and 95% confidence intervals (CI).

Between-patient repeatability ($ICC \geq 0.5$) was not achieved for cluster 4 and 6 in the worst eye, nor for clusters 5 and 6 in the best eye (Table 3). However, the remaining clusters were found to be repeatable between patients. Within-patient repeatability was achieved for clusters 3 and 9 in the worst eye (Figure 4B), and for cluster 3 and 10 in the best eye.

The stability of SRT was sufficient across all 10 clusters for both the worst and the best eye (Table 4). In the worst eye, stability was classified as “Excellent” in 4 clusters, “Very Good” in 2 clusters, and “Good” in 4 clusters (Figures 5A, B). Similarly, for the best eye, 3 clusters were classified as “Excellent”, 2 as “Very Good” and 5 as “Good”.

In the worst eye, clusters 3 and 9 qualify as BCAM biomarkers for GLA (Table 5). Clusters 1, 2, 5, 7, 8 and 10 are potential biomarkers, but require re-testing on the same patient, as within-participant repeatability was classified as “poor”, with sufficient stability. Clusters 4 and 6 are not recommended as BCAM biomarkers for GLA.

Quadrant analysis

SRT was significantly greater among patients compared to HC for all VF quadrants in both the worst and the best eyes (Table 2). SRT was found to be repeatable between and within patients for all quadrants.

Both the IN and the SN quadrants demonstrated high reliability with excellent repeatability (Table 3) and stability (Table 4). Stability classifications were “Excellent” for both quadrants in both eyes. ST and IT quadrants showed “Excellent” and “Very Good” stability in the worst eye, and “Good” and “Excellent” stability in the best eye, respectively.

With the exception of the IT quadrant in the best eye, which is classified as a potential biomarker for GLA due to an AI below 0.5, all VF quadrants are classified as BCAM biomarkers for GLA in both eyes.

Hemifield analysis

The SRT in all hemifields, as well as the absolute difference between opposing hemifields, was significantly greater in GLA patients compared to HC for both eyes (Table 2, Figures 3C, D).

All hemifields were found to be repeatable between patients, with the exception of the absolute difference between the temporal and nasal hemifields (Table 3). Within-patient repeatability was satisfactory for all hemifields in both eyes, except for absolute differences (Figures 4C, D).

Stability was found to be sufficient in all hemifields for both eyes (Table 4). In the worst eye, stability was classified as “Excellent” for all hemifields, except for the absolute difference between the superior and inferior hemifields, which was classified as “Very Good” (Figures 5C, D).

SRT for all hemifields qualifies as BCAM biomarkers for GLA. However, the absolute difference between opposing hemifields has potential as a biomarkers but requires re-testing on the same patient, as the repeatability was “poor” ($AI < 0.5$), with stability classified as “Very Good”.

Whole-area analysis

The mean and SD of SRTs across the entire VF were significantly greater in GLA patients compared to HC for both eyes (Table 1). AUC values in ROC analysis exceeded 0.5 with high confidence for both eyes (Figure 3E).

Both variables demonstrated repeatability between- and within-patients for both eyes (Table 3, Figure 4E); however, within-patient repeatability for SD in the best eye was borderline.

Stability was found to be sufficient for both variables in both eyes (Table 4) and was classified as “Excellent” (Figure 5E). Both the mean and SD of SRTs qualify as BCAM biomarkers for GLA in both eyes.

Discussion

This study demonstrates that multiple SRT-based variables measured with the BCAM system are capable of differentiating GLA from HCs, with sufficient reliability, especially when VF points were aggregated into larger regions. These findings support their potential as biomarkers for GLA. Moreover, our results extend prior work by systematically evaluating both validity and reliability of

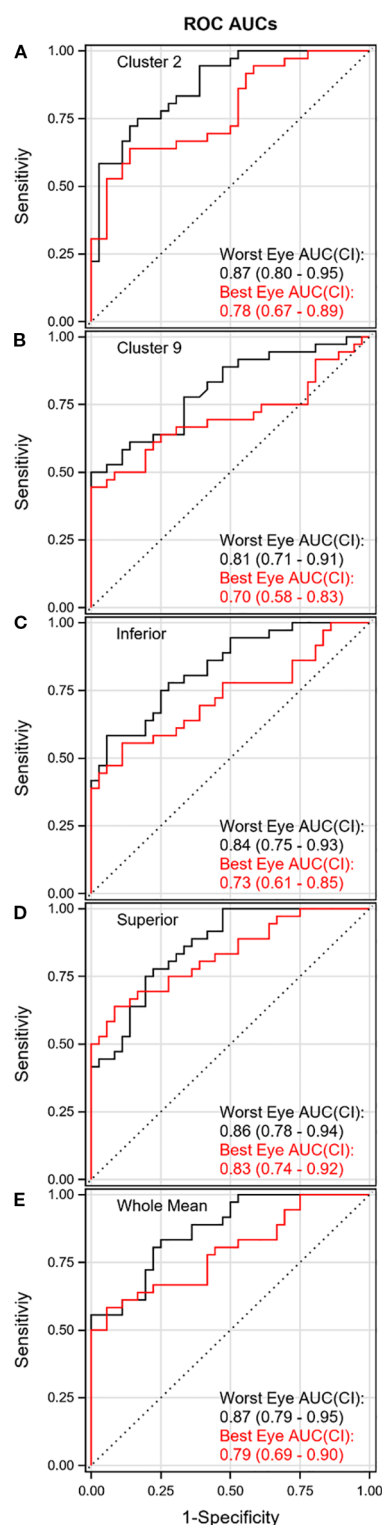


FIGURE 3
ROC curve for worst (black) and best (red) eyes.

SRT-based EMP in a GLA population, an important step for clinical translation.

Prolonged SRT in GLA has been well-documented in the literature, and our findings align with this body of evidence. Mean SRT in our GLA group were 43% (best eye) and 50.7%

(worst eye) longer than in HCs, consistent with prior studies reporting increases ranging from 7.2–54% depending on methodology and disease stage (14–16, 27). The strong discriminatory performance of global mean SRT, even in mild to moderate GLA suggests that SRT may be particularly sensitive to

TABLE 3 Reliability of the BCAM VF test expressed by ICC and AI.

Eye	Items	Variables	Glaucoma patients					Healthy controls				
			M1	M2	M1 – M2	ICC	AI	M1	M2	M1 – M2	ICC	AI
Worst Eye	Cluster	1	711.5	833.7	-122 (-403–159)	0.79	0.31	476.3	520.8	-45 (-142 - 53)	0.25	0.29
		2	854.0	904.9	-51 (-303-202)	0.76	0.41	528.8	524.7	4 (-95 - 103)	0.43	0.40
		3	808.2	791.5	17 (-159-192)	0.85	0.64	542.2	564.2	-22 (-142 - 98)	0.63	0.45
		4	618.5	615.7	3 (-187-192)	0.45	0.05	519.3	470.5	49 (-98 - 196)	0.30	-0.04
		5	836.6	662.3	174 (-212- 561)	0.56	-0.43	467.7	594.5	-127 (-265- 12)	-0.18	-0.18
		6	632.1	580.3	52 (-166- 270)	0.28	-0.27	461.9	583.4	-121 (-285- 42)	0.10	-0.24
		7	803.1	766.5	36.6 (-205-278)	0.70	0.30	544.2	530.9	13 (-119 - 146)	0.10	0.03
		8	817.6	839.0	-21 (-221 - 178)	0.57	0.34	593.6	584.2	9 (-109 - 128)	0.51	0.41
		9	917.4	905.9	11.5 (-261 - 284)	0.90	0.61	537.1	570.1	-33 (-130 - 64)	0.68	0.58
		10	920.5	818.7	102 (-221- 425)	0.88	0.46	498.8	506.7	-8 (-84 - 68)	0.74	0.68
	Quadrant	Superotemporal	742.8	777.5	-35 (-207- 137)	0.84	0.62	511.7	532.0	-20 (-110 - 70)	0.70	0.61
		Inferotemporal	731.4	741.7	-10(-160 - 139)	0.73	0.56	535.5	553.2	-18 (-106 - 71)	0.69	0.62
		Inferonasal	941.6	872.3	69 (-228- 366)	0.96	0.71	526.5	550.6	-24 (-103 - 55)	0.67	0.65
		Superonasal	829.4	840.3	-11 (-265- 243)	0.93	0.66	513.0	543.5	-30 (-118 - 58)	0.50	0.51
	Half	Inferior	847.4	809.7	38 (-164- 239)	0.94	0.75	535.2	559.5	-24 (-104 - 55)	0.74	0.69
		Superior	783.1	806.7	-24 (-209- 162)	0.93	0.75	519.4	540.3	-21 (-103 - 61)	0.80	0.71
		Abs diff/In-Su/	152.5	136.9	16 (-77-108)	0.80	-0.18	35.4	45.3	-10 (-34 - 14)	0.26	-1.17
		Nasal	910.0	878.2	31.9 (-236-299)	0.98	0.83	533.3	561.7	-28 (-113 - 57)	0.61	0.60
		Temporal	724.1	741.4	-17 (-146- 112)	0.82	0.69	520.7	536.8	-16 (-100 - 68)	0.81	0.71
		Abs diff/Na-Te/	215.2	209.6	6 (-149-160)	0.93	0.19	39.9	87.0	-47 (-90 - -4)	-0.14	-2.02
	Whole	Mean	814.6	808.4	6 (-176-189)	0.95	0.80	527.3	549.8	-23 (-101 - 56)	0.80	0.73
		SD	484.0	499.2	-15 (-147-118)	0.91	0.67	193.9	293.3	-99 (-167- -32)	0.62	0.29
Best Eye	Cluster	1	673.3	667.8	6 (-249-260)	0.70	0.31	430.3	512.0	-82 (-164 - 0)	0.14	0.33
		2	724.4	758.2	-34 (-217-149)	0.72	0.45	520.4	479.3	41 (-44 -127)	0.47	0.48
		3	781.9	825.6	-45 (-269-182)	0.86	0.56	485.8	532.7	-47 (-123 - 29)	0.39	0.51
		4	669.0	579.7	89 (-125- 303)	0.55	0.04	449.1	420.4	29 (-56 - 113)	0.03	0.20
		5	662.1	639.6	23 (-199- 244)	0.09	-0.35	511.2	585.6	-74 (-247 - 98)	-0.07	-0.36
		6	534.4	506.0	28 (-102-159)	0.39	0.18	471.7	478.5	-7 (-72 - 59)	0.21	0.49
		7	638.3	700.4	-62 (-293-169)	0.82	0.38	503.8	512.3	-9 (-109 - 91)	0.48	0.41
		8	734.0	832.9	-99 (-349-151)	0.72	0.29	583.9	590.2	-6 (-109 - 96)	0.61	0.55
		9	817.0	797.7	19 (-255- 293)	0.80	0.36	509.7	536.0	-26 (-93 - 41)	0.57	0.65
		10	741.5	685.0	57 (-154- 267)	0.90	0.61	454.5	482.9	-28 (-93 - 36)	0.59	0.63
	Quadrant	Superotemporal	740.3	752.6	-12 (-172- 148)	0.78	0.58	504.3	516.7	-12 (-94 - 69)	0.35	0.46
		Inferotemporal	688.5	709.2	-21 (-203-162)	0.77	0.48	529.4	543.0	-14 (-95 - 67)	0.65	0.63
		Inferonasal	749.6	737.1	13 (-196- 221)	0.91	0.65	485.6	507.2	-22 (-78 - 35)	0.42	0.64
		Superonasal	698.1	708.8	-11 (-204-183)	0.82	0.51	467.7	503.2	-35 (-99 - 28)	0.46	0.60
	Half	Inferior	733.8	726.7	7 (-180-194)	0.95	0.76	509.5	525.7	-16 (-76 - 43)	0.62	0.70

(Continued)

TABLE 3 Continued

Eye	Items	Variables	Glaucoma patients					Healthy controls				
			M1	M2	M1 – M2	ICC	AI	M1	M2	M1 – M2	ICC	AI
		Superior	726.7	723.1	4 (-157- 164)	0.92	0.74	488.5	510.0	-22 (-80 - 36)	0.64	0.71
		Abs(In-Su)	137.8	123.9	14 (-68- 95)	0.54	-0.77	48.6	28.0	21 (0 - 42)	0.10	-1.17
		Nasal	752.0	735.9	16 (-182- 214)	0.96	0.77	491.2	517.1	-26 (-86 - 34)	0.50	0.65
		Temporal	706.8	714.7	-8 (-152- 136)	0.83	0.65	507.8	518.9	-11 (-79 - 57)	0.64	0.67
		Abs (Na-Te)	167.2	105.5	62 (-17- 141)	0.43	-0.83	47.7	85.5	-38 (-72 - -4)	-0.06	-1.20
	Whole	Mean	729.9	725.6	4 (-158-167)	0.96	0.82	499.2	517.9	-19 (-76 - 38)	0.72	0.75
		SD	437.2	423.7	14 (-124 - 151)	0.86	0.50	195.2	236.6	-41 (-111 - 28)	0.24	-0.17

The results are expressed by Least square Mean (LSM) value, SE and 95% CI.

glaucomatous damage. This is in line with previous work showing SRT to be prolonged even in regions with normal light sensitivity (25–27). Elgin also reported prolonged SRTs in preperimetric and moderate GLA, particularly using a kinetic paradigm, suggesting its increased sensitivity to glaucomatous damage. Moreover, by applying a machine learning approach to patterns across multiple SEM variables, the study achieved an AUC of 0.87 for detecting preperimetric GLA (27).

In our study, most VF locations showed significant differentiation between GLA and HCs. Using a similar grid, Meethal et al. (36) found slightly higher mean pointwise AUC of 0.75 (0.05) compared with our findings of 0.67 (0.06) and 0.7 (0.06) for the best and worst eye respectively. This likely reflects their inclusion of advanced glaucoma together with differences in stimulus settings.

Test-retest variability is a well-known limitation of SAP in GLA, where results often fluctuate more than in healthy individuals (11, 37). Pel et al. (28) found low variability of SRT across three measurement series in healthy subjects. In our study, test-retest variability was comparable between GLA and HCs, as reflected by the similar AI classifications. Only 3 of the 22 variables differed between groups, suggesting that, under our protocol, SRT-based measures are not disproportionately susceptible to disease-related variability, consistent with findings from frequency doubling- and motion perimetry (38–40).

Nonetheless, some test-retest variability was present, which may partly be related to stimulus characteristics and fluctuation in fatigue and attention. The overlap paradigm was selected to reduce express saccades, the trade-off may be wider SRT distributions (13, 41, 42). To promote reflexive saccades, flickering and pseudorandom stimuli were applied to enhance salience, yet a proportion of voluntary or predictive saccades likely contributed additional variability (43).

Aggregating single VF locations into larger areas improved the validity and reliability markedly. Although this is an expected consequence of averaging across multiple locations, it still represents a practical way of obtaining more reliable measures, particularly when clusters are organised to reflect the anatomical layout of the RNFL bundles (44).

In our study, only cluster 3 and 9 in the worst eye, and cluster 3 and 10 in the best eye, were reliable and stable. These clusters

correspond to regions commonly affected in early-to-moderate GLA, including the superior arcuate and nasal/inferior paracentral VF (45, 46). The lack of reliability in other clusters likely relates to the relatively few VF points in each cluster and the heterogeneous impact of GLA on the VF responsiveness in the study sample. Aggregating points into VF quadrants and hemifields further improved the validity and reliability, with the nasal regions performing best, consistent with the notion of early glaucomatous damage often affecting on the nasal side of the VF (47–49).

Asymmetrical defects between the superior and inferior hemifields are another hallmark of GLA, related to asymmetric damage of the neuroretinal rim (48). Algorithms such as the glaucoma hemifield test (GHT) in the Humphrey Field Analyzer (HFA) use this principle to help detect early glaucomatous changes. We included a simple variable to highlight such asymmetry - the absolute difference between opposing hemifields. Contrary to our initial expectation, this variable underperformed relative to analysing each hemifield independently. This may reflect relatively symmetric field loss in our study sample, or a global SRT depression in GLA, as suggested previously (15, 25). Mazumdar et al. (15) reported an AUC of 0.78 using a sector-based approach analogous to the GHT. The simpler approach used in our study produced comparable values.

Beyond hemifield asymmetry, early GLA VF loss is often localised and heterogeneous. Variability is typically captured by indices such as the square root of loss variance (Octopus) and pattern standard deviation (HFA). Using the SD of the entire VF plot, we found significantly greater variability in GLA patients compared to HC, with ROC-AUC values of 0.87 for the worst eye and 0.8 for the best eye.

Mean SRT across the VF was the most reliable variable, and produced the strongest discriminatory ability with AUC of 0.79 and 0.87 in the best and worst eye respectively. However, its specificity is limited, as SRT is affected by a range of other diseases (50). For this reason, a function-structure specific approach, such as cluster analysis, are likely to provide a more GLA specific evaluation.

From a practical standpoint, BCAM VF test completed a 24–2 pattern in roughly 2 to 2.5 minutes per eye, comparable to faster SAP strategies (51, 52), and integrates display and eye-tracking in a single unit, simplifying setup relative to earlier EMP systems. These

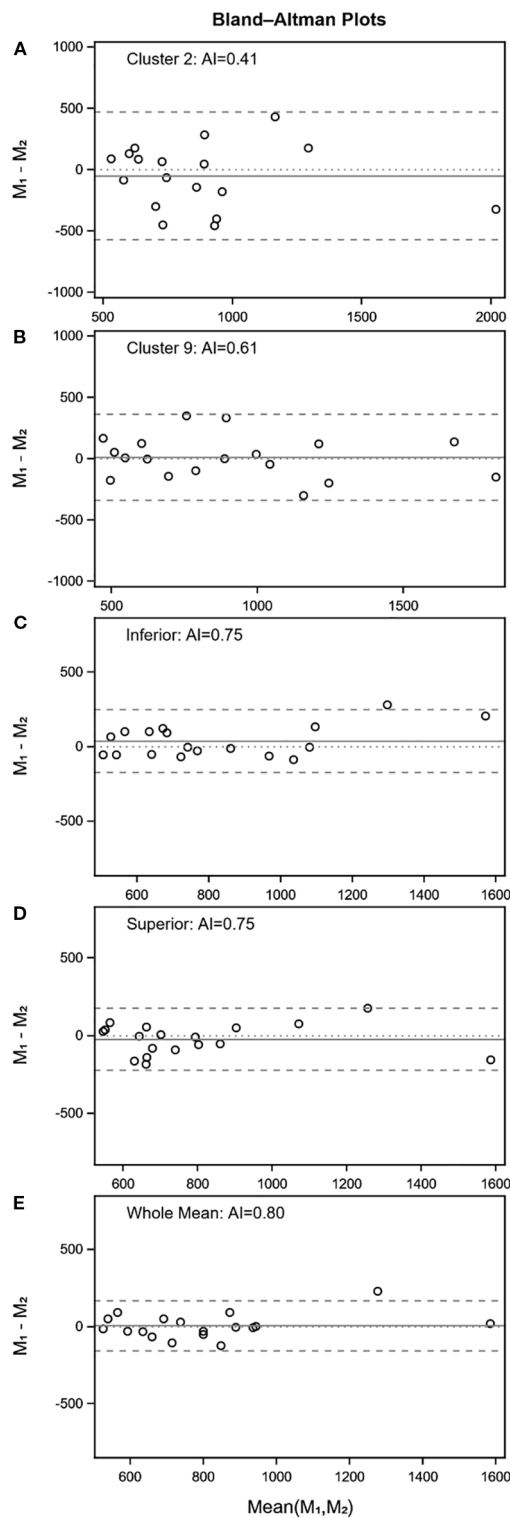


FIGURE 4
Bland-Altman plot with AI (worst eye only) for cluster 2 and 9, inferior and superior hemifield, and for the whole VF.

TABLE 4 Stability of the BCAM VF test.

Eye	Variable		M1	M2	M3	M4	M5	M6	SI (95%CI)	Cla
Worst Eye	Cluster	1	711.5	833.7	896.2	795.3	813.1	814.1	0.50 (0.31-0.69)	VG
		2	854.0	904.9	856.0	847.4	840.0	822.2	0.59 (0.48-0.69)	E
		3	808.2	791.5	739.0	863.3	801.1	864.6	0.43 (0.21-0.65)	G
		4	618.5	615.7	643.8	639.3	617.3	608.2	0.39 (0.23-0.54)	G
		5	836.6	662.3	784.0	693.4	602.5	573.6	0.60 (0.38-0.82)	E
		6	632.1	580.3	571.2	617.4	531.1	588.7	0.52 (0.32-0.73)	VG
		7	803.1	766.5	714.7	722.8	695.2	758.6	0.40 (0.23-0.58)	G
		8	817.6	839.0	840.7	858.1	829.4	842.4	0.35 (0.18-0.53)	G
		9	917.4	905.9	976.1	885.7	892.1	955.3	0.54 (0.42-0.66)	E
		10	920.5	818.7	841.7	910.8	910.9	904.7	0.76 (0.69-0.83)	E
	Quadrante	Superotemporal	742.8	777.5	727.2	756.7	764.9	738.5	0.54 (0.43-0.65)	E
		Inferotemporal	731.4	741.7	735.5	763.9	750.0	745.6	0.43 (0.24-0.62)	G
		Inferonasal	941.6	872.3	894.6	911.9	886.7	941.2	0.75 (0.69-0.81)	E
		Superonasal	829.4	840.3	864.7	848.8	814.8	831.8	0.69 (0.59-0.79)	E
	Half	Inferior	847.4	809.7	828.4	846.6	826.0	847.1	0.73 (0.66-0.79)	E
		Superior	783.1	806.7	807.9	806.0	793.4	782.2	0.70 (0.62-0.79)	E
		Abs diff/In-Su/	152.5	136.9	129.9	134.7	165.5	169.4	0.49 (0.36-0.63)	VG
		Nasal	910.0	878.2	910.6	907.1	874.5	901.5	0.79 (0.74-0.85)	E
		Temporal	724.1	741.4	728.3	746.6	751.0	727.5	0.57 (0.48-0.67)	E
		Abs diff/Na-Te/	215.2	209.6	246.1	285.3	224.1	261.2	0.58 (0.47-0.70)	E
	Whole	Mean	814.6	808.4	818.6	825.9	810.0	814.2	0.77 (0.72-0.83)	E
		SD	484.3	499.2	481.4	516.0	499.7	521.9	0.66 (0.61-0.72)	E
Best Eye	Cluster	1	673.3	667.8	701.1	683.6	670.8	703.5	0.65 (0.51-0.79)	E
		2	724.4	758.2	710.9	705.4	664.3	743.3	0.54 (0.42-0.66)	E
		3	781.9	825.6	777.2	901.8	781.6	847.9	0.46 (0.25-0.67)	VG
		4	669.0	579.7	652.9	654.8	609.4	566.0	0.46 (0.28-0.64)	VG
		5	662.1	639.6	619.1	655.7	608.7	549.5	0.39 (0.14-0.64)	G
		6	534.4	506.0	471.2	537.4	546.4	548.5	0.35 (0.01-0.68)	G
		7	638.3	700.4	659.6	697.6	736.6	583.9	0.41 (0.14-0.67)	G
		8	734.0	832.9	848.7	847.8	809.0	833.3	0.39 (0.24-0.54)	G
		9	817.0	797.7	895.0	739.5	868.7	1011.9	0.44 (0.25-0.64)	G
		10	741.5	685.0	654.4	731.3	682.8	681.0	0.69 (0.61-0.77)	E
	Quadrante	Superotemporal	740.3	752.6	714.3	774.4	747.7	741.6	0.34 (0.18-0.51)	G
		Inferotemporal	688.5	709.2	737.5	709.1	741.7	717.7	0.50 (0.38-0.62)	VG
		Inferonasal	749.6	737.1	717.0	755.3	740.1	784.1	0.67 (0.59-0.76)	E
		Superonasal	698.1	708.8	704.6	709.1	639.6	717.6	0.70 (0.59-0.80)	E
	Half	Inferior	733.8	726.7	731.3	741.5	745.9	759.2	0.70 (0.63-0.77)	E
		Superior	726.7	723.1	716.1	740.9	704.4	730.7	0.68 (0.58-0.77)	E

(Continued)

TABLE 4 Continued

Eye	Variable		M1	M2	M3	M4	M5	M6	SI (95%CI)	Cla
		Abs diff/In-Su/	137.8	123.9	122.2	123.9	137.9	143.0	0.48 (0.33-0.63)	VG
		Nasal	752.0	735.9	728.8	750.0	712.1	770.5	0.75 (0.68-0.82)	E
		Temporal	706.8	714.7	718.5	730.7	739.2	716.0	0.50 (0.38-0.62)	VG
		Abs diff/Na-Te/	167.2	105.5	142.4	178.8	153.9	129.4	0.41 (0.25-0.56)	G
	Whole	Mean	729.9	725.6	723.8	741.1	726.4	743.9	0.72 (0.65-0.80)	E
		SD	437.2	423.7	430.8	421.1	436.6	450.5	0.59 (0.47-0.70)	E

The six measurements are denoted as M1 to M6 and expressed by mean values. Stability is expressed by the Stability Index (SI) with 95% Confidence Intervals (CI). The classification is based on SI. Classification (Cla), E, Excellent; VG, Very Good; G, Good.

features may facilitate clinical use if diagnostic performance is confirmed in broader cohorts.

By engaging natural oculomotor reflexes and increasing retinal image change, SRT-based perimetry may enhance engagement and reduce factors known to compromise reliability in SAP, including inattention, fixation loss, false positive and negative responses, the Troxler fading effect, and Ganzfeld blank-out (53–58).

Our study has several limitations that should be acknowledged. First, the study primarily focused on group-level differences between otherwise healthy GLA patients and controls. While these findings are promising, further research is needed to evaluate the performance of the BCAM VF test across a wider spectrum of patient profiles, including those with comorbidities and a broader range of disease severity. Second, glaucoma diagnoses

TABLE 5 Recommended user area as Bulbicam Visual Field Biomarkers for Glaucoma.

Possible biomarker for glaucoma		ROC	ICC	AI	SI	Recommended user area	
						Worst eye	Best eye
Cluster	1	+	+	–	+	Population + Patient ^{*)}	Population + Patient ^{*)}
	2	+	+	–	+	Population + Patient ^{*)}	Population + Patient ^{*)}
	3	+	+	+	+	Population + Patient	Population + Patient
	4	+	–	–	+	Not recommended	Population + Patient ^{*)}
	5	+	+	–	+	Population + Patient ^{*)}	Not recommended
	6	–	–	–	+	Not recommended	Not recommended
	7	+	+	–	+	Population + Patient ^{*)}	Population + Patient ^{*)}
	8	+	+	–	+	Population + Patient ^{*)}	Population + Patient ^{*)}
	9	+	+	+	+	Population + Patient	Population + Patient ^{*)}
	10	+	+	–	+	Population + Patient ^{*)}	Population + Patient
Quadrant	Superotemporal	+	+	+	+	Population + Patient	Population + Patient
	Inferotemporal	+	+	+	+	Population + Patient	Population + Patient ^{*)}
	Inferonasal	+	+	+	+	Population + Patient	Population + Patient
	Superonasal	+	+	+	+	Population + Patient	Population + Patient
Half	Inferior	+	+	+	+	Population + Patient	Population + Patient
	Superior	+	+	+	+	Population + Patient	Population + Patient
	Abs diff/In-Su/	+	+	–	+	Population + Patient ^{*)}	Population + Patient ^{*)}
	Nasal	+	+	+	+	Population + Patient	Population + Patient
	Temporal	+	+	+	+	Population + Patient	Population + Patient
	Abs diff/Na-Te/	+	+	–	+	Population + Patient ^{*)}	Not recommended
Whole	Mean	+	+	+	+	Population + Patient	Population + Patient
	SD	+	+	+	+	Population + Patient	Population + Patient

^{*)} Need to be repeated on patient level. ROC, ICC, AI, and SI notations are only shown for worst eye.

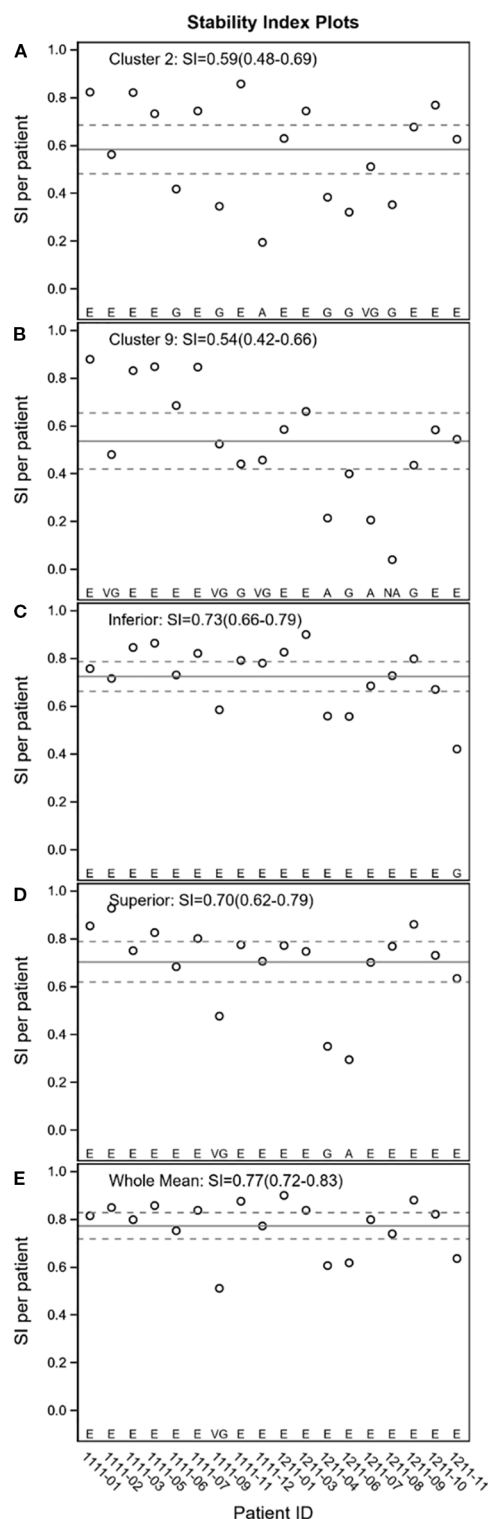


FIGURE 5
Stability Index (worst eye only) for cluster 2 and 9, inferior and superior hemifield, and for the whole VF.

were based on routine clinical judgement without prespecified case definition, introducing risk of misclassification. Third, the study design limits our ability to assess the performance to detect progression over time, which is critical for monitoring the slow

nature of GLA. Fourth, all GLA participants were using at least one form of topical anti-glaucoma medication at the time of testing, and the potential influence of such medications on SRT remains unexplored, however, none reported using systemic medications known to affect SRT (59–62). Fifth, the study did not include a direct comparison with SAP, the current clinical benchmark in functional testing in GLA, which will be important in future evaluations. Finally, participants' eyes were categorized as “worst” and “best”. This approach allowed for independent analysis of each eye, avoiding the need for complex statistical models, preventing the rejection of useful data, and reducing the required number of participants (63, 64). However, this approach also introduced greater variability in disease progression within the two groups, possibly obscuring clear trends.

Conclusion

The findings demonstrate that the majority of the SRT variables studied are not only effective in differentiating glaucomatous eyes from HC, but also exhibit a sufficient level of reliability and stability, which is essential for use in a clinical setting. Furthermore, 19 of the 22 BCAM VF test variables were identified as potential GLA-biomarkers according to pre-specified criteria.

Data availability statement

The datasets presented in this article are not readily available because the raw datasets generated for this study contain sensitive patient information and are subject to strict data privacy regulations, including those derived from the General Data Protection Regulation (GDPR) and relevant Norwegian legislation (e.g., Personopplysningsloven). Consequently, direct public sharing of the full dataset is not possible due to patient confidentiality. The data is stored securely at Meddoc AS, who also performed the analysis. While the raw data cannot be openly shared, reasonable requests for access to de-identified aggregate data or for verification of key findings will be considered by the corresponding author upon approval from the data custodians and in accordance with applicable data protection laws. Requests to access the datasets should be directed to sl@meddoc.no.

Ethics statement

The study was reviewed by the Regional Committees for Medical and Health Research Ethics (REK) nord in Norway, which, in its meeting on 18.08.2022, determined the project to be outside the scope of the Health Research Act and thus not subject to their approval mandate. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

AS: Conceptualization, Investigation, Resources, Writing – review & editing, Validation, Visualization, Data curation, Writing – original draft, Methodology. BH: Data curation, Methodology, Conceptualization, Writing – review & editing, Software. OK: Writing – review & editing, Resources, Methodology, Supervision, Conceptualization. MK: Supervision, Methodology, Conceptualization, Writing – review & editing. SL: Writing – review & editing, Data curation, Project administration, Visualization, Formal Analysis, Writing – original draft, Methodology, Conceptualization, Supervision. GP: Writing – review & editing, Supervision, Resources, Conceptualization, Project administration, Funding acquisition, Methodology.

Funding

The author(s) declare financial support was received for the research and/or publication of this article. This research was supported by institutional funds from Oslo University Hospital and The Research Council of Norway (Project No.: 333925 on “Validation of eye-tracking tests in screening, diagnostics and treatment monitoring in two major, central ophthalmic disorders”).

Conflict of interest

BH is currently employed as Chief Medical Officer at Bulbitech AS, and owns shares, as well as IP (patent: US20240335111A1 –

“Eye Testing Device”). GP serves on the Advisory Board of Bulbitech AS on a pro bono basis.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that Generative AI was used in the creation of this manuscript. Generative AI was used to enhance the fluency and clarity of the manuscript. It was not used for conceptualisation, data analysis, interpretation of results, or the generation of scientific ideas.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Weinreb RN, Khaw PT. Primary open-angle glaucoma. *Lancet*. (2004) 363:1711–20. doi: 10.1016/S0140-6736(04)16257-0
- Quigley HA, Addicks EM, Green WR. Optic nerve damage in human glaucoma. III. Quantitative correlation of nerve fibre loss and visual field defect in glaucoma, ischemic neuropathy, papilledema, and toxic neuropathy. *Arch Ophthalmol*. (1982) 100:135–46. doi: 10.1001/archophth.1982.01030030137016
- Jayaram H, Kolko M, Friedman DS, Gazzard G. Glaucoma: now and beyond. *Lancet*. (2023) 402:1788–801. doi: 10.1016/S0140-6736(23)01289-8
- Garway-Heath DF. Early diagnosis in glaucoma. *Prog Brain Res*. (2008) 173:47–57. doi: 10.1016/S0079-6123(08)01105-9
- Azuara-Blanco A, Traverso CE, Bagnasco L, Bagnis A, Breda JB, Bonzano C, et al. European glaucoma society terminology and guidelines for glaucoma, 5th edition. *Br J Ophthalmol*. (2021) 105(Suppl 1):1–169. doi: 10.1136/bjophthalmol-2021-egsguidelines
- Harwerth RS, Carter-Dawson L, Shen F, Smith EL 3rd, Crawford ML. Ganglion cell losses underlying visual field defects from experimental glaucoma. *Invest Ophthalmol Vis Sci*. (1999) 40:2242–50.
- Harwerth RS, Quigley HA. Visual field defects and retinal ganglion cell losses in patients with glaucoma. *Arch Ophthalmol*. (2006) 124:853–9. doi: 10.1001/archophth.124.6.853
- Kerrigan-Baumrind LA, Quigley HA, Pease ME, Kerrigan DF, Mitchell RS. Number of ganglion cells in glaucoma eyes compared with threshold visual field tests in the same persons. *Invest Ophthalmol Vis Sci*. (2000) 41:741–8.
- Gardiner SK. Differences in the relation between perimetric sensitivity and variability between locations across the visual field. *Invest Ophthalmol Vis Sci*. (2018) 59:3667–74. doi: 10.1167/iovs.18-24303
- Gardiner SK, Swanson WH, Mansberger SL. Long- and short-term variability of perimetry in glaucoma. *Transl Vis Sci Technol*. (2022) 11:3. doi: 10.1167/tvst.11.8.3
- Flammer J, Drance SM, Zulauf M. Differential light threshold. Short- and long-term fluctuation in patients with glaucoma, normal controls, and patients with suspected glaucoma. *Arch Ophthalmol*. (1984) 102:704–6. doi: 10.1001/archophth.1984.01040030560017
- Tahri Sqalli M, Aslonov B, Gafurov M, Mukhammadiev N, Sqalli Houssaini Y. Eye tracking technology in medical practice: a perspective on its diverse applications. *Front Med Technol*. (2023) 5:1253001. doi: 10.3389/fmedt.2023.1253001
- Pierrot-Deseilligny C, Rivaud S, Gaymard B, Müri R, Vermersch AI. Cortical control of saccades. *Ann Neurol*. (1995) 37:557–67. doi: 10.1002/ana.410370504
- Kanjee R, Yücel YH, Steinbach MJ, González EG, Gupta N. Delayed saccadic eye movements in glaucoma. *Eye Brain*. (2012) 4:63–8. doi: 10.2147/EB.S38467
- Mazumdar D, Meethal NSK, George R, Pel JJM. Saccadic reaction time in mirror image sectors across horizontal meridian in eye movement perimetry. *Sci Rep*. (2021) 11:2630. doi: 10.1038/s41598-021-81762-y
- Mazumdar D, Pel JJ, Panday M, Asokan R, Vijaya L, Shantha B, et al. Comparison of saccadic reaction time between normal and glaucoma using an eye movement perimeter. *Indian J Ophthalmol*. (2014) 62:55–9. doi: 10.4103/0301-4738.126182
- Mazumdar D, Pel JJM, Kadavath Meethal NS, Asokan R, Panday M, Steen JVD, et al. Visual field plots: A comparison study between standard automated perimetry and eye movement perimetry. *J Glaucoma*. (2020) 29:351–61. doi: 10.1097/IJG.0000000000001477
- Meethal NSK, Pel JJM, Mazumdar D, Asokan R, Panday M, van der Steen J, et al. Eye Movement Perimetry and Frequency Doubling Perimetry: clinical performance and patient preference during glaucoma screening. *Graefes Arch Clin Exp Ophthalmol*. (2019) 257:1277–87. doi: 10.1007/s00417-019-04311-4
- Tatham AJ, Murray IC, McTrusty AD, Cameron LA, Perperidis A, Brash HM, et al. Speed and accuracy of saccades in patients with glaucoma evaluated using an eye tracking perimeter. *BMC Ophthalmol*. (2020) 20:259. doi: 10.1186/s12886-020-01528-4

20. Yeon JS, Jung HN, Kim JY, Jung KI, Park HL, Park CK, et al. Deviated saccadic trajectory as a biometric signature of glaucoma. *Transl Vis Sci Technol.* (2023) 12:15. doi: 10.1167/tvst.12.7.15
21. Ballal Ganeshrao S, Jaleel A, Madicharla S, Kavya V Sri, Zakir J, Garudadri CS, et al. Comparison of saccadic eye movements among the high-tension glaucoma, primary angle-closure glaucoma, and normal-tension glaucoma. *J Glaucoma.* (2021) 30:e76–82. doi: 10.1097/IJG.0000000000001757
22. Mazumdar D, Meethal NSK, Panday M, Asokan R, Thepass G, George RJ, et al. Effect of age, sex, stimulus intensity, and eccentricity on saccadic reaction time in eye movement perimetry. *Transl Vis Sci Technol.* (2019) 8:13. doi: 10.1167/tvst.8.4.13
23. Meethal NSK, Mazumdar D, Thepass G, Lemij HG, van der Steen J, Pel JJM, et al. Effect of ethnic diversity on the saccadic reaction time among healthy Indian and Dutch adults. *Sci Rep.* (2024) 14:551. doi: 10.1038/s41598-023-50670-8
24. Thepass G, Pel JJ, Vermeer KA, Creten O, Bryan SR, Lemij HG, et al. The effect of cataract on eye movement perimetry. *J Ophthalmol.* (2015) 2015:425067. doi: 10.1155/2015/425067
25. Thepass G, Lemij HG, Vermeer KA, van der Steen J, Pel JJM. Slowed saccadic reaction times in seemingly normal parts of glaucomatous visual fields. *Front Med (Lausanne).* (2021) 8:679297. doi: 10.3389/fmed.2021.679297
26. Lamirel C, Milea D, Cochereau I, Duong MH, Lorenceau J. Impaired saccadic eye movement in primary open-angle glaucoma. *J Glaucoma.* (2014) 23:23–32. doi: 10.1097/IJG.0b013e31825c10dc
27. Elgin CY. Eye-tracking algorithm for early glaucoma detection: analysis of saccadic eye movements in primary open-angle glaucoma. *J Eye Movement Res.* (2025) 18:18. doi: 10.3390/jemr18030018
28. Pel JJ, van Beijsterveld MC, Thepass G, van der Steen J. Validity and repeatability of saccadic response times across the visual field in eye movement perimetry. *Transl Vis Sci Technol.* (2013) 2:3. doi: 10.1167/tvst.2.7.3
29. Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hröbjartsson A, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *J Clin Epidemiol.* (2011) 64:96–106. doi: 10.1016/j.jclinepi.2010.03.002
30. Aronson JK, Ferner RE. Biomarkers-A general review. *Curr Protoc Pharmacol.* (2017) 76:9. doi: 10.1002/cpph.19
31. Wang M, Shen LQ, Boland MV, Wellik SR, De Moraes CG, Myers JS, et al. Impact of natural blind spot location on perimetry. *Sci Rep.* (2017) 7:6143. doi: 10.1038/s41598-017-06580-7
32. Racette L, Fischer M, Bebie H, Holló G, Johnson CA, Matsumoto C. Visual field digest A guide to perimetry and the octopus perimeter. *Octopus.* (2019) 8:112.
33. Aarås A, Veierød MB, Larsen S, Ortgren R, Ro O. Reproducibility and stability of normalized EMG measurements on musculus trapezius. *Ergonomics.* (1996) 39:171–85. doi: 10.1080/00140139608964449
34. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* (2016) 15:155–63. doi: 10.1016/j.jcm.2016.02.012
35. Dalbro SEJ, Elsaïs A, Rydning SL, Toft M, Kerty E, Larsen SE. Repeatability, reliability, and stability of eye movement measurements in Parkinson's disease, cerebellar ataxia, and healthy adults. *Front Neurol.* (2025) 16:1556314. doi: 10.3389/fneur.2025.1556314
36. Kadavath Meethal NS, Mazumdar D, Asokan R, Panday M, van der Steen J, Vermeer KA, et al. Development of a test grid using Eye Movement Perimetry for screening glaucomatous visual field defects. *Graefes Arch Clin Exp Ophthalmol.* (2018) 256:371–9. doi: 10.1007/s00417-017-3872-x
37. Heijl A, Lindgren A, Lindgren G. Test-retest variability in glaucomatous visual fields. *Am J Ophthalmol.* (1989) 108:130–5. doi: 10.1016/0002-9394(89)90006-8
38. Swanson WH, Horner DG, Dul MW, Malinovsky VE. Choice of stimulus range and size can reduce test-retest variability in glaucomatous visual field defects. *Transl Vis Sci Technol.* (2014) 3:6. doi: 10.1167/tvst.3.5.6
39. Wall M, Woodward KR, Doyle CK, Artes PH. Repeatability of automated perimetry: a comparison between standard automated perimetry with stimulus size III and V, matrix, and motion perimetry. *Invest Ophthalmol Vis Sci.* (2009) 50:974–9. doi: 10.1167/iovs.08-1789
40. Chauhan BC, Johnson CA. Test-retest variability of frequency-doubling perimetry and conventional perimetry in glaucoma patients and normal subjects. *Invest Ophthalmol Vis Sci.* (1999) 40:648–56.
41. Summer P. *The oxford handbook of eye movements.* Oxford: Oxford University Press (2011). p. 413–24.
42. Carpenter RH, Williams ML. Neural computation of log likelihood in control of saccadic eye movements. *Nature.* (1995) 377:59–62. doi: 10.1038/377059a0
43. Jóhannesson ÓI, Tagu J, Kristjánsson Á. Asymmetries of the visual system and their influence on visual performance and oculomotor dynamics. *Eur J Neurosci.* (2018) 48:3426–45. doi: 10.1111/ejn.14225
44. Gardiner SK, Mansberger SL, Demirel S. Detection of functional change using cluster trend analysis in glaucoma. *Invest Ophthalmol Vis Sci.* (2017) 58:Bio180–bio90. doi: 10.1167/iovs.17-21562
45. Werner EB, Beraskow J. Peripheral nasal field defects in glaucoma. *Ophthalmology.* (1979) 86:1875–8. doi: 10.1016/s0161-6420(79)35335-0
46. Lewis RA, Phelps CD. A comparison of visual field loss in primary open-angle glaucoma and the secondary glaucomas. *Ophthalmologica.* (1984) 89:41–8. doi: 10.1159/000309383
47. Kim JM, Kyung H, Shim SH, Azarbad P, Caprioli J. Location of initial visual field defects in glaucoma and their modes of deterioration. *Invest Ophthalmol Vis Sci.* (2015) 56:7956–62. doi: 10.1167/iovs.15-17297
48. de Paula A, Perdicchi A, Pocobelli A, Fragiotta S, Scuderi G. The “Topography” of glaucomatous defect using OCT and visual field examination. *J Curr Glaucoma Pract.* (2022) 16:31–5. doi: 10.5005/jp-journals-10078-1353
49. Vandersnickt MF, van Eijgen J, Lemmens S, Stalmans I, Pinto LA, Vandewalle EM. Visual field patterns in glaucoma: A systematic review. *Saudi J Ophthalmol.* (2024) 38:306–15. doi: 10.4103/sjopt.sjopt_143_24
50. Rucker JC, Hudson T, Rizzo JR. Translational neurology of slow saccades. In: Shaikh A, Ghasia F, editors. *Advances in translational neuroscience of eye movement disorders.* Springer International Publishing, Cham (2019). p. 221–54.
51. Lavanya R, Riyazuddin M, Dasari S, Puttaiah NK, Venugopal JP, Pradhan ZS, et al. A comparison of the visual field parameters of SITA faster and SITA standard strategies in glaucoma. *J Glaucoma.* (2020) 29:783–8. doi: 10.1097/IJG.0000000000001551
52. King AJ, Taguri A, Wadood AC, Azuara-Blanco A. Comparison of two fast strategies for the assessment of visual fields in glaucoma patients. *Graefes Arch Clin Exp Ophthalmol.* (2002) 240:481–7. doi: 10.1007/s00417-002-0482-y
53. Toepfer A, Kasten E, Guenther T, Sabel BA. Perimetry while moving the eyes: implications for the variability of visual field defects. *J Neuroophthalmol.* (2008) 28:308–19. doi: 10.1097/WNO.0b013e31818e3cd7
54. Bonneh YS, Donner TH, Cooperman A, Heeger DJ, Sagi D. Motion-induced blindness and Troxler fading: common and different mechanisms. *PLoS One.* (2014) 9:e92894. doi: 10.1371/journal.pone.0092894
55. Fuhr PS, Hershner TA, Daum KM. Ganzfeld blackout occurs in bowl perimetry and is eliminated by translucent occlusion. *Arch Ophthalmol.* (1990) 108:983–8. doi: 10.1001/archophth.1990.01070090085045
56. Trope GE, Eizenman M, Coyle E. Eye movement perimetry in glaucoma. *Can J Ophthalmol.* (1989) 24:197–9.
57. Rao HL, Yadav RK, Begum VU, Addepalli UK, Choudhari NS, Senthil S, et al. Role of visual field reliability indices in ruling out glaucoma. *JAMA Ophthalmol.* (2015) 133:40–4. doi: 10.1001/jamaophth.2014.3609
58. Ishiyama Y, Murata H, Asaoka R. The usefulness of gaze tracking as an index of visual field reliability in glaucoma patients. *Invest Ophthalmol Vis Sci.* (2015) 56:6233–6. doi: 10.1167/iovs.15-17661
59. Reilly JL, Lencer R, Bishop JR, Keedy S, Sweeney JA. Pharmacological treatment effects on eye movement control. *Brain Cogn.* (2008) 68:415–35. doi: 10.1016/j.bandc.2008.08.026
60. Klein C C.OMMAJ.R.X.X.X, Fischer B, Fischer B, Hartnegg K. Effects of methylphenidate on saccadic responses in patients with ADHD. *Exp Brain Res.* (2002) 145:121–5. doi: 10.1007/s00221-002-1105-x
61. Naicker P, Anoopkumar-Dukie S, Grant GD, Modenese L, Kavanagh JJ. Medications influencing central cholinergic pathways affect fixation stability, saccadic response time and associated eye movement dynamics during a temporally-cued visual reaction time task. *Psychopharmacol (Berl).* (2017) 234:671–80. doi: 10.1007/s00213-016-4507-3
62. Fafrowicz M, Unrug A, Marek T, van Luijckelaar G, Noworol C, Coenen A. Effects of diazepam and buspirone on reaction time of saccadic eye movements. *Neuropsychobiology.* (1995) 32:156–60. doi: 10.1159/000119316
63. Armstrong RA. Statistical guidelines for the analysis of data obtained from one or both eyes. *Ophthalmic Physiol Opt.* (2013) 33:7–14. doi: 10.1111/opo.12009
64. Banerjee K, Pramanik S, Mondal LK. Statistical methods for best and worst eye measurements. *Clin Ophthalmol.* (2024) 18:1901–8. doi: 10.2147/OPTH.S461511